

# Structural Characterization of an Equilibrium Unfolding Intermediate in Cytochrome *c*

Ramil F. Latypov<sup>1</sup>, Hong Cheng<sup>1</sup>, Navid A. Roder<sup>1</sup>, Jiaru Zhang<sup>1</sup>  
and Heinrich Roder<sup>1,2\*</sup>

<sup>1</sup>Basic Science Division, Fox Chase Cancer Center, 333 Cottman Avenue, Philadelphia PA 19111, USA

<sup>2</sup>Department of Biochemistry and Biophysics, University of Pennsylvania, Philadelphia PA 19104, USA

Although the denaturant-induced unfolding transition of cytochrome *c* was initially thought to be a cooperative process, recent spectroscopic studies have shown deviations from two-state behavior consistent with accumulation of an equilibrium intermediate. However, little is known about the structural and thermodynamic properties of this state, and whether it is stabilized by the presence of non-native heme ligands. We monitored the reversible denaturant-induced unfolding equilibrium of oxidized horse cytochrome *c* using various spectroscopic probes, including fluorescence, near and far-UV CD, heme absorbance bands in the Soret, visible and near-IR regions of the spectrum, as well as 2D NMR. Global fitting techniques were used for a quantitative interpretation of the results in terms of a three-state model, which enabled us to determine the intrinsic spectroscopic properties of the intermediate. A well-populated intermediate was observed in equilibrium experiments at pH 5 using either guanidine-HCl or urea as a denaturant, both for wild-type cytochrome *c* as well as an H33N mutant chosen to prevent formation of non-native His-heme ligation. For a more detailed structural characterization of the intermediate, we used 2D <sup>1</sup>H-<sup>15</sup>N correlation spectroscopy to follow the changes in peak intensity for individual backbone amide groups. The equilibrium state observed in our optical and NMR studies contains many native-like structural features, including a well-structured  $\alpha$ -helical subdomain, a short Trp59-heme distance and solvent-shielded heme environment, but lacks the native Met80 sulfur-iron linkage and shows major perturbations in side-chain packing and other tertiary interactions. These structural properties are reminiscent of the A-state of cytochrome *c*, a compact denatured form found under acidic high-salt conditions, as well as a kinetic intermediate populated at a late stage of folding. The denaturant-induced intermediate also resembles alkaline forms of cytochrome *c* with altered heme ligation, suggesting that disruption of the native methionine ligand favors accumulation of structurally analogous states both in the presence and absence of non-native ligands.

© 2006 Elsevier Ltd. All rights reserved.

\*Corresponding author

Keywords: protein folding; denaturation; stability; spectroscopy; NMR

Present address: R. F. Latypov, Amgen, Thousand Oaks, CA 91320, USA.

Abbreviations used: cyt *c*, cytochrome *c*; GuHCl, guanidine-HCl; CD, circular dichroism; HSQC, heteronuclear single-quantum correlation; NOE, nuclear Overhauser effect; TOC, total correlation; WT, wild-type.

E-mail address of the corresponding author: [h\\_roder@fccc.edu](mailto:h_roder@fccc.edu)

## Introduction

Small globular proteins often exhibit cooperative unfolding transitions in which only folded (native) and unfolded (denatured) molecules are populated at equilibrium.<sup>1–3</sup> This two-state behavior is thought to be a consequence of the fact that the native structure is stabilized by a large number of weak interactions, which are cooperative in nature, so that partially folded states are inherently unstable. Even when partially structured states

accumulate as transient intermediates in kinetic experiments, they are generally not stable enough to be observable at equilibrium under strongly denaturing conditions. On the other hand, non-native states with spectroscopic and hydrodynamic properties distinct from those of the fully native and unfolded states (the so-called molten-globule state) often accumulate under mildly denaturing conditions, such as acidic or basic pH.<sup>4–6</sup> In some cases, deviations from simple two-state behavior can also be observed in denaturant-induced or thermal unfolding equilibria.<sup>7</sup> Although a stable native state that is energetically well separated from denatured states is expected to favor rapid folding into a unique structure,<sup>8,9</sup> there is growing evidence that kinetic intermediates are very common, even in small, single-domain proteins.<sup>10–14</sup> Thus, conformational states distinct from the fully native and unfolded populations are readily accessible and may result in non-cooperative unfolding transitions.

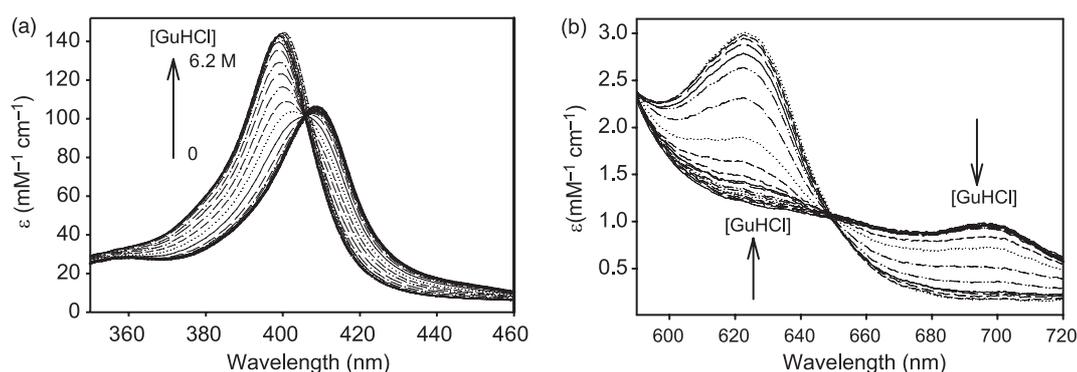
Cytochrome *c* (cyt *c*) has long served as a model protein for developing new concepts and approaches in protein folding.<sup>15–25</sup> The presence of a heme group, covalently linked to the polypeptide chain *via* thioester linkages to Cys14 and Cys17, provides a sensitive spectroscopic probe. While one of the axial heme ligands, His18, is maintained even under denaturing conditions, the second one, Met80, is inherently labile and readily displaced by other side-chains, such as a deprotonated His or Lys,<sup>26</sup> which can become trapped during folding.<sup>21</sup> Early solvent and thermal denaturation studies suggested that cyt *c*, like most other single-domain proteins in its size class (104 residues), exhibits two-state equilibrium behavior.<sup>27–29</sup> More recent spectroscopic and calorimetric studies revealed deviations from a fully cooperative two-state transition for horse cyt *c* at denaturant concentrations below those leading to major unfolding,<sup>30–32</sup> and a small-angle X-ray scattering study found evidence for two or more distinct populations of denatured molecules at higher denaturant concentrations.<sup>33</sup> Mayne & Englander<sup>34</sup> further predicted that low levels of equilibrium intermediates accumulate in the unfolding transition region of horse cyt *c* on the basis of earlier native-state hydrogen exchange data.<sup>22</sup> Most previous work was done at or near neutral pH, conditions where non-native ligand interactions between deprotonated histidine or lysine side-chains with the heme iron can occur.<sup>21,35–37</sup> In fact, Russel & Bren<sup>30,31</sup> reported complex changes in the <sup>1</sup>H NMR spectra of horse cyt *c* indicating that lysine residues (possibly those at positions 72, 73 or 79) can displace the native Met80 ligand below the main guanidine–HCl (GuHCl) or urea-induced unfolding transition while His33 is the predominant ligand at higher denaturant concentrations.<sup>38</sup>

Although evidence for accumulation of equilibrium intermediates in cyt *c* is compelling, little is known about the structural and thermodynamic properties of these states. Moreover, it is unclear

whether accumulation of equilibrium intermediates is linked with the formation of non-native heme ligands, or whether they are inherent structural intermediates. To gain further insight into the structure and stability of non-native equilibrium states, we monitored the reversible denaturant-induced unfolding transition of oxidized horse cyt *c* using a variety of spectroscopic probes, including fluorescence, near and far-UV CD, heme absorbance bands in the Soret, visible and near-IR regions of the spectrum, as well as 2D NMR. We have made extensive use of global fitting techniques for a quantitative interpretation of the results in terms of a three-state equilibrium model, which has made it possible to extract the heme absorbance spectrum of the intermediate. To determine whether heme misligation plays a role in stabilizing intermediates, we prepared a recombinant horse cyt *c* variant in which the predominant non-native ligand, His33, is replaced by Asn,<sup>38,39</sup> and compared the equilibrium unfolding behavior of this H33N variant with those of the wild-type protein at mildly acidic conditions (pH 5). NMR studies as a function of urea concentration, using 2D <sup>1</sup>H–<sup>15</sup>N heteronuclear single-quantum correlation (HSQC) spectroscopy to follow the changes in peak intensity for individual backbone amide groups lead to a detailed structural description of the pre-transition intermediate in the H33N variant.

## Results

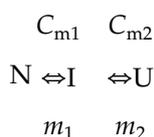
The absorbance spectrum of cyt *c* is highly sensitive to the heme environment. The heme gives rise to bands in the Soret (350–450 nm), visible (500–600 nm) and near-IR (600–750 nm) regions that vary substantially with the oxidation and ligation state of the iron, as well as the hydration and polarity of the heme environment within the protein.<sup>40,41</sup> Heme absorbance is, thus, a particularly sensitive probe for detecting conformational changes associated with folding intermediates. We followed the unfolding equilibrium of WT and H33N cyt *c* by recording absorbance spectra over a wide spectral range (typically 250 nm to 800 nm) as a function of denaturant concentration (GuHCl and/or urea; see Materials and Methods). Figure 1 shows the effect of GuHCl concentration on the spectra for H33N cyt *c* in the Soret (a) and near-IR (b) regions. As in the case of WT cyt *c*, native H33N cyt *c* shows a prominent Soret band with absorbance maximum at 409 nm ( $\epsilon = 106 \text{ mM}^{-1} \text{ cm}^{-1}$ ), which moves to 399 nm and increases in intensity ( $\epsilon = 143 \text{ mM}^{-1} \text{ cm}^{-1}$ ) upon unfolding. The lack of a clear isosbestic point as a function of GuHCl concentration indicates that unfolding is not a simple two-state process. However, the Soret band of the fully unfolded state continues to move to higher wavelength with increasing GuHCl concentration (> 4 M), which makes it difficult to extract the spectral contributions of any intermediates populated near the unfolding transition region. The spectral changes



**Figure 1.** Absorbance changes of the H33N variant of cyt *c* (0.1 M sodium acetate buffer (pH 5.0) at 15 °C) associated with the GuHCl-induced unfolding transition. (a) Absorbance spectra in the Soret region recorded at a protein concentration of 5  $\mu$ M using a 1 cm path length cuvette. (b) Absorbance spectra in the near-IR region, including a band at 625 nm characteristic of high-spin (penta-coordinated) heme and a weak band at 695 nm associated with the Met/His-ligated native state, measured on a 80  $\mu$ M protein solution in a 1 cm cuvette.

within the unfolded baseline region, together with the fact that there is no simple theoretical model for describing electronic spectra, also impede the use of singular value decomposition (SVD) methods.

