Published in final edited form as:

J Mol Biol. 2009 April 24; 388(1): 159–167. doi:10.1016/j.jmb.2009.02.059.

The Determinants of Stability and Folding in Evolutionarily Diverged Cytochromes *c*

Megan C. Thielges, Jörg Zimmermann, Philip E. Dawson, and Floyd E. Romesberg*

Department of Chemistry, The Scripps Research Institute, La Jolla, CA, USA.

Summary

Cytochrome *c* has served as a paradigm for the study of protein stability, folding, and molecular evolution, but it remains unclear how these aspects of the protein are related. For example, while the bovine and equine cytochromes *c* are known to have different stabilities, and possibly different folding mechanisms, it is not known how these differences arise from just three amino acid substitutions introduced during divergence. Using site-selectively incorporated carbon-deuterium bonds we show that like the equine protein, bovine cytochrome *c* is induced to unfold by guanidine-hydrochloride via a stepwise mechanism, but it does not populate an intermediate as is observed with the equine protein. The increased stability also results in more similar free energies of unfolding observed at different sites within the protein, giving the appearance of a more concerted mechanism. Furthermore, we show that the differences in stability and folding appear to result from a single amino acid substitution which stabilizes a helix by allowing for increased solvation of its N-terminus.

Keywords

cytochrome *c*; carbon-deuterium; FT IR; protein folding; evolution

Introduction

Cytochrome *c* functions to transport electrons between the reductase and oxidase components of the respiratory chain and has emerged as one of the most intensively studied of all proteins. It has served as a model system for the study of not only biological electron transfer, but also protein stability, folding, and evolution.^{1–17} However, it remains unclear how evolution has affected the function, stability, or folding of the protein. Equilibrium studies focused on the horse heart protein (hcyt *c*) have shown that it unfolds via a rather complex stepwise process.^{3,15,18–24} Interestingly, the closely related bovine protein (bcyt *c*), which differs from hcyt *c* by only three mutations (T47S, K60G, and T89G; where the equine residue is listed first, but with no implication as to how or when the mutations were acquired), is more stable, and there is also evidence that it unfolds via a somewhat different

© 2009 Elsevier Ltd. All rights reserved.

*Corresponding author. Department of Chemistry, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037, USA. Phone: +1 858 551 4575. Fax: +1 858 784 7472. floyd@scripps.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Supplementary Material

Supplementary methods and data associated with this article can be found in the online version.

mechanism. For example, amide IR absorption studies suggest that both proteins are thermally induced to unfold in a stepwise manner under equilibrium conditions, but the helical structure of bcyt *c* is lost in a cooperative manner while the helical structure of hcyt *c* is lost in a stepwise manner.¹⁵ Moreover, while recent UV/vis and circular dichroism studies suggest that guanidine-hydrochloride (GdnHCl) induces hcyt *c* to unfold in a stepwise manner, they also suggest that bcyt *c* unfolds via a single cooperative transition.¹⁶ These differences emphasize the importance of residue-specific resolution, which is lacking in the techniques employed, and render it difficult to judge whether the diverged proteins fold by similar mechanisms. It thus remains unclear why bcyt *c* is more stable and if the increased stability affects the folding mechanism.

In principle, vibrational spectroscopy provides a direct and bond-specific approach to the characterization of a protein. Unfortunately, the spectral congestion inherent to proteins limits its application. Previous studies have made use of isotopic labeling in conjunction with difference Fourier transform infrared (FT IR) spectroscopy to alleviate some of the spectral complexity and thereby examine protein vibrations, such as those in the amide region of the IR spectrum.^{25–27} However, these studies remain limited by the significant protein absorption in this region of the IR spectrum, as well as the inherent delocalization of these vibrations due to the presence of many other vibrations of commensurate frequency and the strong dipole-dipole interactions that couple them. As a result, linewidths and frequencies of the observable absorptions are difficult to accurately determine and even more difficult to interpret in terms of specific protein motions. As part of a program to develop general probes of proteins, including their microenvironments and dynamics, we have been developing the use of carbon-deuterium (C-D) bonds as FT IR probes.^{28–36} C-D bonds may be incorporated nonperturbatively anywhere throughout a protein and they provide environmentally sensitive IR absorptions in a transparent region ($\sim 2200\text{ cm}^{-1}$) of the otherwise prohibitively congested protein IR spectrum. For example, comparison of the C-D absorption signals as a function of denaturant allows for the characterization of how specific parts of the protein are induced to unfold.^{29,30,34} In addition to their high structural resolution, C-D probes also provide the high temporal resolution inherent to IR spectroscopy such that even the most rapidly interconverting protein conformations may be resolved. Previously, we used this technique to examine the GdnHCl-induced unfolding of hcyt *c*, and we observed a stepwise process with different parts of the protein undergoing local unfolding transitions at different denaturant concentrations,^{29,30} consistent with the conclusions drawn from a large body of previous studies.^{23,37–42}

To address the differences between hcyt *c* and bcyt *c*, we report a residue-specific analysis of GdnHCl-induced unfolding of bcyt *c* for comparison with that already reported for hcyt *c*.²⁹ Specifically, we characterized bcyt *c* selectively labeled with C-D bonds at the native heme ligand Met80, which in hcyt *c* is induced to dissociate from the prosthetic group at relatively low denaturant concentrations, leading to a stable folding intermediate, and at two residues located in helices, Leu68 and Leu94, which in hcyt *c* are induced to locally unfold only at relatively high denaturant concentrations. The data confirm that bcyt *c* is more stable than hcyt *c*, but also demonstrate that like hcyt *c*, bcyt *c* unfolding is not concerted. However, the sequence differences do affect the folding mechanism by precluding the formation of the folding intermediate, and making the free energies of unfolding for the different structural elements more similar, thus making the individual transitions in bcyt *c* more coincident than in hcyt *c*. UV/vis spectra as a function of denaturant suggest that the different behaviors result entirely from one of the three mutations introduced during divergence, K60G. The mutation appears to stabilize the N-terminus of a helix which increases the stability of different parts of the protein via packing and salt-bridge formation, resulting in the more coincident and less heterogeneous unfolding.

