THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

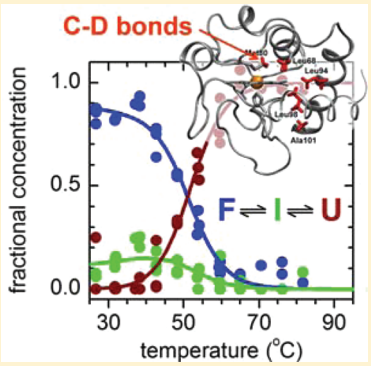
Carbon−Deuterium Bonds as Probes of Protein Thermal Unfolding

Wayne Yu, Phillip E. Dawson, Jörg Zimmermann,* and Floyd E. Romesberg*

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

Supporting Information

ABSTRACT: We report a residue-specific characterization of the thermal unfolding mechanism of ferric horse heart cytochrome *c* using C−D bonds site-specifically incorporated at residues dispersed throughout three different structural elements within the protein. As the temperature increases, Met80 first dissociates from the heme center, and the protein populates a folding intermediate before transitioning to a solvent exposed state. With further increases in temperature, the C-terminal helix frays and then loses structure along with the core of the protein. Interestingly, the data also reveal that the state populated at high temperature retains some structure and possibly represents a molten globule. Elucidation of the detailed unfolding mechanism and the structure of the associated molten globule, both of which represent challenges to conventional techniques, highlights the utility of the C−D technique.



INTRODUCTION

Protein folding is the final step of transferring genetic information into biological function, and understanding its detailed mechanism is a central goal of biophysics. Of particular interest is differentiating between the varying contributions of stepwise and funnel-like folding. A variety of techniques have been applied to characterize the folding process, for example, as a function of added denaturant.[1] However, while most of these techniques are generally sensitive to global changes in structure, they typically lack the site-specific resolution required to elucidate detailed folding mechanisms. While NMR spectroscopy is inherently site-specific, its use to study protein folding is generally limited by reduced dispersion and exchange broadening. Perhaps the most ideal technique is IR spectroscopy, which, like NMR, is bond-specific but has an inherently higher time resolution. However, the IR spectrum of proteins is prohibitively congested. As part of an effort to develop a general and nonperturbative approach to characterizing protein dynamics, we have been exploring the use of proteins with specific carbon−hydrogen bonds replaced with carbon−deuterium (C−D) bonds. C−D bonds are nonperturbative, and the frequencies of their stretching absorptions occur in an otherwise transparent region of the IR spectrum (~2200 cm^{-1}). We have shown that individual C−D bonds may be characterized when incorporated into a variety of different proteins, including cytochrome *c*,[2−7] dihydrofolate reductase,[8,9] and an SH3 domain.[10] Moreover, we have used the C−D absorptions to follow protein folding under both equilibrium and time-resolved conditions.[11]

The characterization of thermal (un)folding is particularly attractive as the process may be conveniently studied under both equilibrium and kinetic conditions, which should help differentiate stepwise and funnel-like folding mechanisms. There is a large body of work characterizing the thermal unfolding of horse heart ferricytochrome *c* (cyt *c*) under equilibrium conditions using UV/vis,[12−16] circular dichroism,[13,17,18] NMR,[19] and FT IR spectroscopy.[20,21] As the temperature is increased, quenching of a weak charge-transfer absorption band at 695 nm occurs with a midpoint temperature, T_m, of ~58 °C, and is interpreted as the formation of an intermediate via dissociation of the Met80 ligand from the iron center.[19,22] A transition observed via far-UV circular dichroism occurs with a T_m of ~70 °C and has been interpreted as the loss of global structure. On the basis of IR amide band deconvolutions, Filosa et al. concluded that turns and extended chains of the protein lose structure prior to some subset of α-helices, which in turn precede the melting of more stable extended chains and helices.[12,23] Unfolding of cyt *c* triggers aggregation into amyloid-like fibers,[14,20] but aggregation is diminished in the presence of high ionic strength or chemical denaturants.[14,16] Furthermore, UV/vis, circular dichroism, and differential scanning calorimetry studies have suggested that unfolding follows a two-state process in the presence of high but not low concentrations of guanidine hydrochloride (GdnHCl).[13,24]

While these studies have been illuminating, the techniques employed do not provide high structural resolution, leaving the details of the process unclear. To elucidate these details, we now report the thermal denaturation of cyt *c* with C−D bonds site-specifically incorporated throughout various secondary structural elements (Figure 1). The results reveal both localized and cooperative transitions that, as the temperature is increased, eventually culminate in the population of an apparent molten globule state, that itself then appears to unfold via multiple steps.