In contrast, it is relatively straightforward to model the denaturant dependence of unfolding equilibria, even if intermediates are populated, and we were able to overcome these problems simply by replotting the data in Figure 1, choosing denaturant concentration as an independent variable rather than wavelength. Figure 2(a) shows a series of unfolding transition curves for H33N cyt *c* obtained by plotting the extinction coefficient at selected wavelengths spanning the Soret region as a function of GuHCl concentration. The transition curves at individual wavelengths can be approximated reasonably well by a two-state model (not shown). However, the midpoint concentrations,  $C_m$ , and slopes ( $m$ -values) vary significantly with the wavelength, which is a clear indication that a two-state model is inadequate to describe the data. In addition, closer inspection of the transition curves in the 405 nm region (Figure 2(b)) reveals significant deviations from the symmetric sigmoidal behavior expected for a two-state transition. On the other hand, a three-state model (Scheme 1)



(Scheme 1)

consisting of coupled equilibria between native (N), intermediate (I) and unfolded (U) species, respectively, can fully account for these complexities. Each of the two coupled transitions are described in terms of a midpoint concentration,  $C_{mi}$ , and slope,  $m_i$ , assuming that the free energy for each transition,  $\Delta G_i$ , varies linearly with

denaturant concentration,  $c$ :

$$\Delta G_i(c) = \Delta G_i(0) - m_i c \quad (1)$$

Since  $\Delta G_i(C_{mi}) = 0$ , equation (1) is equivalent to:

$$\Delta G_i(c) = m_i (C_{mi} - c) \quad (2)$$

The fractional populations of the three states in Scheme 1,  $f_N$ ,  $f_I$  and  $f_U$ , can be expressed as:

$$f_N = 1/(1 + K_{NI} + K_{NI}K_{IU}) \quad (3)$$

$$f_I = K_{NI}f_N$$

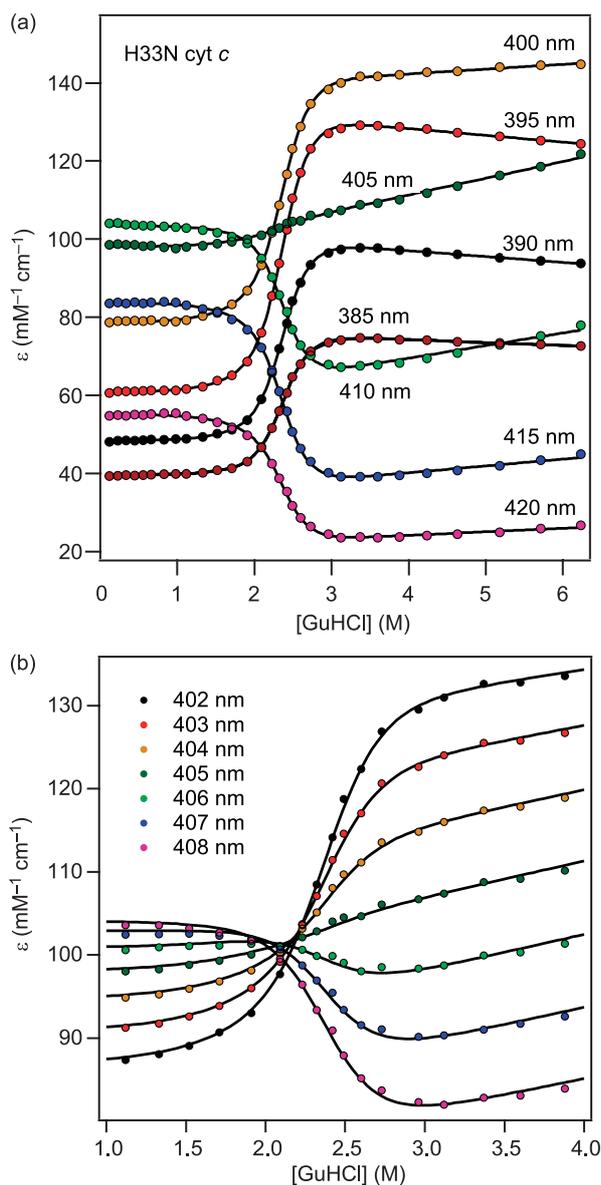
$$f_U = K_{IU}f_I$$

where  $K_{NI}$  and  $K_{IU}$  are the equilibrium constants for the two transitions given by  $K_i = \exp(-\Delta G_i/RT)$ . Thus, the changes in extinction coefficient associated with denaturant-induced unfolding transition,  $\varepsilon(c)$ , are given by:

$$\varepsilon(c) = \varepsilon_N f_N + \varepsilon_I f_I + \varepsilon_U f_U \quad (4)$$

where  $\varepsilon_N$ ,  $\varepsilon_I$  and  $\varepsilon_U$  are the extinction coefficients of the N, I and U-state, respectively, which reflect the intrinsic absorbance spectra of the three states. In general, the absorbance properties of each state may vary with denaturant concentration, especially if the chromophore is exposed to the solvent. However, we found that excellent fits of the cyt *c* unfolding transitions monitored by heme absorbance could be obtained with the following simplifying assumptions: (a)  $\varepsilon_N$  and  $\varepsilon_I$  are independent of denaturant concentrations (consistent with the solvent-shielded environment of the heme group in compact forms of cyt *c*); (b)  $\varepsilon_U$  varies linearly with denaturant concentration, i.e.  $\varepsilon_U(c) = \varepsilon_U^0 + s_U c$ , where  $s_U$  is the slope of the unfolded baseline versus denaturant concentration.

The continuous lines in Figure 2 were obtained by simultaneously fitting a three-state equilibrium to the combined data set consisting of 61 transition curves ( $\varepsilon$  versus [GuHCl]) at 1 nm increments over



**Figure 2.** Absorbance-detected unfolding transition of H33N cyt *c* at selected wavelengths *versus* GuHCl concentration. The lines represent the best global fit of a three-state equilibrium model (Scheme 1) to the combined absorbance data. (a) Molar extinction at selected wavelengths in the Soret region (from Figure 1(a)). (b) Expanded plots of the transition curves monitored between 402 nm and 408 nm showing wavelength-dependent changes inconsistent with a two-state behavior.

the range from 370 nm to 440 nm). The four equilibrium parameters,  $C_{m1}$ ,  $m_1$ ,  $C_{m2}$  and  $m_2$ , were used as global fitting parameters. In addition, the four spectral parameters ( $\epsilon_N$ ,  $\epsilon_I$ ,  $\epsilon_U^0$  and  $\epsilon_U$ ) were used as local parameters to describe the transition curves at each wavelength. Note that the absorbance changes at low GuHCl concentrations are fully accounted for by the accumulation of an intermediate in the pre-transition region, and there is no need to introduce additional parameters for any denaturant-dependence of  $\epsilon_N$  and  $\epsilon_I$ . The optimized fitting parameters obtained by using the global fitting routine of the IGOR Pro software package (WaveMetrics, Inc., Lake Oswego, OR 97035) are listed in Table 1.

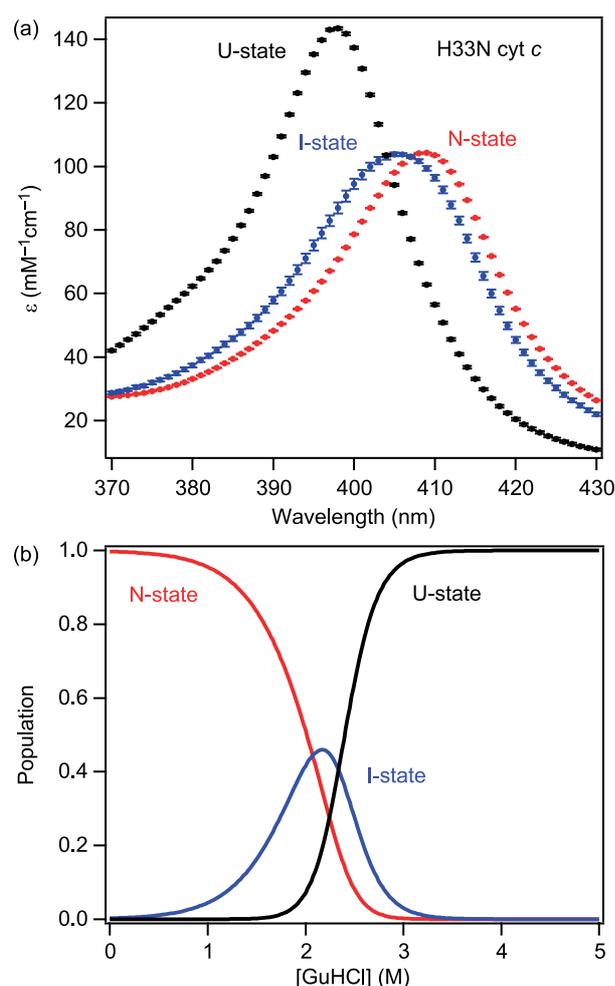
Figure 3(a) shows a plot of the optimized local parameters,  $\epsilon_N$ ,  $\epsilon_I$  and  $\epsilon_U^0$ , as a function of wavelength,  $\lambda$ . As expected,  $\epsilon_N(\lambda)$  is an exact representation of the Soret spectrum of native cyt *c* (cf Figure 1).  $\epsilon_U^0$  is the intercept of a linear fit to the unfolded baseline, and its dependence on wavelength,  $\epsilon_U(\lambda)$ , is similar in shape and position to the spectrum of the unfolded protein at high GuHCl concentrations (Figure 1). However, the absorbance maximum occurs at a slightly lower wavelength (398 nm, compared to 399 nm in 4.2 M GuHCl), since it represents the spectrum of the unfolded state in the absence of denaturant. The spectrum of the I-state,  $\epsilon_I(\lambda)$ , is similar in magnitude and shape to that of the native state, but its absorbance maximum lies at 405 nm, corresponding to a 4 nm blue shift relative to the N-state. These differences are far outside errors as indicated by the error bars in Figure 3, which are based on the covariance of the individual local fitting parameters. Thus, our global fitting procedure is a robust and straightforward approach for extracting the spectral contributions of intermediate states in multi-state folding equilibria. At the same time, the four global fitting parameters provide an accurate measure of the denaturant-dependence of the individual equilibrium transitions, and thus define the populations of the three states as a function of denaturant concentration (Figure 3(b)). The I-state reaches a maximum population of 45% near 2 M GuHCl.

To obtain further structural insight, we followed the GuHCl-induced unfolding transition of H33N cyt *c* using several other spectroscopic parameters, including tryptophan fluorescence, near and far-UV CD, aromatic absorbance and heme absorbance bands at 625 and 695 nm. Figure 4 compares the changes in ellipticity at 225 nm, which reflects  $\alpha$ -helix content, with the changes in the near-UV

**Table 1.** Thermodynamic parameters by global fitting of a three-state model to the GuHCl and urea-induced unfolding transitions of cyt *c* monitored by heme absorbance, near and far-UV CD and fluorescence

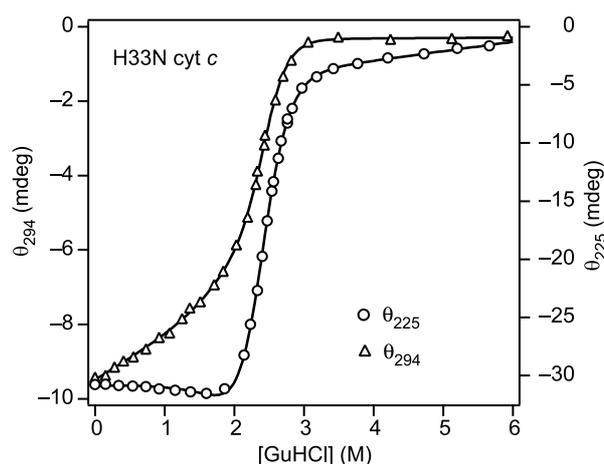
Protein/ denaturant	$C_{m1}$ (M)	$m_1$ (kcal $\text{mol}^{-1}\text{M}^{-1}$ )	$\Delta G_1$ (kcal $\text{mol}^{-1}$ )	$C_{m2}$ (M)	$m_2$ (kcal $\text{mol}^{-1}\text{M}^{-1}$ )	$\Delta G_2$ (kcal $\text{mol}^{-1}$ )	$\Delta G_{\text{tot}}$ (kcal $\text{mol}^{-1}$ )	$m_{\text{tot}}$ (kcal $\text{mol}^{-1}\text{M}^{-1}$ )
H33N/GuHCl	$2.07 \pm 0.06$	$1.7 \pm 0.08$	$3.5 \pm 0.2$	$2.34 \pm 0.01$	$3.08 \pm 0.06$	$7.3 \pm 0.15$	$10.8 \pm 0.35$	$4.8 \pm 0.14$
H33N/urea	$4.0 \pm 0.2$	$0.59 \pm 0.06$	$2.4 \pm 0.3$	$4.81 \pm 0.03$	$1.84 \pm 0.02$	$8.9 \pm 0.1$	$10.5 \pm 0.4$	$2.43 \pm 0.08$
WT/GuHCl	$1.50 \pm 0.05$	$1.34 \pm 0.11$	$2.0 \pm 0.2$	$2.55 \pm 0.01$	$3.05 \pm 0.03$	$7.8 \pm 0.1$	$9.8 \pm 0.3$	$4.39 \pm 0.14$

Measured at 15 °C in 0.1 M sodium acetate buffer (pH 5.0).