Results

We semi-synthesized bcyt *c* site-specifically labeled with C-D bonds by incorporating (*methyl-d*₃) methionine at Met80 ((*d*₃)Met80 bcyt *c*) or a 1:1 mixture of C_{δ1}-*d*₃ and C_{δ2}-*d*₃ labeled leucine isotopomers at Leu68 and Leu94 ((*d*₃)Leu68 bcyt *c* and ((*d*₃)Leu94 bcyt *c*, respectively) (Figure 1). These residues were selected because in hcyt *c* they are sensitive to different unfolding transitions that contribute to the stepwise unfolding process. As mentioned above, dissociation of Met80 from the heme center is induced at relatively low GdnHCl concentrations, while structural changes associated with a loss of helical structure at Leu68 and Leu94 occur only at higher concentrations. Thus, characterizing the GdnHCl-induced transitions at these residues in bcyt *c* was expected to test whether these two diverged proteins fold via similar mechanisms.

The spectra of both (*d*₃)Leu68 and (*d*₃)Leu94 bcyt *c* in the absence of denaturant show relatively intense overlapping absorptions at ~2215 cm⁻¹ (Figure 2) and a weaker absorption at 2135 cm⁻¹ (data not shown). By analogy to hcyt *c*²⁹ these absorptions are assigned as the two asymmetric stretches and the symmetric stretch of the *d*₃-methyl group, respectively. Previously, the overlapping asymmetric stretches of these residues in hcyt *c* were fit to two Gaussian functions; however, in bcyt *c* fitting them requires either four Gaussian functions or two Gaussian functions and a quasi-Voigt function (the difference in fits between the two proteins is unlikely to reflect any actual difference between them, but rather appears to result from the increased signal to noise in the bcyt *c* spectra). We deemed the four Gaussian fit to be more physically realistic as it deconvolutes the observed spectrum into two pairs of equally intense absorptions, which we assume correspond to either two pairs of asymmetric stretches that experience two different microenvironments or two pairs of asymmetric stretches of diastereomeric methyl groups of the 1:1 mixture of C_{δ1}-*d*₃ and C_{δ2}-*d*₃ labeled leucine isotopomers (we also note that fitting the spectra in either manner results in the same unfolding behavior). Upon unfolding bcyt *c* in 6 M GdnHCl, the spectra change significantly, becoming virtually identical to that of the free amino acid for both (*d*₃)Leu68 and (*d*₃)Leu94 bcyt *c*, and were fit by a combination of a quasi-Voigt and a Gaussian function (Figure 2). All spectra recorded at intermediate GdnHCl concentrations are well fit by a superposition of the folded and high denaturant protein spectra, allowing us to determine the fraction of Leu68 and Leu94 that experience a folded-like environment as a function of denaturant (Figure 2 and 3). Clear two-state behavior is apparent for the transitions observed at both residues, with virtually identical midpoints of 2.65 ± 0.01 and 2.63 ± 0.01 M GdnHCl, respectively, for (*d*₃)Leu68 and (*d*₃)Leu94.

In the spectra of (*d*₃)Met80 bcyt *c*, we observed an intense absorption at ~2130 cm⁻¹ and weaker overlapping absorptions around 2215 cm⁻¹, which, based on previous studies of hcyt *c*, are assigned as the symmetric and overlapping asymmetric stretching vibrations²⁹ (Figure 2). In the absence of denaturant, the symmetric stretch is best fit by a sum of two Gaussian functions; however, one was very low in intensity and based on previous studies,³⁴ can be attributed to the presence of a small amount of unfolded protein. In 6.0 M GdnHCl, the symmetric stretching vibration is best fit by a quasi-Voigt function and is virtually identical to that observed for the solvent-exposed free amino acid in the same solvent. The frequencies and linewidths of the (*d*₃)Met80 absorptions in the folded state of both proteins are similar, as are the absorptions in the unfolded state. However, all of the bcyt *c* spectra at intermediate GdnHCl concentrations were well fit by a superposition of the folded and high denaturant spectra (Figure 2), which is not the case with hcyt *c*, where the spectra required an additional Gaussian function at intermediate denaturant concentrations.³⁰ When an extra function was included in the (*d*₃)Met80 bcyt *c* fits, depending on the details of the fitting procedure, its amplitude as a function of denaturant did not show the behavior expected for an intermediate, either mimicking the native state, suggesting that it was simply fitting some

of the native state's absorption, or showing no dependence on denaturant concentration. Thus, unlike Met80 hcyt *c*, which populates a folding intermediate that peaks in concentration at ~1 M GdnHCl,²⁹ there is no evidence that a similar intermediate is populated at Met80 in bcyt *c*. The midpoint of the observed two-state transition at Met80 is 2.51 ± 0.02 M GdnHCl (Figure 3). While the difference in unfolding midpoints observed at Met80 and that observed at Leu68 and Leu94 are small, they appear to be significant; the transition induced by GdnHCl at Met80, which likely corresponds to a conformational change or dissociation from the heme center, appears to occur at slightly lower denaturant concentration than the transitions induced at Leu68 and Leu94, which correspond to changes in the structure of the corresponding helices or their environments.