Received: April 12, 2012
Revised: May 11, 2012
Published: May 24, 2012

   dx.doi.org/10.1021/jp303521t | *J. Phys. Chem. B* 2012, 116, 6397−6403

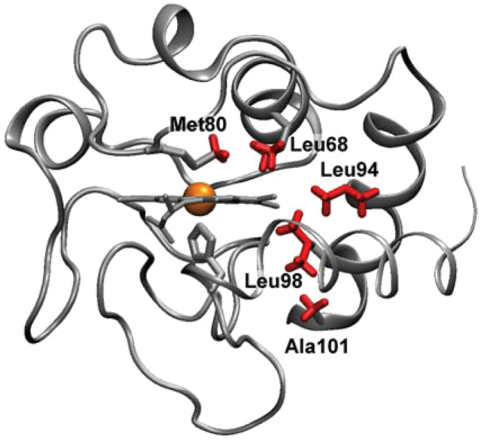


Figure 1. Structure of cyt *c* (PDB 1GIW) showing the heme iron (orange) and side-chains of deuterated residues. The positions where C−D bonds were incorporated are shown in red. The side chain of His18 is also shown.

EXPERIMENTAL METHODS

Sample Preparation. Site-selectively deuterated cyt *c* variants were prepared via semisynthesis, as described previously.[4,6,25,26] Briefly, the homoserine lactone peptide fragment corresponding to residues 1−65 of cyt *c*, with a covalently bound heme, was generated via cyanogen bromide cleavage of the intact protein (Sigma-Aldrich, St. Louis, MO) by adding 5% w/v cyanogen bromide to 2.5% w/v cyt *c* dissolved in thoroughly degassed 30% formic acid (aqueous) followed by stirring for up to four hours under an argon atmosphere at room temperature. The crude reaction mixture was loaded onto a Sephadex G-50 gel exclusion column equilibrated with 7% formic acid (aqueous) to remove excess cyanogen bromide and further purified by reverse phase HPLC using a preparative column (Phenomenex Jupiter C-18, 10 μm particle size, 300 Å pore size). Fractions with greater than 90% purity as determined by ESI-MS were combined, lyophilized, and stored at 4 °C.

The peptide corresponding to residues 66−104 of cyt *c* was synthesized via solid phase peptide synthesis employing Boc-chemistry[27] and Boc-L-Glu(OcHx)-PAM resin, during which one specifically deuterated amino acid (Cambridge Isotope Laboratories, Boston, MA) was incorporated at the desired position. Boc-protected amino acids were used in 2- to 4-fold excess at each step, and the coupling times were 15−30 min. After Boc-deprotection of the last residue, the resin was neutralized with DIEA, washed with dichloromethane, and dried under vacuum overnight. The dried resin was cleaved with 10 mL of HF and 0.5 mL of anisole in an ice-bath for 1 h. The cleaved peptide was precipitated in ice-cold diethyl ether and filtered before being resolubilized in 27% acetonitrile and 0.1% TFA. The crude product was inspected by ESI-MS to confirm the presence of the correct peptide and then purified by reverse-phase HPLC as described above. Fractions with greater than 90% purity as determined by ESI-MS were pooled, lyophilized, and stored at 4 °C.

Coupling the N-terminal homoserine lactone and the C-terminal synthetic peptide was achieved by dissolving the C- and N-terminal peptides (2.3:4) in degassed 50 mM sodium phosphate buffer, pH 7, and 10 mM sodium dithionate. The reaction was sealed under argon and stirred at room temperature for 24 h. The resulting semisynthetic cyt *c* was purified by reverse-phase HPLC and gel filtration chromatography (Sephadex G-25). Fractions with greater than 95% purity by ESI-MS were combined, lyophilized, oxidized with a saturated, aqueous solution of bis(dipicolinato)cobaltate (III), and desalted over a 25 mL Sephadex G25 column equilibrated with 5 mM sodium phosphate, pH 7. The sample was concentrated using an Amicon Ultra-4 centrifugal filter unit, divided in single-use aliquots, lyophilized, and stored at −20 °C. The semisynthetic cyt *c* is identical to the wild type cyt *c* except at residue 65, which is a homoserine instead of the native methionine. This mutation has been shown to be inconsequential to the biological function of the protein.[25]