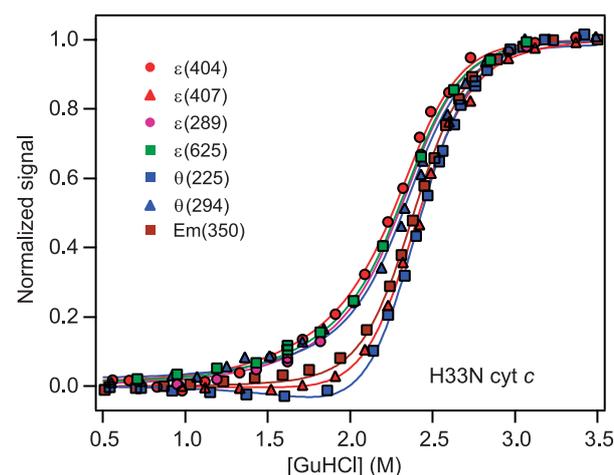
**Figure 3.** Intrinsic heme absorbance spectra (a) and populations (b) of the three predominant states populated in the GuHCl-induced unfolding equilibrium of H33N cyt *c*. The symbols in (a) represent the extinction coefficients in the absence of denaturant obtained by global fitting of the data shown in Figures 1 and 2. Error bars indicate the co-variance of (local) fitting parameters. The populations in (b) were calculated using the global fitting parameters ( $C_{m1}$ ,  $m_1$ ,  $C_{m2}$ ,  $m_2$ ) listed in Table 1.

region due to aromatic side-chains. Under native conditions, this region of the CD spectrum features two sharp negative bands at 287 nm and 294 nm attributed to Trp59,<sup>42</sup> which begin to decrease in intensity already between 1 and 2 M GuHCl. On the other hand, the far-UV signal shows major changes indicative of helix loss only during the second transition above 2 M GuHCl. In fact, the signal at 225 nm becomes slightly more negative from 0 to 2 M GuHCl, suggesting a small increase in helix content prior to main unfolding transition. In order to detect changes in heme coordination associated with unfolding, we monitored the GuHCl-induced changes in the absorbance band centered at 625 nm, which is indicative of a non-native high-spin state (i.e. the sixth heme coordination site is unoccupied or contains a weak ligand, such as a hydroxide ion), and the weak



**Figure 4.** GuHCl-induced unfolding equilibrium of H33N cyt *c* at pH 5.0, 15 °C monitored *via* the changes in the far-UV CD signal at 225 nm (right scale, 2 mm path length, 13  $\mu\text{M}$  protein) and the near-UV CD band at 294 nm (left scale, 1 cm path length, 40  $\mu\text{M}$  protein). The lines represent a three-state unfolding transition using the global fit parameters listed in Table 1. The intrinsic CD signal of the N and I-states was assumed to be independent of GuHCl concentration.

charge-transfer band at 695 nm, which correlates with the presence of a native Met80 sulfur-iron bond.<sup>41,43</sup> Finally, we measured the GuHCl-induced changes in tryptophan fluorescence, which are dominated by resonant energy transfer to the heme (a highly efficient fluorescence quencher), and thus provide



**Figure 5.** Normalized transition curves for GuHCl-induced unfolding of H33N cyt *c* monitored by absorbance ( $\epsilon$ ), CD ( $\theta$ ) and tryptophan fluorescence (Em). The normalized signal  $n(c)$  versus denaturant concentration  $c$  at a given wavelength was calculated from the raw signal  $s(c)$  using the relation  $n(c) = [s(c) - s_N] / [s_U(c) - s_N]$ , where  $s_N$  is the fitted signal for the N-state (generally assumed to be independent of  $c$ ) and  $s_U(c)$  represents the GuHCl-dependent signal of the unfolded state defined by the fitted intercept and slope of the unfolded baseline.

**Table 2.** Spectroscopic properties of the equilibrium intermediate in wild-type and H33N cyt *c*

Protein/denaturant	$\lambda_{\max}$ (nm)	$\varepsilon(\lambda_{\max})$ ( $\text{mM}^{-1}\text{cm}^{-1}$ )	$f(\theta_{225})$	$f(\theta_{\text{arom}})$	$f(\text{Em}_{350})$	$f(A_{625})$	$\alpha_1 = m_1/m_{\text{tot}}$
H33N/GuHCl	405	104	-0.15	0.31	0.06	0.40	0.35
H33N/urea	408	110	0.04	0.28	-	-	0.24
WT/GuHCl	407	107	-0.12	0.42	0.04	-	0.31

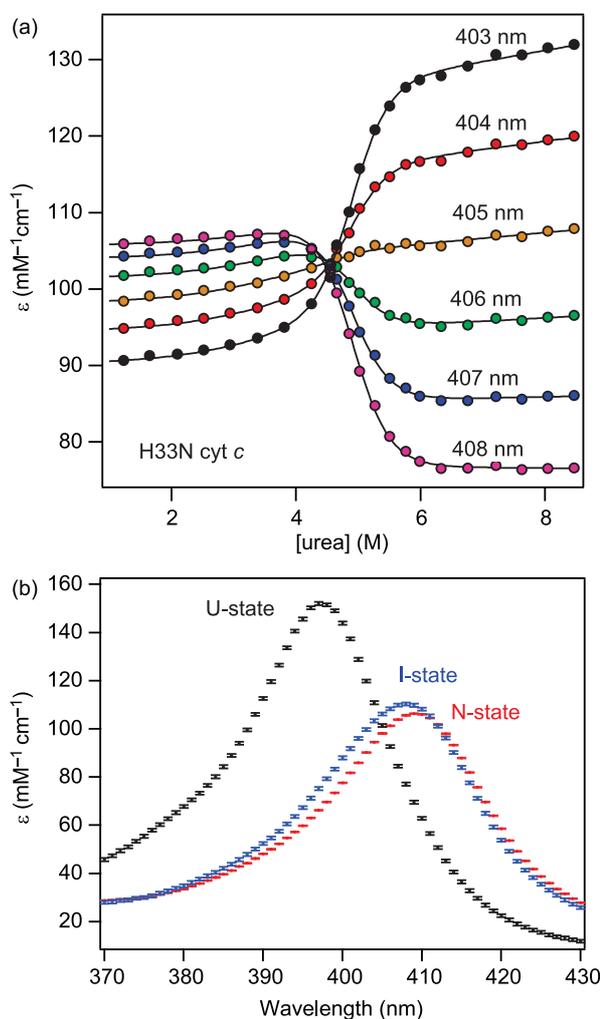
a measure of the average distance between Trp59 and the heme.<sup>44</sup> Figure 5 compares the unfolding transitions for the various spectroscopic parameters obtained by normalization (see Materials and Methods). To avoid any bias due to the choice of baseline values, we used the offsets and slopes of the native and unfolded baselines obtained in a global three-state fit of the raw data (see Figures 2 and 4) to normalize the measured transition curves.

The fact that the transition curves for different spectroscopic probes diverge between 1 M and 2.5 M GuHCl confirms our conclusion that an intermediate is populated at intermediate denaturant concentrations. The different signals show two types of behavior: (1) the Trp59 fluorescence, far-UV CD and Soret absorbance at 407 nm remain essentially at their native levels up to 2 M GuHCl followed by steep transitions centered at 2.34 M GuHCl; (2) the remaining probes, including the near-UV CD changes at 294 nm and the normalized absorbance changes at 289 nm, 404 nm and 625 nm begin to increase above 1 M GuHCl, reaching relative values of about 0.3 at 2 M GuHCl before joining the other probes (the 695 nm absorbance band shows similar behavior to that at 625 nm, but is less reliable due to its small extinction coefficient and large underlying baseline changes). The continuous curves in Figure 5 represent a global fit of a three-state unfolding equilibrium to the combined data, using the equilibrium parameters obtained in the analysis of the Soret data (Table 1). Table 2 summarizes the spectroscopic properties of the intermediate state in terms of the wavelength and extinction coefficient of the Soret band and the relative contributions of the other parameters.

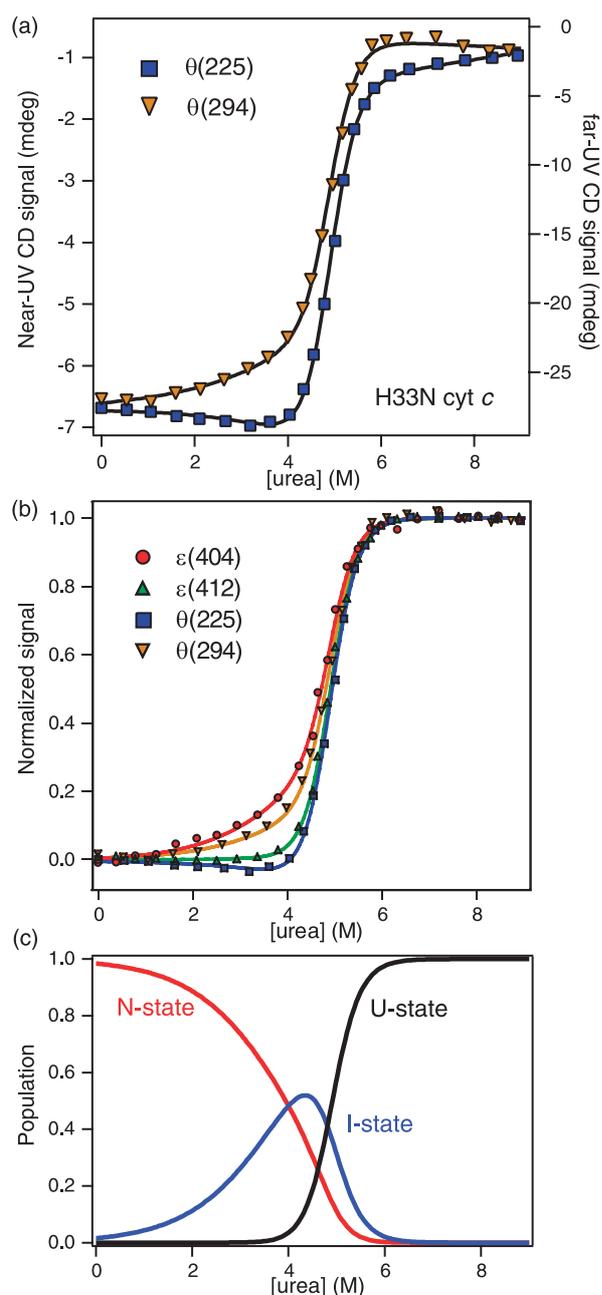
To rule out the possibility that specific binding of a guanidinium ion to the native protein is responsible for accumulation of the equilibrium intermediate, we carried out a parallel series of unfolding experiments using urea as a denaturant. As in the case of GuHCl, a detailed analysis of heme absorbance in the Soret region as a function of urea provides clear evidence for accumulation of an intermediate in the unfolding equilibrium of H33N cyt *c*. The wavelength-dependent variations in the shape of the unfolding transitions, which are particularly pronounced in the wavelength range between 404 nm and 408 nm (Figure 6(a)), are inconsistent with a two-state transition. The spectrum of the I-state obtained by global analysis of the Soret data (Figure 6(b)) shows again a blue shift relative to the N-state (Table 2). These spectral differences are well beyond errors and fully account for the complex shape of the individual unfolding

curves. For example, between 405 nm and 409 nm, the extinction coefficient of the I-state is higher than both N and U-states, giving rise to a local maximum in the unfolding curves near 4 M urea (Figure 6(a)).

In the pre-transition region (1.5 M–4.5 M urea), we also found major deviations between the unfolding curves for different regions of the CD spectrum (cf Figure 4). In particular, intermediate urea concentrations result in substantial changes in the aromatic region of the CD spectrum

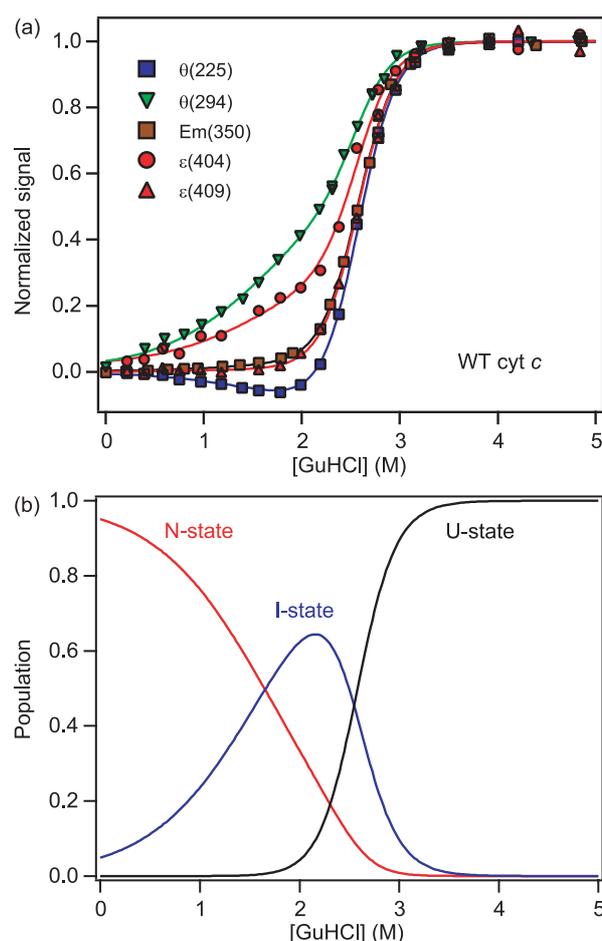


**Figure 6.** Urea-induced unfolding of H33N cyt *c* (pH 5.0 at 15 °C) as monitored by heme absorbance. (a) Unfolding curves at selected wavelengths in the Soret region (symbols), along with global three-state fits (lines). (b) Intrinsic Soret absorbance spectra for the native (red), intermediate (blue) and unfolded (black) states obtained by global analysis of the combined absorbance data versus urea concentration. Fitting errors are indicated by the height of the symbols.