To further quantify the observed transitions at Met80, Leu68, and Leu94 of bcyt *c*, we analyzed the FT IR data using the linear extrapolation method.⁴³ The values of $\Delta G(0)$ range from 8.8 to 11.0 kcal/mol, and those for *m* range from 3.3 to 4.4 kcal/M/mol (Table 1). Differences outside of error are found in both the $\Delta G(0)$ and *m* values for Met80 and Leu68, further reinforcing the conclusion that neither protein unfolds via a simple two-state transition. However, the equilibrium folding mechanisms are not identical in the two proteins, largely due to more similar *m* values and a relative increase in the free energy for unfolding at Leu94. Along with the absence of the stable folding intermediate at Met80, this data suggests that unfolding occurs in a more homogeneous manner with bcyt *c*.

To explore how the specific mutations introduced during the divergence of these two proteins affect stability and folding, we recombinantly expressed hcyt *c* with the three bovine mutations introduced individually: T47S, K60G, and T89G (Figure 1). (For technical reasons, these proteins also contain the mutation M65L, which slightly destabilizes the protein). The stability of each wild-type and mutant protein was then quantified using its UV/vis absorption at 530 and 695 nm as a function of GdnHCl (Figure 4) and the data was analyzed using the linear extrapolation method (Table 2).⁴³ Comparison of the wild-type proteins reveals that the bovine variant is ~1 kcal/mol more stable than the horse variant, in good agreement with previous reports.^{6,8,11} Comparison of the hcyt *c* mutants reveals that T47S and T89G have little effect on stability, while K60G increases $\Delta G(0)$ by ~1 kcal/mol.

Discussion

Cytochrome *c* has served as a paradigm for the study of protein stability and folding, and especially with hcyt *c*, many details are now understood. However, while effort has also been directed toward understanding its evolution (see for example, Refs. 6–17), it remains unclear how the specific amino acid differences introduced during protein divergence affect stability and/or folding. For example, while numerous studies have revealed that the three mutations introduced during the divergence of bcyt *c* and hcyt *c*, T47S, K60G, and T89G, stabilize the bovine variant,^{6,8,11} the origins of this stabilization are unknown. Moreover, conflicting data exists in the literature regarding the similarity of the folding mechanisms of the two proteins,^{15,16} and thus it remains unclear if or how the increased stability of bcyt *c* affects its folding.

To generate a more detailed understanding of the differences in folding mechanisms, we characterized the GdnHCl-induced transitions of the IR absorptions in bcyt *c* that arise from C-D bonds selectively incorporated at Met80, Leu68, and Leu94, and compared them to those previously reported for hcyt *c*. These experiments take advantage of the high structural and 9 temporal resolution inherent to IR spectroscopy and thus are able to identify the population of even partially unfolded intermediates that are in rapid exchange with each other or with the native state. In all cases, except for hcyt *c* (*d*₃)Met80, the absorptions show a two-state transition from spectra corresponding to the native folded state to the high

denaturant state. This does not imply that unfolding of the entire protein is two-state, but rather that the residues under observation are sensitive to a two-state transition that may be either local or more global. Indeed, the observation that the two-state transitions observed at each residue do not share the same midpoints reveals that neither hcyt *c* nor bcyt *c* unfold via a simple two-state process. Rather each appears to unfold via a series of different transitions, localized to different parts of the protein, which collectively induce the loss of native structure, and ultimately result in an unfolded protein. While both proteins are induced to unfold via stepwise mechanisms, the transitions induced with bcyt *c* are much more coincident than those with hcyt *c*. Moreover, in the case of bcyt *c* there is no evidence that an intermediate accumulates at low denaturant concentration as observed with hcyt *c*.³⁴

To further compare the folding mechanisms of the diverged proteins, we analyzed the residue-specific denaturant-induced transitions using the linear extrapolation method⁴³ (Table 1). Fitting the titration data to this model yields $\Delta G(0)$, the free energy difference between the states connected by the transition in the absence of denaturant, and m , the sensitivity of the transition to the addition of denaturant. Both hcyt *c*²⁹ and bcyt *c* show residue-dependent $\Delta G(0)$ and m values, confirming that both proteins unfold via stepwise mechanisms. Similarly to hcyt *c*, the transition midpoint observed at Met80 in bcyt *c* occurs at a relatively lower denaturant concentration, despite its greater $\Delta G(0)$, due to a greater m value, likely reflecting exposure of the large hydrophobic heme surface. However, while the increased stability of bcyt *c*, relative to hcyt *c*, is apparent in the $\Delta G(0)$ and m values for each transition, the increases are not equal for the transitions observed at different residues. The greatest increase in $\Delta G(0)$ and m is observed for the transition at Leu94, which is actually the least stable of the transitions in hcyt *c*. Along with the absence of the low denaturant intermediate, it is this differential stabilization of the individual transitions and their m values that results in the more coincident unfolding of bcyt *c*.

The UV/vis absorption at 695 is known to depend on a native Met80-heme interaction, while the absorption at 530 nm depends on more global interactions between the protein and the heme. Thus, following both of these absorptions as a function of denaturant concentration allowed us to deconvolute the effects of the individual mutations that differentiate bcyt *c* and hcyt *c* on folding events that occur at both low and high denaturant concentrations. The data suggests that K60G is responsible for virtually all of the differences in stability and folding between the diverged proteins, with T47S and T89G contributing little if anything to the differences.