FT IR Measurements and Data Analysis. Lyophilized samples were dissolved in aqueous buffer containing 100 mM sodium phosphate pH 7, 100 mM NaCl, and 1 M GdnHCl. The final protein concentration was 4 mM as confirmed by the absorbance at 530 nm. The sample was allowed to equilibrate for 10 minutes and then loaded into a temperature-controllable demountable liquid cell with CaF_2 windows and a 75 μm Teflon spacer (Harrick Scientific Products, Inc., model TFC-M13, connected to a Watlow Series 989 1/8 DIN temperature controller). FT IR spectra were collected using a Bruker Equinox 55 FT IR spectrometer equipped with a liquid nitrogen cooled MCT detector, purged with dry nitrogen gas to remove atmospheric carbon dioxide and water vapor. Each interferogram was the average of 10 000 scans with 4 cm^{-1} resolution and was Fourier transformed using Blackman−Harris 3-term apodization. For automated spectra collection, both the spectrometer and the temperature cell were controlled using Labview (National Instruments Corp.). Spectra of proteo cyt *c* (Sigma) were obtained under identical conditions.

FT IR spectra were acquired between 25 and 90 at 5 °C intervals. Because of the low intensity of C−D stretches, raw FT IR spectra need to be carefully background corrected to minimize spectral distortion due to background variations.[28] After autosubtracting the corresponding proteo spectrum from the deuterium-labeled spectrum in the OPUS program, the resulting spectra excluding the C−D absorption band were approximated by an *n*th order polynomial, which was then subtracted from the data to yield the baseline-corrected spectra using Matlab (Mathworks, Inc.). Gaussian(s) were fit to the baseline-corrected spectra to approximate the C−D absorption, with the number of Gaussians determined by an F-test at the 99% confidence level (see Supporting Information, Figures S4−S7) or in accordance to previous reports.[4] The sum of Gaussian(s) used to fit the room temperature C−D absorptions was assigned to the spectrum of the folded state, and the sum of Gaussian(s) used to fit the high temperature C−D absorption was assigned to that of the high-temperature state. All spectra at intermediate temperatures were then deconvoluted into linear combinations of the limiting spectra of folded and high temperature states, varying only their respective intensities. This yielded adequate fits for all labeled proteins except for (d_3)Met80, where an intermediate state was present, which was fit with a single Gaussian as previously reported.[4]

Temperature-dependent absorption intensities, $A_i(T)$, of the folded, intermediate, and high-temperature species were calculated assuming a three-state equilibrium (folded, intermediate, and high-temperature state) for (d_3)Met80, and a two-state equilibrium (folded and high-temperature state) for all other proteins, according to

$$A_i(T) = a_i \times \frac{\exp[\Delta H^\circ_{Si\rightarrow H}(T^{-1} - {T_{m,i}}^{-1})/k_B]}{\sum_j^N \exp[\Delta H^\circ_{Si\rightarrow H}(T^{-1} - {T_{m,i}}^{-1})/k_B]}$$

where a_i is the temperature-independent absorbance of the ith species at 100% population, $\Delta H^\circ_{Si\rightarrow H}$ is the enthalpy of unfolding, and $T_{m,i} = \Delta H^\circ_{Si\rightarrow H}/\Delta S^\circ_{Si\rightarrow H}$ is the midpoint temperature of unfolding from state i, k_B is the Boltzmann constant, and T denotes temperature. The Gibbs free energy of unfolding is defined as $\Delta G^\circ_{Si\rightarrow H} = \Delta H^\circ_{Si\rightarrow H} - T\Delta S^\circ_{Si\rightarrow H}$, and we assume that $\Delta H^\circ_{Si\rightarrow H}$ and $\Delta S^\circ_{Si\rightarrow H}$ are temperature independent in the temperature range of interest. $\Delta H^\circ_{Si\rightarrow H}$ and $T_{m,i}$ were determined by least-squares fits to the thermal unfolding plots using Matlab (Mathworks, Inc.).

UV/Vis, Fluorescence, and Circular Dichroism Spectroscopy. UV/vis absorption spectra were acquired with a Cary300 spectrometer (Varian, Inc.). Temperature-dependent Trp fluorescence spectra were collected between 25 and 93 in 4 °C intervals on a photon counting spectrofluorimeter (PC1, ISS, Inc.). The signal intensity integrated between 300 and 400 nm as a function of temperature was fit to a sigmoidal function and two linear functions to correct for signal changes at low and high temperature. Far-UV circular dichroism spectra were recorded between 25 and 93 in 4 °C intervals on an AVIV 62DS spectrometer using a quartz sample cell with 2 mm path length (0.5 nm step size, 3 s integration per step, and average of 3 scans). The mean residue ellipticity, $[\theta]_{mrw}$, was calculated from $[\theta]_{mrw} = \theta/(206c)$, where θ is the raw ellipticity and c is the sample concentration. Sample concentrations, determined from the UV/vis spectrum of each sample, were 5−10 μM. The signal intensity integrated between 215 and 230 nm as a function of temperature was fit to a sigmoidal function. All experiments were performed in triplicate to yield averages and standard deviations for T_m and ΔH°.