**Figure 7.** Multi-parameter analysis of the three-state unfolding equilibrium of H33N cyt *c* versus urea concentration. (a) Raw unfolding transitions monitored in the near-UV (left scale) and far-UV (right scale) regions of the CD spectrum. (b) Urea dependence of the fractional changes in absorbance and CD at the wavelengths indicated. (c) Urea dependence of the populations for the N-state (red), I-state (blue) and U-state (black) calculated from the thermodynamic parameters listed in Table 1.

(Figure 7(a)), whereas the far-UV CD region shows only a small decrease below 4 M urea consistent with a slight increase in helix content prior to the main unfolding transition. Figure 7(b) compares the normalized transition curves (fractional signal change) for several different spectroscopic



**Figure 8.** Normalized unfolding curves for selected optical probes (a) and populations (b) obtained by global fitting of the GuHCl-unfolding equilibrium of WT cyt *c*.

parameters, including the 294 nm band in the near-UV region of the CD spectrum attributed to Trp59, the CD signal at 225 nm, and selected wavelengths in the Soret region. The curves in Figure 7(b) were obtained by constrained fitting of a three-state model, using the set of equilibrium parameters derived from the global fit of the heme absorbance data (Figure 6; Table 1). The fractional population of the I-state calculated for these global equilibrium parameters reaches a peak value of 50% at 4 M urea (Figure 7(c)).

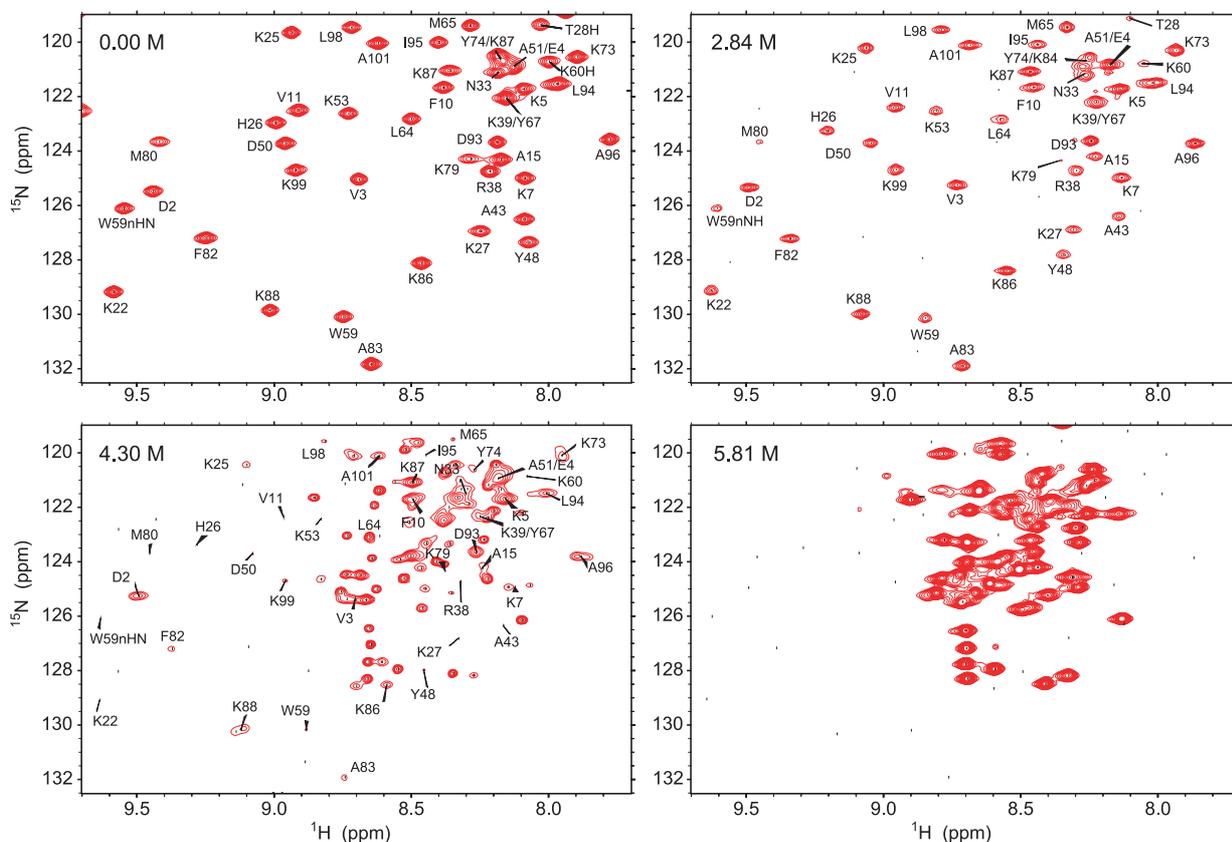
To determine whether the H33N mutation has any effect on the equilibrium properties of cyt *c*, we used the same combination of optical methods to follow the GuHCl-induced unfolding transition of wild-type horse cyt *c* (obtained from Sigma Co., St Louis, MO) under identical conditions (in 0.1 M acetate buffer (pH 5.0), 15 °C). The results of our global three-state analysis are summarized in Figure 8, and the set of equilibrium parameters consistent with the combined data is included in Table 1. As in the case of the H33N mutant (Figures 3 and 6), the intrinsic spectrum for the intermediate obtained by global fitting of absorbance data in

the Soret region shows a small, but significant blue-shift relative to that of the native state (Table 2). The normalized unfolding curves (Figure 8(a)) again show major variations among the different spectroscopic parameters in the pre-transition region (0.5 M to 2.5 M GuHCl). At 2 M GuHCl, the equilibrium intermediate reaches a maximum population of about 70% (Figure 8(b)).

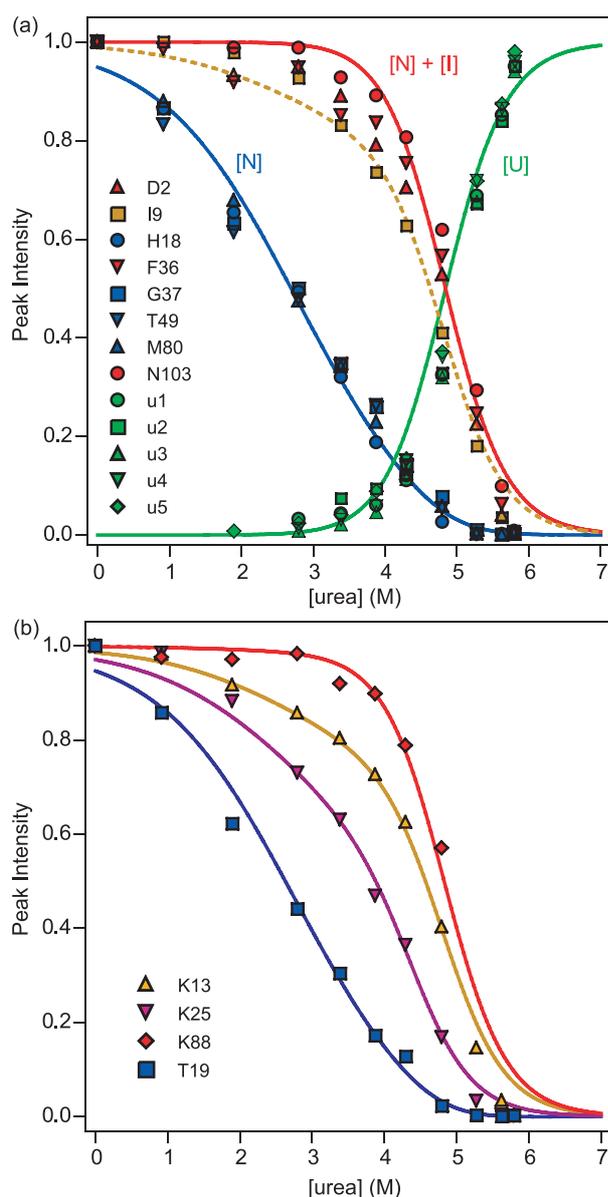
Our observation that a stable intermediate accumulates in the denaturant-induced unfolding equilibrium of *cyt c*, accounting for over 50% of all molecules under certain conditions (denaturant concentrations just below the main unfolding transition), suggests the possibility of using 2D NMR methods for obtaining more detailed structural insight into this partially unfolded state. Because high concentrations of GuHCl result in severe NMR signal loss due to its ionic character, we chose urea as a denaturant for NMR equilibrium unfolding experiments. Working with a uniformly  $^{15}\text{N}$ -labeled sample of H33N *cyt c*, we recorded a series of 11  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra over a wide range of urea concentrations (0 to 5.81 M). As indicated by the examples in Figure 9, the spectrum in the absence of urea (upper left) shows a subset of the well dispersed backbone  $^1\text{H}$ - $^{15}\text{N}$  cross-peaks in the native state, which were assigned to individual residues using standard  $^{15}\text{N}$ -based 2D and 3D NMR techniques (see Materials and Methods).

The spectrum at 5.81 M urea (lower right) contains a new set of peaks with narrower line widths and limited chemical shift dispersion within the region expected for a random coil (a total of 81 peaks can be resolved). Although residue-specific assignments for the U-state are not available at this time, these properties are characteristic of a disordered, largely unfolded conformation. In the spectrum recorded at 2.84 M urea (upper right panel), we still observe the majority of native peaks, although many of them with diminished intensity. One notable example among several residues with strongly reduced peak intensity is Met80, which serves as the sixth iron ligand in the native structure. The spectrum at 4.30 M urea (lower left panel), where the population of the I-state is at a maximum (based on the optical results in Figure 7), is dominated by peaks assigned to the fully unfolded state, although a subset of the native peaks is still visible.

Figure 10(a) shows a plot of the normalized peak intensity for a representative set of resolved N-state peaks as a function of urea concentration. Also shown are a few resolved peaks attributed to the unfolded state (labeled u1–u5). If unfolding of *cyt c* was governed by a two-state mechanism, all peaks assigned to the native state would exhibit the same dependence on denaturant concentration. In contrast, the urea-induced changes in



**Figure 9.** HSQC spectra (expanded region) of H33N *cyt c* (in 0.1 M sodium acetate, 5%  $^2\text{H}_2\text{O}$  (pH 5.0)) recorded at 15 °C in the absence of denaturant (top left) and in the presence of increasing concentrations of urea, as indicated. Assignments are shown for resolved  $^1\text{H}$ - $^{15}\text{N}$  cross-peaks of the native state.



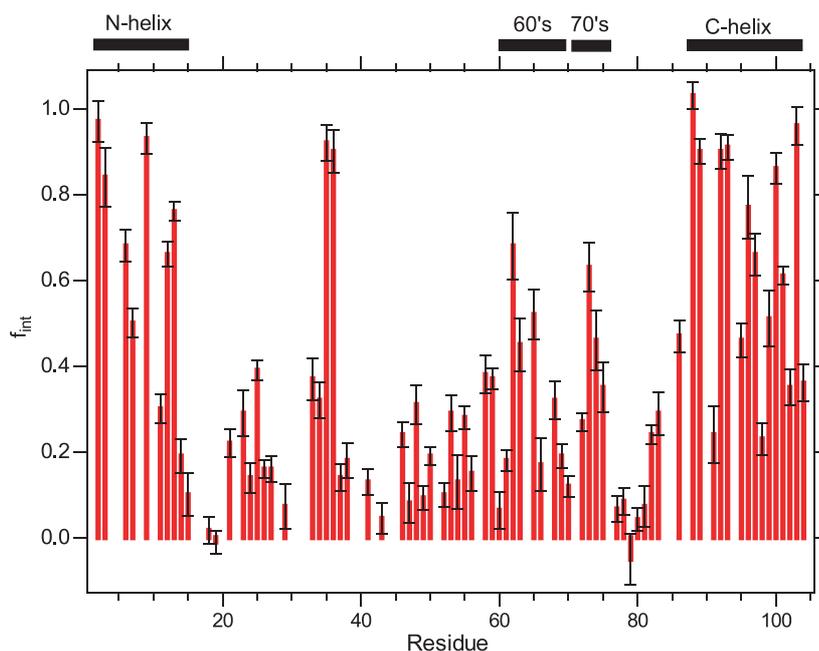
**Figure 10.** Urea dependence of the peak intensity (volume) of resolved cross-peaks in the HSQC spectra of H33N cyt *c*. (a) Representative residues in group 1 (blue symbols) following the urea dependence of the N-state predicted by a three-state model (blue line), residues in group 2 (red symbols) approaching the predicted for the sum of N and I-state populations, and normalized peak intensity for selected U-state peaks (green). The yellow broken line indicates the range of behaviors included in group 2. (b) Normalized peak intensity for four residues illustrating the range of unfolding transitions observed. Continuous lines were calculated with equation (5) for  $f_{\text{int}}$  values of 0 (T19, blue), 0.39 (K25, purple), 0.76 (K13, yellow) and 1.0 (K88, red).