Interestingly, the structures of bcyt *c*⁴⁴ and hcyt *c*⁴⁵ provide a likely explanation for the stabilization associated with K60G. Residue 60 is located at the N-terminus of the 60's helix (Figure 5), and such N-cap residues are known to have a particularly important effect on helix stability.^{46–49} In general, an N-cap Gly is more stabilizing than other residues such as Lys because the presence of a side chain can block solvation of the non-H-bonded amide group of the first turn of the helix^{46–48} or prevent its H-bonding with other donors.⁴⁹ Indeed, comparison of the structures of bcyt *c* and hcyt *c* reveals that mutation to Gly in the bovine variant introduces a crystallographically observable water molecule that H-bonds to the N-H backbone of residue Thr63 (Figure 5). In contrast, this H-bond donor is desolvated in the hcyt *c* protein due to its occlusion by the Lys side chain. It is also interesting to note that other stabilizing N-cap residues (Asp, Asn, Ser, or Thr)^{46–49} are present at position 60 in more than 70% of the reported cytochrome *c* sequences.⁵ Moreover, the structures of the tuna (PDB ID 5cyt) and yeast (PDB IDs 1yea and 2ycc) variants, which have an Asn or Asp at position 60, show a similar, presumably stabilizing H-bond between the side chain and the backbone of residue 63.

How might the increased stability of the N-terminus of the 60's helix stabilize the protein to unfolding at Leu94, which is in the middle of another helix (the C-terminal helix), more than it stabilizes the protein to denaturation at Leu68, which is at the C-terminus of the same helix? Interestingly, the structures reveal intimate interactions between the N-terminus of the 60's helix and the C-terminal helix. Specifically, the helices interact via extensive packing and a salt bridge between Glu61 and Lys99 (Figure 5). Similar interactions are known to underlie the specificity of helix dimerization in peptides and other proteins.^{50,51} The preferential stabilization of at least part of the C-terminal helix may be analogous to helix dimer formation in other proteins where limited helical formation facilitates dimerization and then complete helix formation after dimerization.^{52,53}

Both bcyt *c* and hcyt *c* appear to unfold via stepwise mechanisms; however, the mutation K60G makes the bovine variant more stable. Because this stability appears to reduce the population of a folding intermediate and to be more manifest at the least stable parts of the protein, it also causes the unfolding to be more homogeneous with more coincident transitions. While the detailed mechanism of this stabilization remains speculative, the residue-specific resolution of the C-D based IR probe is well suited for its further testing, both in a steady-state format to study equilibrium folding, and in time-resolved format to study folding kinetics. In addition, along with the large number of available cytochrome *c* sequences, both extant and extinct, the technique and its high structural and temporal resolution should allow for the further exploration of how protein stability, folding, and perhaps even function, were tailored by evolution.

Materials and Methods

Semisynthesis of bovine cytochrome *c*

Residue-specific deuterated bcyt *c* was semisynthesized as described previously for hcyt *c*.^{30,54} Briefly, (*methyl-d*₃)methionine ((*d*₃)methionine) or a 1:1 mixture of C_{δ1}-*d*₃ and C_{δ1}-*d*₃ labeled leucine isotopomers ((*d*₃)leucine) (Cambridge Isotopes) were incorporated during the Boc-based solid phase synthesis of C-terminal residues 66–104, according to standard procedures.^{55,56} The N-terminal fragment of residues 1–65, containing the covalently bound heme, was prepared by cyanogen bromide cleavage of bcyt *c* (Sigma). The full protein was generated by ligation of the labeled peptide to the 1–65 fragment and proceeded under anaerobic conditions in the presence sodium dithionite in 50 mM sodium phosphate, pH 7. The synthesized peptide, the 1–65 fragment, and the ligated protein were purified by HPLC and assayed by electrospray mass spectrometry. The product bcyt *c* was oxidized with bis(dipicolinato)cobaltate(III), purified by size exclusion chromatography (Sephadex G25) in 10 mM sodium acetate (pH 5.8), concentrated by filtration (10K MWCO, Amicon) to 6 mM, separated into 15–20 μL aliquots, lyophilized, and stored at –20 °C.

Sample preparation and FT IR data acquisition

Individual IR samples containing 6 mM bcyt *c* were prepared by adding solutions of GdnHCl (Ultra Pure, MP Biomedicals, LLC) in 100 mM sodium acetate (pH 6.2) to the aliquots of lyophilized bcyt *c*. Concentrations of GdnHCl solutions were verified by refractive index measurements. Solutions of 100 mM (*methyl-d*₃) methionine or C_{δ1}-*d*₃ leucine with varying amounts of GdnHCl were also prepared to account for any salt effects on the C-D absorptions. The amino acid or protein samples were placed in a liquid sample cell with CaF₂ windows and a 50 μm Teflon spacer. FT IR spectra were recorded on a N₂ purged Bruker Equinox 55 spectrometer equipped with a liquid nitrogen cooled MCT detector, as described previously.^{30,31} Proteo bcyt *c* and amino acids were used for acquisition of the reference interferogram for the protein and amino acid spectra, respectively. The absorption spectra were obtained with the OPUS software (Bruker) from

8192 averaged interferograms, and 3–6 absorption spectra were measured for each sample. Spectra were acquired at 4 cm^{-1} resolution, which is appropriate for condensed phase spectra,^{57,58} and were identical to those obtained with 2 cm^{-1} resolution. Spectra were analyzed using Matlab, version 14, with the curvefitting toolbox (Mathworks, Inc.). To approximate the background spectrum, a polynomial was fit to a $100\text{--}200\text{ cm}^{-1}$ spectral region, excluding a $\sim 30\text{ cm}^{-1}$ window around the C-D absorption band(s), and subtracted to yield baseline-corrected spectra.

Data analysis

The deuterated proteins in zero or 6.0 M GdnHCl were first analyzed to determine the absorption spectra for native (folded) and high denaturant conditions. The spectra were fit to the minimum number of Gaussian and/or quasi-Voigt functions (approximated by a linear combination of a Gaussian and Lorentzian functions) that were statistically justified using F-tests. The spectra obtained at all intermediate GdnHCl concentrations were fit to a superposition of the folded and high denaturant spectra (for the (d_3)methionine spectra the frequencies and linewidths were allowed to vary by $\pm 0.5\text{ cm}^{-1}$). This is justified as long as the particular transition is two state (which in each case was ultimately justified by the quality of the fits) due to the time scale inherent to IR spectroscopy (*i.e.* the spectra are not expected to change as a function of denaturant, only their relative contribution is expected to vary). The fraction of folded protein as a function of added GdnHCl was determined from the amplitude of the folded absorptions divided by the total amplitude of all absorptions, with each amplitude normalized by the relative intensities of the absorptions at zero and 6.0 M GdnHCl.