N-state peak intensity shown in Figure 10 segregate into two distinct groups: residues in group 1 (blue) show continuously decreasing intensity between 0 and 5 M urea while residues in group 2 (red) exhibit nearly urea-independent peak intensity up to 3 M followed by a sharp decrease,

mirroring the urea dependence of peaks assigned to the unfolded state (green). The lines in Figure 10 indicate the populations predicted for a three-state model, using equilibrium parameters ( $C_{m1}=2.9$  M,  $m_1=0.6$  kcal mol<sup>-1</sup> M<sup>-1</sup>,  $C_{m2}=4.8$  M,  $m_2=1.4$  kcal mol<sup>-1</sup> M<sup>-1</sup>) similar to those obtained by global analysis of the optical data (Figure 7; Table 1). Group 1 peaks follow the urea-dependence of the native population, which begins to decline already at low urea concentrations. On the other hand, group 2 peaks approach the urea-dependence predicted for the sum of the populations of the N and I-states, indicating that these residues maintain a native-like structural environment in the intermediate state. The same three-state equilibrium parameters also describe the steep increase in U-state peaks above 4 M urea. Thus, the NMR data directly confirm accumulation of a partially unfolded intermediate state and are fully consistent with our optical evidence presented above. If the structural environment in the I-state around a given residue were either native-like or completely unfolded, all residues should follow either the N-state population (group 1) or the sum of N and I-state populations (group 2). While most residues segregate into these two categories, others fall between (Figure 10(b)), suggesting that partial unfolding or increased mobility within I-state molecules can contribute to the N-state resonance intensity (a more detailed analysis of the NMR data will be presented elsewhere).

The HSQC spectra at each urea concentration are dominated by peaks assigned to either the native state or the fully unfolded ensemble. In addition, a number of small cross-peaks appear at intermediate urea concentrations (3.4 M–5.3 M). They are well dispersed in both frequency dimensions and are somewhat broader than the peaks assigned to the native and unfolded populations. The intensity of these minor peaks goes through a maximum between 4.3 M and 4.7 M urea, but reaches no more than about 15% of the full intensity of resolved native or unfolded peaks. In several cases, the minor peaks appear in close proximity to resolved N-state peaks, indicating a population of molecules in slow exchange with the native state. However, the maximum intensity of these peaks is too low to account for the population of the I-state detected by optical methods (Figure 7(c)). In a recent NMR study of another cyt *c* variant (H.C., L. Wang & H.R., unpublished observations), we found that minor peaks appear only after prolonged incubation in the presence of urea, indicating that they are the result of a slow and irreversible chemical modification or aggregation process.

The fact that the HSQC spectra show two predominant sets of peaks with denaturant-dependent intensities, but only small urea-induced chemical shift changes, indicates that there are two ensembles of molecules in slow exchange (relative to the frequency difference between N and U-states). Therefore, the I-state population must be in rapid exchange with N or U (in most cases



**Figure 11.** Plot of  $f_{\text{int}}$  versus residue number representing the relative contribution of the I-state ( $f_{\text{int}}$ ) to the population of resolved (N-state) peaks in the HSQC spectra of H33N cyt *c*.  $f_{\text{int}}$  was obtained by fitting equation (5) to the measured urea dependence of peak volumes (cf Figure 10).

the N-state, based on the urea-dependence of native peaks). Empirically we can describe the relative intensity of the N-state peak for a given residue versus denaturant concentration,  $c$ , as:

$$I_{\text{rel}}(c) = P_{\text{N}}(c) + f_{\text{int}} \times P_{\text{I}}(c) \quad (5)$$

where  $P_{\text{N}}(c)$  and  $P_{\text{I}}(c)$  represent the normalized urea-dependent populations of the native and intermediate states, respectively, and  $f_{\text{int}}$  describes the fractional contribution of the I-state to the N-state NMR peak intensity. Figure 10 shows that this simple relationship is sufficient to describe the urea-dependent intensity changes for the majority of peaks, including those belonging to group 1 ( $f_{\text{int}}=0$ ), group 2 ( $f_{\text{int}}=1$ ), and those falling between these extremes. In Figure 11, the  $f_{\text{int}}$  values obtained by using equation (5) to fit the urea-dependent intensities for all resolved N-state peaks, are plotted versus residue number.

## Discussion

A number of previous studies questioned whether the denaturant-induced unfolding equilibrium of cyt *c* can be described adequately by a two-state model. For example, Myer<sup>45</sup> concluded that urea denaturation of horse heart cyt *c* is characterized by several optically monitored transitions implicating partially unfolded states. Especially clear evidence that an equilibrium intermediate accumulates prior to global unfolding of cyt *c* was presented by Ferri *et al.*,<sup>46</sup> who found that protein species distinct from the native and the unfolded states, in terms of their redox properties, accumulate above 1 M GuHCl, within the region previously considered as the native baseline region.<sup>44,47</sup> Heme CD and absorbance

measurements at 695 nm further confirmed the complex mechanism of cyt *c* equilibrium denaturation.<sup>46</sup> Conventional and magnetic CD measurements in the Soret region were subsequently used to characterize reduced and oxidized forms of cyt *c* and provided additional support for the existence of the partially unfolded state below global unfolding region.<sup>48</sup> Another recent study made use of CD and heme absorbance to show non-coincidence of denaturation curves monitored at different wavelengths.<sup>32</sup> Although these studies clearly document the non-cooperative nature of the cyt *c* unfolding equilibrium, they provide only limited structural insight. Moreover, all these experiments have been performed at neutral pH using GuHCl as the only denaturant. Here, we studied both urea and GuHCl-induced unfolding of WT (Sigma) and recombinant cyt *c*, H33N, at pH 5, which effectively suppresses non-native heme ligation.

Most prior reports of equilibrium intermediates in protein unfolding have been based on the non-coincidence of denaturation curves obtained by monitoring various spectroscopic probes.<sup>1,49,50</sup> One limitation of this approach is that small populations of intermediate states are not expected to result in significant differences in the apparent  $C_{\text{m}}$  and  $m$ -values, and intermediates can be detected only if their spectroscopic properties show measurable differences from both the native and unfolded states. Moreover, comparison of different probes requires normalization of the raw data, which can be model-dependent. In the simplest case the pre and post-transition baselines are linear and shallow, and a two-state model with linear baselines (six-parameter fit) is adequate to describe the data.<sup>51</sup> However, non-linear and/or highly sloped baselines are not unusual, and additional information is

required for a meaningful analysis of the data. We have been able to overcome these problems by using a systematic global analysis approach to fit a multi-state equilibrium model with a common set of thermodynamic parameters to a family of unfolding curves recorded as a function of wavelength using different spectroscopic methods. Global fitting methods have been used extensively to analyze time-resolved fluorescence and other spectroscopic data, including denaturant and pH-induced two-state unfolding transitions.<sup>48,52,53</sup> However, there have been few systematic applications of global least-squares fitting methods to study the spectral properties of intermediates in three-state folding equilibria.<sup>54</sup> In the present study, this strategy has provided clear evidence for accumulation of a partially unfolded equilibrium state in two forms of cyt *c* (WT and H33N variant) at low to intermediate denaturant (GuHCl and/or urea) concentrations. The observed discrepancies between the transition curves as monitored by different spectroscopic observables, including heme absorbance and CD changes at multiple wavelengths, tryptophan fluorescence quenching by the heme (Figures 5, 7 and 8), as well as NMR peak intensity (Figure 10) are clearly inconsistent with a two-state unfolding transition. Moreover, individual unfolding curves for many of the probes used (e.g. near-UV CD and heme absorbance at certain wavelengths) exhibit a complex shape that cannot be fitted by a simple two-state model. In contrast, a three-state model can fully account for all of these observations, including the dispersion among the transition curves for different probes and the biphasic denaturant dependence of individual probes. Although small amounts of additional states cannot be ruled out, the quality of the fits obtained in our global fitting procedure indicates that a three-state equilibrium mechanism (Scheme 1) is sufficient to account for all of the available optical data as a function of both urea and GuHCl.

There are examples in the literature where an apparent equilibrium intermediate has been attributed to specific binding of a denaturant molecule to a folded state with native-like structure.<sup>55,56</sup> In the case of cyt *c*, Thomas *et al.*<sup>48</sup> proposed that a guanidinium ion may displace the side-chain of Arg38 in cyt *c*, which forms a functionally important salt bridge involving the buried heme propionate side-chain.<sup>57</sup> To address this possibility, we carried out a detailed comparison of the GuHCl and urea-induced unfolding transitions for H33N cyt *c*. In the absence of specific binding effects, we expect that the free energy of unfolding extrapolated to  $c=0$  is independent of the denaturant used. Indeed, Table 1 shows very similar values for the total unfolding energy,  $\Delta G_{\text{tot}}$ , for the GuHCl and urea-induced unfolding transitions of H33N cyt *c*. However, the free energy changes for the individual transitions,  $\Delta G_1$  ( $N \rightleftharpoons I$ ) and  $\Delta G_2$  ( $I \rightleftharpoons U$ ) show significant denaturant-dependent differences (Table 1). This behavior may be due to the fact that GuHCl is an ionic denaturant, which apparently

raises the free energy of the intermediate relative to the native state (by 1.1 kcal mol<sup>-1</sup>; Table 1) without significant effect on the total free energy of unfolding (0.3 kcal mol<sup>-1</sup>). Despite these differences, comparison of Figures 3(b) and 7(c) indicates that in both denaturants the intermediate reaches a peak population of about 50% below the midpoint of the major unfolding transition. Moreover, there are striking similarities between the GuHCl and urea-induced intermediates in terms of their spectroscopic properties (Table 2), including a blue-shifted Soret band (cf Figures 3(a) and 6(b)), a native-like far-UV CD signal and diminished intensity in the near-UV CD region (cf Figures 4 and 7(a)).

It is well known that folding of cyt *c* is intimately linked with the presence of the heme and its ligands.<sup>21,58</sup> Studies of oxidized cyt *c* from several species have shown that the native methionine sulfur-iron ligand is replaced by a His, Lys or N-terminal NH<sub>2</sub> (if unprotected) under denaturing conditions near neutral pH or above.<sup>30,35,38</sup> The presence of a non-native heme ligand can result in the trapping of transient folding intermediates if folding experiments are conducted near neutral pH.<sup>20,21,59</sup> Misligation of the heme has also been linked with the accumulation of equilibrium unfolding intermediates.<sup>30,31,36,60</sup> This study was designed to minimize such complications by working at mildly acidic pH where His and Lys side-chains are protonated and no longer available as potential non-native heme ligands. As an additional control, we compared the equilibrium unfolding behavior of WT cyt *c* in GuHCl with that of the H33N variant, which lacks the primary non-native His ligand.<sup>38</sup> While the H33N variant is slightly more stable than the WT in terms of  $\Delta G_{\text{tot}}$  (Table 1), both proteins unfold *via* highly populated equilibrium states (cf Figures 3(b) and 8(b)) with very similar spectroscopic properties (Table 2). Thus, we can clearly rule out the possibility that coordination of His33 to the heme iron is responsible for accumulation of an equilibrium intermediate. Since the pK<sub>a</sub> values of other potential heme ligands, including lysine residues and the N-terminal amino group (which is not acetylated in the case of the recombinant H33N variant), are higher than His, they are even less likely to be deprotonated and bound to the heme iron in the I-state at pH 5. However, the fact that a prominent absorbance band at 625 nm appears at GuHCl concentrations well below 2 M (Figure 1) where the changes in the far-UV CD and fluorescence signals are minimal (Figure 5) is a strong indication that the native ligand interaction between the Met80 sulfur and the heme iron has been severed in the intermediate. Together with the comparison of urea and GuHCl discussed above, these observations indicate that the I-state represents a genuine conformational intermediate that accumulates under moderately denaturing conditions irrespective of the denaturant used or the presence of any non-native heme ligands.