The data were then analyzed by the linear extrapolation method, which is commonly used in folding studies.^{43,59} The fraction of folded protein as a function of GdnHCl concentration

was fit to the equation: $\frac{a}{1 + \exp\left(\frac{-\Delta G(0) - m[\text{GdnHCl}]}{RT}\right)} + b$. This method assumes that ΔG of unfolding depends linearly on GdnHCl concentration and yields $\Delta G(0)$, the free energy of unfolding in the limit of zero GdnHCl; m , the dependence of ΔG on denaturant concentration; and

$c_{1/2} = \frac{\Delta G(0)}{m}$, the denaturant concentration at which the folded and unfolded protein are present at equal concentrations. IR data were acquired for three independent titration curves, the curves were individually fit, and the resulting parameter values were averaged and standard deviations calculated. The parameter a was found to vary between 0.86 and 0.94, indicating the presence of a small amount of unfolded protein even at low GdnHCl concentration, as observed previously for hcyt *c*.³⁴ The parameter b was used to correct for error in the offset, and was approximately zero in all cases.

Preparation and GdnHCl-induced unfolding of T47S, K60G, and T89G horse cyt c

The sequence of the hcyt *c* expression vector (kindly provided by Dr. Gary Pielak, University of North Carolina, Chapel Hill) was modified using QuikChange (Stratagene), and verified by DNA sequencing. The proteins were expressed in minimal media supplemented with 400 mg/L amino acids and $1\times$ Gibco MEM vitamin solution (Invitrogen) and purified as described previously.²⁸ Briefly, cells were lysed by sonication, the proteins were purified from the crude extract by ammonium sulfate precipitation (326 g/L), further purified by ion exchange chromatography (Macroprep S, Bio-Rad) and by reverse phase HPLC, and samples stored as described above.

UV/vis spectra were monitored as function of GdnHCl concentration with 60 μM hcyt *c* in 100 mM sodium acetate, pH 6.2 with a 1 cm quartz cuvette on a Cary300

Spectrophotometer. The absorptions at 530 or 695 nm were analyzed using the linear extrapolation method^{16,43} and fit to

$$A = \varepsilon_{UF} + \varepsilon_{UF'}[GdnHCl] + \frac{(\varepsilon_F + \varepsilon_{F'}[GdnHCl]) \exp \frac{-(\Delta G(0) + m[GdnHCl])}{RT}}{1 + \exp \frac{-(\Delta G(0) + m[GdnHCl])}{RT}}$$

where $\varepsilon_F + \varepsilon_{F'}[GdnHCl]$ and $\varepsilon_{UF} + \varepsilon_{UF'}[GdnHCl]$ are the extinction coefficient of the folded and unfolded protein, respectively, as a function of GdnHCl concentration. At least three independent titration curves were acquired and analyzed, and the average and standard deviation of the folding parameters were calculated.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by the National Science Foundation under Grant No. MCB 0346967 (to F.E.R.) and a graduate fellowship (to M.C.T.) and by the National Institutes of Health (GM059380 to P.E.D.).

References

1. Moore, GR.; Pettigrew, GW. Cytochromes c. Evolutionary, Structural and Physiochemical Aspects. Berlin: Springer-Verlag; 1990.
2. Scott, RA.; Mauk, AG., editors. Cytochrome c: A Multidisciplinary Approach. Sausalito, CA: University Science Books; 1995.
3. Hoang L, Maity H, Krishna MM, Lin Y, Englander SW. Folding units govern the cytochrome c alkaline transition. *J. Mol. Biol* 2003;331:37–43. [PubMed: 12875834]
4. Dickerson RE. The structures of cytochrome c and the rates of molecular evolution. *J. Mol. Evol* 1971;1:26–45. [PubMed: 4377446]
5. Banci L, Bertini I, Rosato A, Varani G. Mitochondrial cytochromes c: a comparative analysis. *J. Biol. Inorg. Chem* 1999;4:824–837. [PubMed: 10631615]
6. Knapp JA, Pace CN. Guanidine hydrochloride and acid denaturation of horse, cow, and *Candida krusei* cytochromes c. *Biochemistry* 1974;13:1289–1294. [PubMed: 4360785]
7. Ferguson-Miller S, Brautigan DL, Margoliash E. Correlation of the kinetics of electron transfer activity of various eukaryotic cytochromes c with binding to mitochondrial cytochrome c oxidase. *J. Biol. Chem* 1976;251:1104–1115. [PubMed: 2600]
8. McLendon G, Smith M. Equilibrium and kinetic studies of unfolding of homologous cytochromes c. *J. Biol. Chem* 1978;253:4004–4008. [PubMed: 206558]
9. Fredericks ZL, Pielak GJ. Exploring the interface between the N-terminal and C-terminal helices of cytochrome c by random mutagenesis within the C-terminal helix. *Biochemistry* 1993;32:929–936. [PubMed: 8380714]
10. Pielak GJ, Auld DA, Beasley JR, Betz SF, Cohen DS, Doyle DF, Finger SA, Fredericks ZL, Hilgenwillis S, Saunders AJ, Trojak SK. Protein thermal-denaturation, side-chain models, and evolution - amino-acid substitutions at a conserved helix-helix interface. *Biochemistry* 1995;34:3268–3276. [PubMed: 7880822]
11. Yuan X, Hawkrigde FM, Chlebowsky JF. Thermodynamic and kinetic studies of cytochrome c from different species. *J. Electroanal. Chem* 1993;350:29–42.
12. Fisher A, Shi Y, Ritter A, Ferretti JA, Perez-Lamboy G, Shah M, Shiloach J, Taniuchi H. Functional correlation in amino acid residue mutations of yeast iso-2-cytochrome c that is consistent with the prediction of the concomitantly variable codon theory in cytochrome c evolution. *Biochem. Genet* 2000;38:181–200. [PubMed: 11091908]

13. Filosa A, Ismail AA, English AM. FTIR-monitored thermal titration reveals different mechanisms for the alkaline isomerization of tuna compared to horse and bovine cytochromes *c*. *J. Biol. Inorg. Chem* 1999;4:717–726. [PubMed: 10631603]
14. Filosa A, English AM. Probing local thermal stabilities of bovine, horse, and tuna ferricytochromes *c* at pH 7. *J. Biol. Inorg. Chem* 2000;5:448–454. [PubMed: 10968615]
15. Filosa A, Wang Y, Ismail AA, English AM. Two-dimensional infrared correlation spectroscopy as a probe of sequential events in the thermal unfolding of cytochromes *c*. *Biochemistry* 2001;40:8256–8263. [PubMed: 11444971]
16. Moza B, Qureshi SH, Ahmad F. Equilibrium studies of the effect of difference in sequence homology on the mechanism of denaturation of bovine and horse cytochromes-*c*. *Biochem. Biophys. Acta* 2003;1646:49–51. [PubMed: 12637011]
17. Rahaman H, Khan KA, Hassan I, Wahid M, Singh SB, Singh TP, Moosavi-Movahedi AA, Ahmad F. Sequence and stability of the goat cytochrome *c*. *Biophys. Chem* 2008;138:23–28. [PubMed: 18814948]
18. Myer YP. Conformation of cytochromes. 3. Effect of urea, temperature, extrinsic ligands, and pH variation on the conformation of horse heart ferricytochrome *c*. *Biochemistry* 1968;7:765–776. [PubMed: 5689422]
19. Tsong TY. Detection of three kinetic phases in the thermal unfolding of ferricytochrome *c*. *Biochemistry* 1973;12:2209–2214. [PubMed: 4351055]
20. Jeng MF, Englander SW, Elöve GA, Wand AJ, Roder H. Structural description of acid-denatured cytochrome *c* by hydrogen exchange and 2D NMR. *Biochemistry* 1990;29:10433–10437. [PubMed: 2176867]
21. Rossell FI, Ferrer JC, Mauk AG. Proton-linked protein conformational switching: definition of the alkaline conformational transition of yeast iso-1-ferricytochrome *c*. *J. Am. Chem. Soc* 1998;120:11234–11245.
22. Hostetter DR, Weatherly GT, Beasley JR, Bortone K, Cohen DS, Finger SA, Hardwidge P, Kakouras DS, Saunders AJ, Trojak SK, Waldner JC, Pielak GJ. Partially formed native tertiary interactions in the A-state of cytochrome *c*. *J. Mol. Biol* 1999;289:639–644. [PubMed: 10356334]
23. Latypov RF, Cheng H, Roder NA, Zhang J, Roder H. Structural characterization of an equilibrium unfolding intermediate in cytochrome *c*. *J. Mol. Biol* 2006;357:1009–1025. [PubMed: 16473367]
24. Maity H, Maity M, Englander SW. How cytochrome *c* folds, and why: submolecular foldon units and their stepwise sequential stabilization. *J. Mol. Biol* 2004;343:223–233. [PubMed: 15381432]
25. Tadesse L, Nazarbaghi R, Walters L. Isotopically enhanced infrared spectroscopy: a novel method for examining secondary structure at specific sites in conformationally heterogeneous peptides. *J. Am. Chem. Soc* 1991;113:7036–7037.
26. Barth A. Selective monitoring of 3 out of 50,000 protein vibrations. *Biopolymers* 2002;67:237–241. [PubMed: 12012437]
27. Decatur SM. Elucidation of residue-level structure and dynamics of polypeptides via isotope-edited infrared spectroscopy. *Acc. Chem. Res* 2006;39:169–175. [PubMed: 16548505]
28. Chin JK, Jimenez R, Romesberg FE. Direct observation of protein vibrations by selective incorporation of spectroscopically observable carbon-deuterium bonds in cytochrome *c*. *J. Am. Chem. Soc* 2001;123:2426–2427. [PubMed: 11456893]
29. Sagle LB, Zimmermann J, Dawson PE, Romesberg FE. A high-resolution probe of protein folding. *J. Am. Chem. Soc* 2004;126:3384–3385. [PubMed: 15025440]
30. Sagle LB, Zimmermann J, Dawson PE, Romesberg FE. Direct and high resolution characterization of cytochrome *c* equilibrium folding. *J. Am. Chem. Soc* 2006;128:14232–14233. [PubMed: 17076477]
31. Thielges MC, Case DA, Romesberg FE. Carbon-deuterium bonds as probes of dihydrofolate reductase. *J. Am. Chem. Soc* 2008;130:6597–6603. [PubMed: 18412341]
32. Chin JK, Jimenez R, Romesberg F. Protein dynamics and cytochrome *c*: Correlations between ligand vibrations and redox activity. *J. Am. Chem. Soc* 2002;124:1846–1847. [PubMed: 11866585]