Despite differences in solution conditions, the thermodynamic parameters for the major (second) unfolding transition of WT cyt *c* observed here (Table 1) show reasonable agreement with values reported in the literature.<sup>28,47,61</sup> However, the total free energy ( $\Delta G_{\text{tot}} = 9.8(\pm 0.3) \text{ kcal mol}^{-1}$ ) and *m*-value ( $m_{\text{tot}} = 4.39(\pm 0.14) \text{ kcal mol}^{-1} \text{ M}^{-1}$ ) obtained in our global three-state fit are substantially higher than published values, since they include contributions from the pre-transition that has historically been dismissed as a baseline effect. Thus, the use of a two-state model to fit a multi-state equilibrium with overlapping transitions can seriously underestimate the overall thermodynamic stability of a protein. To illustrate this point, we used the standard two-state model with six free parameters ( $C_m$  and *m*, intercepts and slopes for the pre and post-transition baselines) to fit individual unfolding curves for H33N cyt *c* in GuHCl, including those monitored by CD (Figure 4), fluorescence and heme absorbance at selected wavelengths (cf Figure 2). The apparent  $C_m$  values are close to the global value for the second transition ( $2.34(\pm 0.1) \text{ M}$ ). However, the two-state *m*-values range from 2.8–3.4  $\text{kcal mol}^{-1} \text{ M}^{-1}$  resulting in apparent free energies (6.6–8.2  $\text{kcal mol}^{-1}$ ) well below the total free energy obtained in the global three-state fit (10.8  $\text{kcal mol}^{-1}$ ). Moreover, for some of the data (e.g. the near-UV CD curve in Figure 4) a two-state fit results in a steep slope for the pre-transition baseline implying that the native-state signal varies strongly with denaturant concentration. By contrast, global fitting of a three-state equilibrium model (Scheme 1) yields a single, self-consistent, set of thermodynamic parameters (Table 1) and accounts for all of the data without invoking highly sloping native baselines.

In addition to accurately defining the thermodynamic parameters for each unfolding transition (Table 1), our systematic global analysis of the unfolding transitions as monitored by a variety of spectroscopic parameters also provides insight into the structural properties of the equilibrium intermediate (Table 2). The fractional change in the far-UV CD signal,  $f(\theta_{225})$ , associated with the  $\text{N} \rightleftharpoons \text{I}$  transition is close to zero (H33N cyt *c* in urea) or slightly negative (WT and H33N in GuHCl), indicating that the helix content of the I-state is similar to that of the N-state, or even somewhat enhanced. The low values for  $f(\text{Em}_{350})$  in both the N and I-states are consistent with a compact ensemble of states with an average Trp59–heme distance significantly smaller than 35 Å (the Förster distance for Trp–heme energy transfer). On the other hand, the  $\text{N} \rightleftharpoons \text{I}$  transition is accompanied by significant changes in the aromatic region of the CD spectrum, including a nearly complete loss of the sharp features near 290 nm assigned to Trp59, indicating that some of the specific side-chain packing interactions are disrupted in the I-state. The observed shift in Soret absorption band to lower wavelengths (Table 2) is also consistent with a loosening of the protein structure and a concomitant

increase in solvent exposure of the heme. At the same time, the presence of a prominent absorption band at 625 nm in the I-state correlates with the loss of the Met80 ligand, which results in a high-spin heme state;<sup>62</sup> the sixth coordination site may remain vacant or occupied by a weak ligand, such as water.<sup>63,64</sup> This conclusion is further supported by the observed loss in NH cross-peak intensity at low urea concentrations for Met80 and adjacent residues (Figures 9–11). It is also consistent with previous reports of changes in the magnetic CD<sup>48</sup> and paramagnetic NMR spectra<sup>30,31</sup> preceding the main unfolding transition of oxidized cyt *c*.

Our NMR analysis of the urea-induced unfolding transition provides novel insight into the structural properties of equilibrium intermediates at the level of individual residues. When we plotted the intensity of  $^1\text{H}$ – $^{15}\text{N}$  peaks versus urea concentration (Figure 10), we found a clear segregation of the residues into two distinct groups. (i) Group 1 residues (blue symbols in Figure 10) exhibit a gradual loss in peak intensity over the 0–4 M urea range, tracking the population of the N-state. These residues are located in regions of the protein where the native structure has been disrupted in the I-state, either through local unfolding or increased mobility, and their relative contribution to the intensity of N-state peaks,  $f_{\text{int}}$ , is low (Figure 11). Residues in this category include the axial heme ligands, His18 and Met80, and adjacent residues (14, 19, 78–82), as well as a majority of the residues in the non-helical regions of cyt *c*. (ii) Group 2 residues (red symbols in Figure 10) retain full peak intensity up to about 3 M urea followed by a steep drop, following the behavior expected for the sum of N and I-state populations. These residues are located in regions of the protein where the native structure is largely preserved in the I-state. Their NH groups experience minimal chemical shift perturbations, and the I-state population contributes significantly to the intensity of N-state peaks ( $f_{\text{int}} > 0.5$ ). This category comprises most of the residues in the N and C-terminal helices, a few residues in the 60s and 70s helices, as well as two residues (L35, F36) involved in hydrophobic contacts with the heme (Figure 11). Thus, the two interacting  $\alpha$ -helices near each end of the cyt *c* molecule appear to be largely intact in the intermediate state while other helical regions appear to be partially disrupted. On the other hand, with the exception of a cluster of residues contacting the heme, the non-helical loop regions of the native cyt *c* structure appear to be largely disordered. These structural patterns agree well with those described by Englander and colleagues, based on native-state hydrogen exchange data.<sup>22,65,66</sup>

In those cases where the relative intensity of a resolved (conformationally shifted)  $^1\text{H}$ – $^{15}\text{N}$  cross-peak remains high up to 3 M urea ( $f_{\text{int}} \sim 1$ ), we can conclude with confidence that the local structural environment of a given residue remains intact in the I-state. Likewise, a steady decrease in cross-peak

intensity with increasing urea concentration indicates a major disruption of the local environment in the I-state. The fact that we found no resolved I-state peaks of sufficient size to account for the intensity loss of N-state peaks implies that the I-state undergoes fast exchange with either the N or U-state. A likely scenario is that for group 1 residues, which experience large chemical shift changes on formation of the I-state, the  $N \rightleftharpoons I$  exchange rate falls into the intermediate exchange regime (compared to the frequency difference between N and I-state resonances), resulting in severe line broadening and loss in measurable peak volume. On the other hand, for group 2 residues, which are characterized by small chemical shift differences between the N and I-state, the rate of the  $N \rightleftharpoons I$  transition falls into the fast-exchange regime, giving rise to a sharp peak at an average chemical shift (close to that of the N-state) with an intensity corresponding to the sum of N and I-state populations. The residues falling in between these limiting cases (Figure 10(a)) experience partial line broadening in an intermediate-exchange regime, so that the I-state makes a fractional contribution to the intensity of the observed peak.

In conclusion, we have detected and structurally characterized a non-native state of oxidized horse cyt *c* that accumulates at equilibrium in the presence of moderate concentrations of GuHCl or urea, accounting for as much as 50% of the molecules at denaturant concentrations below the major unfolding transition. The relative sensitivity to denaturant,  $\alpha_1 = m_1/m_{\text{tot}}$  (Table 2), indicates that the N–I transition is accompanied by a  $\sim 30\%$  increase in solvent accessible surface area relative to the total change associated with unfolding. The I-state appears native-like in terms of its  $\alpha$ -helix content, which is  $\sim 10\%$  higher than that of the N-state, low Trp59 fluorescence consistent with a short Trp59-heme distance, and a Soret band indicative of a solvent-shielded heme environment. On the other hand, the I-state shows major structural differences from the N-state, including the loss of the native Met80 sulfur–iron linkage revealed by near-IR absorption and NMR spectroscopy and disruption in side-chain packing and other tertiary interactions resulting in changes in the near-UV CD spectrum and chemical shift perturbations for most non-helical amide groups. These structural properties are reminiscent of the A-state of cyt *c*, a compact denatured form found under acidic high-salt conditions.<sup>18,67–70</sup> The denaturant-induced intermediate also shares some common features with the alkaline state and other partially unfolded forms of cyt *c* stabilized by non-native heme ligands.<sup>63,71</sup> For example, in their studies on a K73H mutant of yeast iso-1 cyt *c*, Bowler and colleagues found that the transition from the native state to a His73–ligated intermediate is accompanied by a 30% increase in solvation,<sup>72,73</sup> which agrees well with the  $\alpha_1$  values in Table 2. Moreover, recent NMR data on the alkaline transition of iso-1 cyt *c* showed that the

$\alpha$ -helical core remains intact and the loop regions become partially disordered when Lys73 replaces Met80 at alkaline pH.<sup>74</sup> Thus, different means of destabilization that disrupt the native Met80–iron bond give rise to structurally analogous intermediate states.

What is the relationship between this solvent-induced equilibrium intermediate and the transient intermediate states detected in kinetic folding experiments? Although the kinetic intermediates populated over the 100  $\mu\text{s}$ –1000  $\mu\text{s}$  time range are compact, based on time-resolved fluorescence and small-angle X-ray scattering data,<sup>25,75–77</sup> continuous-flow CD measurements indicate that only a fraction of the native helical secondary structure is formed at this stage.<sup>78</sup> In addition, the shallow denaturant-dependence of the observed rates on the sub-millisecond time scale<sup>25,79</sup> suggest that a smaller fraction of residues are shielded from the solvent during the initial folding phases compared to the equilibrium intermediate. Thus, the denaturant-induced equilibrium intermediate appears more highly structured than the intermediate states encountered during early stages of folding. On the other hand, it resembles a late folding (or early unfolding) intermediate,  $N^*$ , which we have previously introduced in order to account for the kinetics of unfolding of cyt *c*.<sup>21,61</sup> The transition from N to  $N^*$  becomes rate-limiting for unfolding at high denaturant concentrations and gives rise to a rate profile (chevron plot) with a highly non-linear unfolding branch<sup>61</sup> (similar behavior has also been reported for two bacterial c-type cytochromes<sup>80,81</sup>). The shallow denaturant dependence of the N to  $N^*$  transition and effect of imidazole (an extrinsic heme ligand) on the kinetics of unfolding confirm that  $N^*$  is a highly structured state lacking the native methionine–iron bond.<sup>61,80</sup>

## Materials and Methods

Horse heart cytochrome *c* (highest grade from Sigma-Aldrich Corp., St Louis, MO) was used without further purification. Urea and GuHCl were obtained from ICN Biomedicals, Inc., Aurora, OH (ultra pure grade). Unless stated otherwise, the standard solution conditions used for biophysical measurements were 0.1 M sodium acetate buffer (NaAc) at pH 5.0 (titrated by adding glacial acetic acid to 0.1 M solution of sodium acetate salt), 15 °C.

## Protein expression and purification

All of the genetic manipulations were performed according to conventional molecular biology or manufacturers' protocols.<sup>82,83</sup> A previously published procedure has been used for expression and purification of unlabeled H33N cyt *c*.<sup>39</sup> Briefly, an expression vector pBP(H33N)/3 was constructed *via* introduction of the horse H33N cyt *c* gene, prepared using PCR-based mutagenesis, into a plasmid pBP(XhoI/BamHI)/3. Ligation of the amplified DNA fragment at XhoI and BamHI sites resulted in the placement of the cyt *c* gene in tandem with the gene of yeast cytochrome *c* heme lyase, CYC3.<sup>84,85</sup> The protein was expressed using the JM109

*Escherichia coli* strain in the presence of ampicillin. The purification protocol used was as described.<sup>39</sup>

In an effort to increase production of uniformly <sup>15</sup>N-labeled H33N *cyt c*, a new expression vector was constructed based on the pET-24a plasmid offering kanamycin resistance. The region of pBP(H33N)/3 containing the H33N and CYC3 genes (~1.3 kb) was PCR amplified using Platinum Pfx DNA Polymerase (Invitrogen Corp.) and inserted into pET-24a *via* newly introduced NdeI and HindIII sites. RapidTrans™ TAM1 competent *E. coli* cells (Vinci-Biochem) were used for plasmid DNA amplification and BL21(DE3) strain (Novagen) was used for *cyt c* expression. In order to achieve optimal *cyt c* production a quick semi-quantitative test procedure was developed. Typically, 1.5 ml of the culture prior to and after IPTG induction was centrifuged to pellet the cells. The pellet was resuspended in 600 µl of 50 mM Tris-HCl, 2 mM EDTA (pH 7.5), plus 10 µl of a 50 mg/ml solution of lysozyme. Following 10 min incubation at 30 °C, 5 µl of 1 M MgCl<sub>2</sub> and 10 µl of 20 mg/ml DNAase I were added, and the incubation was continued for another 5 min at 30 °C. After 2 min centrifugation the supernatant was transferred into a new tube, and the absorbance spectrum was taken in the absence and presence of sodium dithionite. Every few hours time points were analyzed for the amount of holo-*cyt c* produced.