33. Weinkam P, Zimmermann J, Sagie LB, Matsuda S, Dawson PE, Wolynes PG, Romesberg FE. Characterization of alkaline transitions in ferricytochrome c using carbon-deuterium infrared probes. *Biochemistry* 2008;47:13470–13480. [PubMed: 19035653]
34. Sagie LB, Zimmermann J, Matsuda S, Dawson PE, Romesberg FE. Redox-coupled dynamics and folding in cytochrome c. *J. Am. Chem. Soc* 2006;128:7909–7915. [PubMed: 16771505]
35. Kinnaman CS, Cremeens ME, Romesberg FE, Corcelli SA. Infrared line shape of an alpha-carbon deuterium-labeled amino acid. *J. Am. Chem. Soc* 2006;128:13334–13335. [PubMed: 17031927]
36. Cremeens ME, Fujisaki H, Zhang Y, Zimmermann J, Sagie LB, Matsuda S, Dawson PE, Straub JE, Romesberg FE. Efforts toward developing direct probes of protein dynamics. *J. Am. Chem. Soc* 2006;128:6028–6029. [PubMed: 16669659]
37. Roder H, Elöve GA, Englander SW. Structural characterization of folding intermediates in cytochrome c by H-exchange labelling and proton NMR. *Nature* 1988;335:700–704. [PubMed: 2845279]
38. Bai YW, Sosnick TR, Mayne L, Englander SW. Protein-folding intermediates - native-state hydrogen-exchange. *Science* 1995;269:192–197. [PubMed: 7618079]
39. Ferri T, Poscia A, Ascoli F, Santucci R. Direct electrochemical evidence for an equilibrium intermediate in the guanidine-induced unfolding of cytochrome c. *BBA* 1996;1298:102–108. [PubMed: 8948494]
40. Xu Y, Mayne L, Englander SW. Evidence for an unfolding and refolding pathway in cytochrome c. *Nat. Struct. Biol* 1998;5:774–778. [PubMed: 9731770]
41. Maity H, Maity M, Krishna MM, Mayne L, Englander SW. Protein folding: the stepwise assembly of foldon units. *Proc. Natl. Acad. Sci. U.S.A* 2005;102:4741–4746. [PubMed: 15774579]
42. Krishna MM, Maity H, Rumbley JN, Lin Y, Englander SW. Order of steps in the cytochrome c folding pathway: evidence for a sequential stabilization mechanism. *J. Mol. Biol* 2006;359:1410–1419. [PubMed: 16690080]
43. Greene RFJ, Pace CN. Urea and guanidine hydrochloride denaturation of ribonuclease, lysozyme, alpha-chymotrypsin, and beta-lactoglobulin. *J. Biol. Chem* 1974;249:5388–5393. [PubMed: 4416801]
44. Mirkin N, Jaconic J, Stojanoff V, Moreno A. High resolution X-ray crystallographic structure of bovine heart cytochrome c and its application to the design of an electron transfer biosensor. *Proteins* 2008;70:83–92. [PubMed: 17634981]
45. Bushnell GW, Louie GV, Brayer GD. High-resolution three-dimensional structure of horse heart cytochrome c. *J. Mol. Biol* 1990;214:585–595. [PubMed: 2166170]
46. Harpaz Y, Elmasry N, Fersht AR, Henrick K. Direct observation of better hydration at the N terminus of an alpha-helix with glycine rather than alanine as the N-cap residue. *Proc. Natl. Acad. Sci. U.S.A* 1994;91:311–315. [PubMed: 8278384]
47. Chakraborty A, Doig AJ, Baldwin RL. Helix capping propensities in peptides parallel those in proteins. *Proc. Natl. Acad. Sci. U.S.A* 1993;90:11332–11336. [PubMed: 8248248]
48. Serrano L, Fersht AR. Capping and alpha-helix stability. *Nature* 1989;342:296–299. [PubMed: 2812029]
49. Chen YW, Fersht AR. Stability and solvation of Thr/Ser to Ala and Gly mutations at the N-cap of alpha-helices. *FEBS Lett* 1994;347:304–309. [PubMed: 8034023]
50. Vinson CR, Hai T, Boyd SM. Dimerization specificity of the leucine zipper-containing bZIP motif on DNA binding: prediction and rational design. *Genes Dev* 1993;7:1047–1058. [PubMed: 8504929]
51. Pervushin K, Billeter M, Siegal G, Wüthrich K. Structural role of a buried salt bridge in the 434 repressor DNA-binding domain. *J. Mol. Biol* 1996;264:1002–1012. [PubMed: 9000626]
52. Kammerer RA, Schulthess T, Landwehr R, Lustig A, Engel J, Aepli U. An autonomous folding unit mediates the assembly of two-stranded coiled coils. *Proc. Natl. Acad. Sci. U.S.A* 1998;95:13419–13424. [PubMed: 9811815]
53. Kammerer RA, Jaravine VA, Frank S, Schulthess T, Landwehr R, Lustig A, Garcia-Echeverria C, Alexandrescu AT, Engel J, Steinmetz MO. An intrahelical salt bridge within the trigger site stabilizes the GCN4 leucine zipper. *J. Biol. Chem* 2001;276:13685–13688. [PubMed: 11134036]

54. Wallace CJ, Offord RE. The semisynthesis of fragments corresponding to residues 66–104 of horse heart cytochrome *c*. *Biophys. J* 1979;179:169–182.
55. Wallace CJ, Mascagni P, Chait BT, Collawn JF, Paterson Y, Proudfoot AE, Kent SB. Substitutions engineered by chemical synthesis at three conserved sites in mitochondrial cytochrome *c*. Thermodynamic and functional consequences. *J. Biol. Chem* 1989;264:15199–15209. [PubMed: 2475497]
56. Hackeng TM, Griffin JH, Dawson PE. Protein synthesis by native chemical ligation: expanded scope by using straightforward methodology. *Proc. Natl. Acad. Sci. U.S.A* 1999;96:10068–10073. [PubMed: 10468563]
57. Griffiths, PR.; deHaseth, J. *Fourier Transform Infrared Spectrometry*. New York: John Wiley & Sons; 1986.
58. Smith, BC. *Fundamentals of Fourier Transform Infrared Spectroscopy*. Boca Raton: CRC Press; 1995.
59. Morar AS, Kakouras D, Young GB, Boyd J, Pielak GJ. Expression of ¹⁵N-labeled eukaryotic cytochrome *c* in *Escherichia coli*. *J. Biol. Inorg. Chem* 1999;4:220–222. [PubMed: 10499094]
60. Petersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. UCSF Chimera: a visualization system for exploratory research and analysis. *J. Comput. Chem* 2004;25:1605–1612. [PubMed: 15264254]

Figure 1.