Following transformation of BL21(DE3) a single colony was inoculated in 50 ml of LB media with 1% (w/v) glucose and 50 µg/ml of kanamycin and grown overnight at 37 °C. Then, 1 ml of the starting culture was inoculated in one liter of a semi-minimal media (11 g/l Na<sub>2</sub>HPO<sub>4</sub>, 3 g/l KH<sub>2</sub>PO<sub>4</sub>, 0.5 g/l NaCl, 0.5 g/l sodium citrate, 1 g/l Yeast Nitrogen Base (BD and Co.), 1 g/l <sup>15</sup>NH<sub>4</sub>Cl, 1 mM CaCl<sub>2</sub>, 2 mM MgSO<sub>4</sub>, 1 ml of a vitamin mix (Basal Medium Eagle Vitamin Solution; Invitrogen/Gibco, Carlsbad, CA), 50 µM FeSO<sub>4</sub>, 1% glucose and 50 µg/ml of kanamycin) and bacteria were grown at 37 °C with good aeration until A<sub>600</sub> ~1. At this point IPTG was added up to 0.8 mM and the growth was continued at 30 °C with good aeration for another 36–40 h. *cyt c* purification was performed according to the previously published protocol.<sup>39</sup> At least twice the amount of pure *cyt c* has been obtained using the new expression system, compared to the pBP(XhoI/BamHI)/3-based system, both in minimal and rich media.

### Denaturant-induced unfolding experiments

To ensure constant protein concentration during denaturant (GuHCl or urea) titration experiments, two solutions were prepared at carefully matched protein concentration, one containing the native protein in buffer (0.1 M NaAc (pH 5)) and the other containing fully unfolded protein in the presence of denaturant (6–7 M GuHCl or 9–10 M urea in 0.1 M NaAc (pH 5)). The denaturant concentration was increased at constant volume by replacing an aliquot of the previously measured sample with the same volume of the unfolded protein solution. This experimental design helped to reduce protein consumption. Samples were equilibrated for 3 to 5 min at 15 °C prior to every measurement. Denaturant titration experiments on wild-type *cyt c* (Sigma) were performed on separately prepared samples incubated overnight in the presence of various amounts of denaturant. The denaturant concentration for each point of the titration was determined by measuring refractive index ( $\pm 3 \times 10^{-4}$ ), using a Leica Abbe Mark II refractometer (Leica Microsystems, Exton, PA).

### Optical spectroscopy

Absorbance measurements were performed on a Lambda 6 Perkin-Elmer (Boston, MA) and a Hitachi U-3010 UV/VIS (Digilab Hitachi, Randolph, MA) spectrophotometers. In order to monitor spectral changes over wide 250 nm–800 nm range two separate experiments with different path lengths and protein concentrations were performed. Spectra in the Soret region (350 nm–500 nm) were recorded at protein concentrations of 5 µM or 25 µM using 1 cm or 0.2 cm quartz cuvettes, respectively. For the near-UV (250 nm–350 nm) and visible/near-IR regions (500 nm–800 nm), 80 µM protein solutions and a 1 cm path length were standard. Typical scan rate of 1–2 nm/s and a constant 1 nm bandwidth were used in both cases.

Fluorescence measurements were performed on a PTI fluorimeter (Photon Technology International, Lawrenceville, NJ). Excitation and emission wavelengths were 280 nm and 350 nm, with the excitation and emission bandwidths set to 4 nm and 8 nm, respectively. Typical protein concentration was 10 µM. Denaturation curves were obtained by scanning complete emission spectra or by 100 s kinetic trace measurements at 350 nm to improve quality of the data.

CD measurements were performed on a model 62A DS AVIV circular dichroism spectrometer (Lakewood, NJ) equipped with a thermoelectric temperature control unit. Denaturant-induced changes in the far-UV region were monitored at 225 nm with 0.2 cm path length; changes in the near-UV (aromatic) region were typically followed at 294 nm corresponding to one of two negative peaks originating from Trp<sup>59,42</sup> using a 1 cm cuvette. The bandwidth was usually set to 2 nm for both UV regions and a 120 s kinetic trace was recorded at each denaturant concentration. Protein concentrations were 13 µM and 40 µM for far and near-UV CD measurements, respectively.

### NMR measurements

<sup>15</sup>N-labeled H33N *cyt c* was exchanged extensively with 50 mM sodium bicarbonate (pH 7–8) *via* ultrafiltration (Centriplus YM-10, Millipore) and lyophilized. 1 mM solutions of native and unfolded *cyt c* were prepared by dissolving the lyophilized protein in 0.1 M NaAc, 5% <sup>2</sup>H<sub>2</sub>O (pH 5) containing no urea or 9.7 M urea, respectively, and a small amount of DSS. The total of 7 mg of *cyt c* were used for the equilibrium unfolding experiment. <sup>1</sup>H–<sup>15</sup>N HSQC spectra were collected initially for the fully native and unfolded (9.7 M urea) samples, followed by titration from 0 to 5.8 M urea.

NMR spectra were collected at 15 °C on a Bruker DMX 600 MHz spectrometer equipped with a 5 mm *x,y,z*-shielded pulsed-field gradient triple-resonance probe. To confirm resonance assignments, <sup>15</sup>N edited NOE-HSQC and TOC-HSQC,<sup>86</sup> HNHA,<sup>87</sup> and HNHB<sup>88</sup> spectra were collected. The previously published <sup>1</sup>H assignments for WT horse *cyt c*<sup>89</sup> were used as a reference for assigning the <sup>1</sup>H–<sup>15</sup>N cross-peaks. Urea-induced unfolding was studied by recording <sup>1</sup>H–<sup>15</sup>N HSQC spectra<sup>90</sup> for 11 samples at urea concentrations of 0, 0.92 M, 1.90 M, 2.84 M, 3.39 M, 3.88 M, 4.29 M, 4.80 M, 5.28 M, 5.63 M, and 5.81 M. The samples were prepared by proportional mixing of the native protein stock (1 mM protein in 100 mM sodium acetate (pH 5)) and the unfolded protein stock (1 mM protein in 9.7 M urea with 100 mM sodium acetate (pH 5)). The final urea concentration was

determined by its refractive index.  $^1\text{H}$ - $^{15}\text{N}$  HSQC experiments were run with 256 experiments in  $^{15}\text{N}$  dimension ( $t_1$ ) consisting of 40 scans and 4096 data points in  $^1\text{H}$  dimension ( $t_2$ ). The spectra processing and contour peak integration were done by using Felix (Accelrys, San Diego, CA, USA).

## Acknowledgements

This work was supported by NIH grants GM056250 and CA06927, NSF grant MCB-079148, and an Appropriation from the Commonwealth of Pennsylvania. We thank Dimitry Dolgikh for his expert advice and help with protein expression. We are grateful to the Spectroscopy Support Facility, the Biochemistry and Biotechnology Facility and the DNA Sequencing Facility for their support.

## References

1. Tanford, C. (1968). Protein denaturation. *Advan. Protein Chem.* **23**, 121–282.
2. Pace, C. N. (1975). The stability of globular proteins. *CRC Crit. Rev. Biochem.* **2**, 1–43.
3. Privalov, P. L. (1979). Stability of proteins. *Advan. Protein Sci.* **33**, 167–241.
4. Ptitsyn, O. B. (1995). Molten globule and protein folding. *Advan. Protein Chem.* **47**, 83–229.
5. Fink, A. L. (1995). Molten globules. In *Protein Stability and Folding* (Shirley, B. A., ed.), vol. 40, pp. 343–361, Humana Press Inc., Totowa, NJ.
6. Arai, M. & Kuwajima, K. (2000). Role of the molten globule state in protein folding. *Advan. Protein Chem.* **53**, 209–282.
7. Privalov, P. L. (1996). Intermediate states in protein folding. *J. Mol. Biol.* **258**, 707–725.
8. Dill, K. A., Bromberg, S., Yue, K., Fiebig, K. M., Yee, D. P., Thomas, P. D. & Chan, H. S. (1995). Principles of protein folding—a perspective from simple exact models. *Protein Sci.* **4**, 561–602.
9. Sali, A., Shakhovich, E. & Karplus, M. (1994). How does a protein fold? *Nature*, **369**, 248–251.
10. Roder, H. & Colón, W. (1997). Kinetic role of early intermediates in protein folding. *Curr. Opin. Struct. Biol.* **7**, 15–28.
11. Clarke, A. R. & Waltho, J. P. (1997). Protein folding pathways and intermediates. *Curr. Opin. Biotechnol.* **8**, 400–410.
12. Bilsel, O. & Matthews, C. R. (2000). Barriers in protein folding reactions. *Advan. Protein Chem.* **53**, 153–207.
13. Sanchez, I. E. & Kiefhaber, T. (2003). Evidence for sequential barriers and obligatory intermediates in apparent two-state protein folding. *J. Mol. Biol.* **325**, 367–376.
14. Roder, H., Maki, K., Latypov, R. F., Cheng, H. & Shastry, M. C. R. (2005). Early events in protein folding explored by rapid mixing methods. In *Protein Folding Handbook* (Buchner, J. & Kiefhaber, T., eds), pp. 491–535, Wiley-VCH, Weinheim (part I).
15. Stellwagen, E., Rysavy, R. & Babul, G. (1972). The conformation of horse heart apocytochrome c. *J. Biol. Chem.* **247**, 8074–8077.
16. Ikai, A., Fish, W. W. & Tanford, C. (1973). Kinetics of unfolding and refolding of proteins. II. Results for cytochrome c. *J. Mol. Biol.* **73**, 165–184.
17. Roder, H., Elöve, G. A. & Englander, S. W. (1988). Structural characterization of folding intermediates in cytochrome c by H-exchange labelling and proton NMR. *Nature*, **335**, 700–704.
18. Jeng, M.-F., Englander, S. W., Elöve, G. A., Wand, A. J. & Roder, H. (1990). Structural description of acid-denatured cytochrome c by hydrogen exchange and 2D NMR. *Biochemistry*, **29**, 10433–10437.
19. Jones, C. M., Henry, E. R., Hu, Y., Chan, C.-K., Luck, S. D., Bhuyan, A. *et al.* (1993). Fast events in protein folding initiated by nanosecond laser photolysis. *Proc. Natl Acad. Sci. USA*, **90**, 11860–11864.
20. Sosnick, T. R., Mayne, L., Hiller, R. & Englander, S. W. (1994). The barriers in protein folding. *Nature Struct. Biol.* **1**, 149–156.
21. Elöve, G. A., Bhuyan, A. K. & Roder, H. (1994). Kinetic mechanism of cytochrome c folding: involvement of the heme and its ligands. *Biochemistry*, **33**, 6925–6935.
22. Bai, Y., Sosnick, T. R., Mayne, L. & Englander, S. W. (1995). Protein folding intermediates: native-state hydrogen exchange. *Science*, **269**, 192–197.
23. Pascher, T., Chesick, J. P., Winkler, J. R. & Gray, H. B. (1996). Protein folding triggered by electron transfer. *Science*, **271**, 1558–1560.
24. Takahashi, S., Yeh, S.-R., Das, T. K., Chan, C.-K., Gottfried, D. S. & Rousseau, D. L. (1997). Folding of cytochrome c initiated by submillisecond mixing. *Nature Struct. Biol.* **4**, 44–50.
25. Shastry, M. C. R. & Roder, H. (1998). Evidence for barrier-limited protein folding kinetics on the microsecond time scale. *Nature Struct. Biol.* **5**, 385–392.
26. Babul, J. & Stellwagen, E. (1971). The existence of heme-protein-coordinate-covalent bonds in denaturing solvents. *Biopolymers*, **10**, 2359–2361.
27. Knapp, J. A. & Pace, C. N. (1974). Guanidine hydrochloride and acid denaturation of horse, cow, and *Candida krusei* cytochromes c. *Biochemistry*, **13**, 1289–1294.
28. McLendon, G. & Smith, M. (1978). Equilibrium and kinetic studies of unfolding of homologous cytochromes c. *J. Biol. Chem.* **253**, 4004–4008.
29. van Osdol, W. W., Mayorga, O. L. & Freire, E. (1991). Multifrequency calorimetry of the folding/unfolding transition of cytochrome c. *Biophys. J.* **59**, 48–54.
30. Russell, B. S., Melenkivitz, R. & Bren, K. L. (2000). NMR investigation of ferricytochrome c unfolding: detection of an equilibrium unfolding intermediate and residual structure in the denatured state. *Proc. Natl Acad. Sci. USA*, **97**, 8312–8317.
31. Russell, B. S. & Bren, K. L. (2002). Denaturant dependence of equilibrium unfolding intermediates and denatured state structure of horse ferricytochrome c. *J. Biol. Inorg. Chem.* **7**, 909–916.
32. Moza, B., Qureshi, S. H. & Ahmad, F. (2003). Equilibrium studies of the effect of difference in sequence homology on the mechanism of denaturation of bovine and horse cytochromes-c. *Biochim. Biophys. Acta*, **1646**, 49–56.
33. Segel, D. J., Fink, A. L., Hodgson, K. O. & Doniach, S. (1998). Protein denaturation: a small-angle X-ray scattering study of the ensemble of unfolded states of cytochrome c. *Biochemistry*, **37**, 12443–12451.
34. Mayne, L. & Englander, S. W. (2000). Two-state vs. multistate protein unfolding studied by optical melting and hydrogen exchange. *Protein Sci.* **9**, 1873–1877.

35. Muthukrishnan, K. & Nall, B. T. (1991). Effective concentrations of amino acid side chains in an unfolded protein. *Biochemistry*, **30**, 4706–4710.
36. Godbole, S., Dong, A., Garbin, K. & Bowler, B. E. (1997). A lysine 73 → histidine variant of yeast iso-1-cytochrome *c*: evidence for a native-like intermediate in the unfolding pathway and implications for *m* value effects. *Biochemistry*, **36**, 119–126.
37. Kurchan, E., Roder, H. & Bowler, B. E. (2005). Kinetics of loop formation and breakage in the denatured state of Iso-1-cytochrome *c*. *J. Mol. Biol.* **353**, 730–743.
38. Colón, W., Wakem, L. P., Sherman, F. & Roder, H. (1997). Identification of the predominant non-native histidine ligand in unfolded cytochrome *c*. *Biochemistry*, **36**, 12535–12541.
39. Hagen, S. J., Latypov, R. F., Dolgikh, D. A. & Roder, H. (2002). Rapid intrachain binding of histidine-26 and histidine-33 to heme in unfolded ferrocycytochrome *C*. *Biochemistry*, **41**, 1372–1380.
40. Margoliash, E. & Schejter, A. (1966). Cytochrome *c*. *Advan. Protein Chem.* **21**, 113–286.
41. Schejter, A. & Eaton, W. A. (1984). Charge-transfer optical spectra, electron paramagnetic resonance, and redox potentials of cytochromes. *Biochemistry*, **23**, 1081–1084.
42. Davies, A. M., Guillemette, J. G., Smith, M., Greenwood, C., Thurgood, A. G. P., Mauk, A. G. & Moore, G. R. (1993). Redesign of the interior hydrophilic region of mitochondrial cytochrome *c* by site-directed mutagenesis. *Biochemistry*, **32**, 5431–5435.
43. Schechter, E. & Saludjian, P. (1967). Conformation of ferricytochrome *c*. IV. Relationship between optical absorption and protein conformation. *Biopolymers*, **5**, 788–790.
44. Tsong, T. Y. (1974). The Trp-59 fluorescence of ferricytochrome *c* as a sensitive measure of the overall protein conformation. *J. Biol. Chem.* **249**, 1988–1990.
45. Myer, Y. P., MacDonald, L. H., Verma, B. C. & Pande, A. (1980). Urea denaturation of horse heart ferricytochrome *c*. Equilibrium studies and characterization of intermediate forms. *Biochemistry*, **19**, 199–207.
46. Ferri, T., Poscia, A., Ascoli, F. & Santucci, R. (1996). Direct electrochemical evidence for an equilibrium intermediate in the guanidine-induced unfolding of cytochrome *c*. *Biochim. Biophys. Acta*, **1298**, 102–108.
47. Tsong, T. Y. (1976). Ferricytochrome *c* chain folding measured by the energy transfer of tryptophan 59 to the heme group. *Biochemistry*, **15**, 5467–5473.
48. Thomas, Yg., Goldbeck, Ra. & Klinger, Ds. (2000). Characterization of equilibrium intermediates in denaturant-induced unfolding of ferrous and ferric cytochromes *c* using magnetic circular dichroism, circular dichroism, and optical absorption spectroscopies. *Biopolymers*, **57**, 29–36.
49. Kim, P. S. & Baldwin, R. L. (1982). Specific intermediates in the folding reactions of small proteins and the mechanism of protein folding. *Annu. Rev. Biochem.* **51**, 459–489.
50. Fink, A. L. (1995). Compact intermediate states in protein folding. *Annu. Rev. Biophys. Biomol. Struct.* **24**, 495–522.
51. Santoro, M. M. & Bolen, D. W. (1992). A test of the linear extrapolation of unfolding free energy changes over an extended denaturant concentration range. *Biochemistry*, **31**, 4901–4907.
52. Ionescu, R. M. & Eftink, M. R. (1997). Global analysis of the acid-induced and urea-induced unfolding of staphylococcal nuclease and two of its variants. *Biochemistry*, **36**, 1129–1140.
53. Hamada, D. & Goto, Y. (1997). The equilibrium intermediate of beta-lactoglobulin with non-native alpha-helical structure. *J. Mol. Biol.* **269**, 479–487.
54. Maki, K., Cheng, H., Dolgikh, D. A., Shastry, M. C. & Roder, H. (2004). Early events during folding of wild-type staphylococcal nuclease and a single-tryptophan variant studied by ultrarapid mixing. *J. Mol. Biol.* **338**, 383–400.
55. Dunbar, J., Yennawar, H. P., Banerjee, S., Luo, J. & Farber, G. K. (1997). The effect of denaturants on protein structure. *Protein Sci.* **6**, 1727–1733.
56. Pike, A. C. W. & Acharya, K. R. (1994). A structural basis for the interaction of urea with lysozyme. *Protein Sci.* **3**, 706–710.
57. Bushnell, G. W., Louie, G. V. & Brayer, G. D. (1990). High-resolution three-dimensional structure of horse heart cytochrome *c*. *J. Mol. Biol.* **214**, 585–595.
58. Babul, J. & Stellwagen, E. (1972). Participation of the protein ligands in the folding of cytochrome *c*. *Biochemistry*, **11**, 1195–1200.
59. Elöve, G. A. & Roder, H. (1991). Structure and stability of cytochrome *c* folding intermediates. *ACS Symp. Ser.* **470**, 50–63.
60. Hammack, B., Godbole, S. & Bowler, B. E. (1998). Cytochrome *c* folding traps are not due solely to histidine-heme ligation: direct demonstration of a role for N-terminal amino group-heme ligation. *J. Mol. Biol.* **275**, 719–724.
61. Colón, W., Elöve, G. A., Wakem, L. P., Sherman, F. & Roder, H. (1996). Side chain packing of the N- and C-terminal helices plays a critical role in the kinetics of cytochrome *c* folding. *Biochemistry*, **35**, 5538–5549.
62. Lambeth, D. O., Campbell, K. L., Zand, R. & Palmer, G. (1973). The appearance of transient species of cytochrome *c* upon rapid oxidation or reduction at alkaline pH. *J. Biol. Chem.* **248**, 8130–8136.
63. Rosell, F. I., Harris, T. R., Hildebrand, D. P., Dopner, S., Hildebrandt, P. & Mauk, A. G. (2000). Characterization of an alkaline transition intermediate stabilized in the Phe82Trp variant of yeast iso-1-cytochrome *c*. *Biochemistry*, **39**, 9047–9054.
64. Yeh, S.-R. & Rousseau, D. L. (1998). Folding intermediates in cytochrome *c*. *Nature Struct. Biol.* **5**, 222–228.
65. Milne, J. S., Xu, Y., Mayne, L. C. & Englander, S. W. (1999). Experimental study of the protein folding landscape: unfolding reactions in cytochrome *c*. *J. Mol. Biol.* **290**, 811–822.
66. Maity, H., Maity, M., Krishna, M. M., Mayne, L. & Englander, S. W. (2005). Protein folding: the stepwise assembly of foldon units. *Proc. Natl Acad. Sci. USA*, **102**, 4741–4746.
67. Robinson, J. B., Jr, Strottmann, J. M. & Stellwagen, E. (1983). A globular high spin form of ferricytochrome *c*. *J. Biol. Chem.* **258**, 6772–6776.
68. Goto, Y., Calciano, L. J. & Fink, A. L. (1990). Acid-induced folding of proteins. *Proc. Natl Acad. Sci. USA*, **87**, 573–577.
69. Hamada, D., Kuroda, Y., Kataoka, M., Aimoto, S., Yoshimura, T. & Goto, Y. (1996). Role of heme axial ligands in the conformational stability of the native and molten globule states of horse cytochrome *c*. *J. Mol. Biol.* **256**, 172–186.

70. Kataoka, M., Hagihara, Y., Mihara, K. & Goto, Y. (1993). Molten globule of cytochrome *c* studied by small angle X-ray scattering. *J. Mol. Biol.* **229**, 591–596.
71. Davis, L., Schejter, A. & Hess, G. P. (1974). Alkaline isomerization of oxidized cytochrome *c*. Equilibrium and kinetic measurements. *J. Biol. Chem.* **249**, 2624–2632.
72. Nelson, C. J. & Bowler, B. E. (2000). pH dependence of formation of a partially unfolded state of a Lys 73→His variant of iso-1-cytochrome *c*: implications for the alkaline conformational transition of cytochrome *c*. *Biochemistry*, **39**, 13584–13594.
73. Nelson, C. J., LaConte, M. J. & Bowler, B. E. (2001). Direct detection of heat and cold denaturation for partial unfolding of a protein. *J. Am. Chem. Soc.* **123**, 7453–7454.
74. Assfalg, M., Bertini, I., Dolfi, A., Turano, P., Mauk, A. G., Rosell, F. I. & Gray, H. B. (2003). Structural model for the alkaline form of ferricytochrome *C*. *J. Am. Chem. Soc.* **125**, 2913–2922.
75. Pollack, L., Tate, M. W., Darnton, N. C., Knight, J. B., Gruner, S. M., Eaton, W. A. & Austin, R. H. (1999). Compactness of the denatured state of a fast-folding protein measured by submillisecond small-angle x-ray scattering. *Proc. Natl Acad. Sci. USA*, **96**, 10115–10117.
76. Akiyama, S., Takahashi, S., Kimura, T., Ishimori, K., Morishima, I., Nishikawa, Y. & Fujisawa, T. (2002). Conformational landscape of cytochrome *c* folding studied by microsecond-resolved small-angle x-ray scattering. *Proc. Natl Acad. Sci. USA*, **99**, 1329–1334.
77. Hagen, S. J. & Eaton, W. A. (2000). Two-state expansion and collapse of a polypeptide. *J. Mol. Biol.* **297**, 781–789.
78. Akiyama, S., Takahashi, S., Ishimori, K. & Morishima, I. (2000). Stepwise formation of alpha-helices during cytochrome *c* folding. *Nature Struct. Biol.* **7**, 514–520.
79. Shastry, M. C. R., Sauder, J. M. & Roder, H. (1998). Kinetic and structural analysis of submillisecond folding events in cytochrome *c*. *Acc. Chem. Res.* **31**, 717–725.
80. Sauder, J. M., MacKenzie, N. E. & Roder, H. (1996). Kinetic mechanism of folding and unfolding of *Rhodobacter capsulatus* cytochrome *c*. *Biochemistry*, **35**, 16852–16862.
81. Gianni, S., Travaglini-Allocatelli, C., Cutruzzola, F., Brunori, M., Shastry, M. C. & Roder, H. (2003). Parallel pathways in cytochrome *c*(551) folding. *J. Mol. Biol.* **330**, 1145–1152.
82. Ausubel, P. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. & Struhl, K. (1993). *Current Protocols in Molecular Biology*, Greene Publishing Associates, Inc., Wiley, New York.
83. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
84. Dolgikh, D. A., Latypov, R. F., Abdullaev, Z. K., Colon, W., Roder, H. & Kirpichnikov, M. P. (1998). Expression of mutant horse cytochrome *c* genes in *Escherichia coli*. *Bioorg. Khim.* **24**, 756–759.
85. Pollock, W. B. R., Rosell, F. I., Twitchett, M. B., Dumont, M. E. & Mauk, A. G. (1998). Bacterial expression of a mitochondrial cytochrome *c*. Trimethylation of Lys72 in yeast iso-1-cytochrome *c* and the alkaline conformational transition. *Biochemistry*, **37**, 6124–6131.
86. Talluri, S. & Wagner, G. (1996). An optimized 3D NOESY-HSQC. *J. Magn. Reson.* **112**, 200–205.
87. Kuboniwa, H., Grzesiek, S., Delaglio, F. & Bax, A. (1994). Measurement of HN-H alpha J couplings in calcium-free calmodulin using new 2D and 3D water-flip-back methods. *J. Biomol. NMR*, **4**, 871–878.
88. Bax, A., Vuister, G., Grzesiek, S., Delaglio, F., Wang, A., Tschudin, R. & Zhu, G. (1994). Measurement of homo- and heteronuclear J couplings from quantitative J correlation. *Methods Enzymol.* **239**, 79–105.
89. Feng, Y., Roder, H., Englander, S. W., Wand, A. J. & Di Stefano, D. L. (1989). Proton resonance assignments of horse ferricytochrome *c*. *Biochemistry*, **28**, 195–203.
90. Bax, A. & Pochapsky, S. S. (1992). Optimized recording of heteronuclear multidimensional NMR spectra using pulsed field gradients. *J. Magn. Reson.* **99**, 638–643.

Edited by K. Kuwajima

(Received 11 November 2005; received in revised form 11 January 2006; accepted 15 January 2006)  
Available online 3 February 2006