Structure of hcyt *c* (PDB ID 1hrc) showing motifs and residues characterized (image produced using the UCSF Chimera package from the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco60). The residues used to probe unfolding transitions are shown in blue. The 60's and C-terminal helices as well as the hcyt *c* residues that are altered in bcyt *c* are shown in gold.

Figure 2.

IR spectra and fits. The asymmetric stretching absorptions of $(d_3)\text{Leu68}$ and $(d_3)\text{Leu94}$, and the symmetric stretching absorptions of $(d_3)\text{Met80}$, are shown at zero (native state), intermediate, and high denaturant (unfolded state) concentrations. See text for details.

Figure 3. Residue-specific GdnHCl-induced equilibrium unfolding of bcyt c; (A) (d₃)Met80, (B) (d₃)Leu68, (C) (d₃)Leu94, and (D) superposition of (d₃)Met80 (filled circles), (d₃)Leu68 (circles), and (d₃)Leu94 (squares).

Figure 4. GdnHCl-induced unfolding of hcyt *c* and point mutants. (A) 530 nm and (B) 695 nm absorption as a function of GdnHCl concentration for hcyt *c* M65L (triangle) and the T47S/M65L (square), K60G/M65L (filled circle), and T89G/M65L (circle) mutants.

Figure 5. Proposed mechanism of divergence. The unique water molecule observed in bcyt *c* (PDB ID 2b4z), and excluded by the lysine side chain in hcyt *c* (PDB ID 1hrc), is shown as a light blue sphere.

Table 1LEM analysis of site-specific unfolding data for bcyt *c* and hcyt *c*.^a

Protein	residue	$\Delta G(0)^b$ (kcal/mol)	m (kcal/M/mol)	$c_{1/2}$ (M)
bcyt <i>c</i>	(<i>d</i> ₃)Leu68 ^d	8.8 ± 0.6	3.3 ± 0.2	2.65 ± 0.01
	(<i>d</i> ₃)Leu94 ^d	10.4 ± 1.0	3.9 ± 0.4	2.63 ± 0.01
	(<i>d</i> ₃)Met80 ^d	11.0 ± 0.8	4.4 ± 0.3	2.51 ± 0.02
hcyt <i>c</i> ^c	(<i>d</i> ₃)Leu68 ^d	5.6 ± 0.5	2.3 ± 0.3	2.37 ± 0.05
	(<i>d</i> ₃)Leu94 ^d	4.0 ± 0.2	1.7 ± 0.1	2.31 ± 0.05
	(<i>d</i> ₃)Met80 ^e F $\rightleftharpoons$ I	5.3 ± 0.3	2.6 ± 0.3	2.0 ± 0.2
	(<i>d</i> ₃)Met80 ^e F $\rightleftharpoons$ U	7.5 ± 0.5	3.2 ± 0.3	2.31 ± 0.05

^aSee text and Supplementary Material for details.^bCalculated from $\Delta G(0) = m \times c_{1/2}$.^cData from Ref. 30.^dFit to two-state equilibrium.^eFit to three-state equilibrium between folded (F), intermediate (I) and unfolded (U) states.

Table 2LEM analysis of UV/Vis transitions.^a

530 nm	$\Delta G(0)^b$ (kcal/mol)	m (kcal/M/mol)	$c_{1/2}$ (M)
hcyt <i>c</i>	8.5 ± 0.08	3.5 ± 0.03	2.47 ± 0.01
bcyt <i>c</i>	9.5 ± 0.2	3.7 ± 0.07	2.60 ± 0.01
hcyt <i>c</i> (M65L)	8.2 ± 0.2	3.6 ± 0.1	2.28 ± 0.02
hcyt <i>c</i> (T47S/M65L)	8.5 ± 0.2	3.8 ± 0.06	2.26 ± 0.02
hcyt <i>c</i> (K60G/M65L)	9.0 ± 0.4	3.7 ± 0.2	2.45 ± 0.04
hcyt <i>c</i> (T89G/M65L)	8.3 ± 0.2	3.7 ± 0.08	2.22 ± 0.03
695 nm	$\Delta G(0)^b$ (kcal/mol)	m (kcal/M/mol)	$c_{1/2}$ (M)
hcyt <i>c</i>	8.9 ± 0.2	3.6 ± 0.07	2.48 ± 0.01
bcyt <i>c</i>	9.7 ± 0.4	3.7 ± 0.2	2.62 ± 0.02
hcyt <i>c</i> (M65L)	8.8 ± 0.4	3.8 ± 0.2	2.30 ± 0.02
hcyt <i>c</i> (T47S/M65L)	9.0 ± 0.2	3.9 ± 0.07	2.27 ± 0.02
hcyt <i>c</i> (K60G/M65L)	9.9 ± 0.4	4.0 ± 0.1	2.50 ± 0.04
hcyt <i>c</i> (T89G/M65L)	8.7 ± 0.5	3.9 ± 0.2	2.23 ± 0.03

^a See text and Supplementary Material for details.^b Calculated from $\Delta G(0) = m \times c_{1/2}$.