RESULTS AND DISCUSSION

Cyt c is known to irreversibly aggregate at high temperatures in aqueous buffer,[20] but aggregation can be minimized by the addition of salt or chemical denaturants.[14,16] Thus, to extend the accessible temperature range, spectra were acquired in the presence of 1 M GdnHCl. Under these conditions, cyt c is still mostly folded at room temperature (except for ~10% population of a folding intermediate, see below). UV/vis and FT IR spectra confirmed that the presence of 1 M GdnHCl eliminates irreversible aggregation of cyt c up to at least 90 °C and that the temperature-induced unfolding is fully reversible (see Figures S8 and S9, Supporting Information). To assess the effect of denaturant on protein stability, we monitored the temperature-induced unfolding of cyt c with far-UV circular dichroism and fluorescence spectroscopy (Figure 2). Using far-UV circular dichroism, the midpoint for temperature-induced unfolding was determined to be 59.5 ± 1.4 °C (Table 1). Thus, the denaturant reduces T_m for global unfolding by ~10 °C.[17,23] Interestingly, the unfolding transition monitored by fluorescence of the lone Trp59, which in the folded protein is quenched due to photoenergy transfer to the heme, shows a midpoint for recovery of Trp59 fluorescence of 65.3 ± 0.3 °C, significantly higher than the T_m associated to loss of secondary structure measured via circular dichroism. This suggests that the protein remains in a compact, collapsed state after at least a significant part of the helices have unfolded.

To generate a higher resolution view of the unfolding process, we synthesized five different cyt c variants with

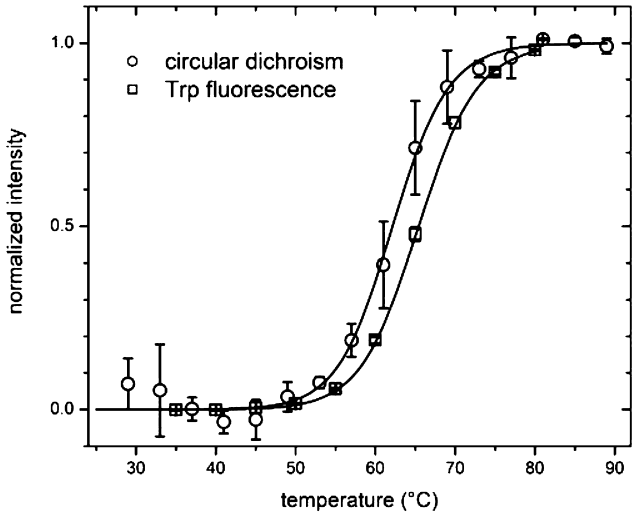


Figure 2. Temperature-induced change in the circular dichroism spectrum (circles) and Trp fluorescence intensity (squares). Lines are fits of the data to sigmoidal functions.

Table 1. Thermodynamic Fit Parameters

		T_m (°C)	ΔH° (kcal mol^{-1})	ΔS° (cal mol^{-1} K^{-1})
(d_7)Leu68		58.3 ± 0.4	47 ± 9	142 ± 28
(d_3)Met80	F↔H[a]	51.0 ± 1.1	47 ± 1	142 ± 6
	F↔I$_1$[a]	n.d.[b]	3.1 ± 0.3	6.6 ± 0.6
(d_7)Leu94		57.4 ± 1.0	41 ± 4	124 ± 14
(d_3)Leu98		57.5 ± 1.2	36 ± 2	109 ± 8
(d_3)Ala101		53.0 ± 0.9	46 ± 4	141 ± 15
circular dichroism		59.5 ± 1.4	52 ± 6	155 ± 14
fluorescence		65.3 ± 0.3	62 ± 1	182 ± 4

[a]Data fit to three-state equilibrium (F, folded state; H, high temperature state; I$_1$, intermediate state). [b]Not determined. Because of their similar temperature dependencies, if not replaced by the other species, both F and I$_1$ would coexist until very high temperatures. For fitting purposes, a value of 200 °C was used, but this value is so poorly defined as to be physically meaningless.

($C_{\gamma,\delta1,\delta2}$-$d_7$) leucine incorporated at Leu68 ((d_7)Leu68) or Leu94 ((d_7)Leu94), (C_ε-d_3) methionine incorporated at Met80 ((d_3)Met80), a 1:1 mixture of ($C_{\delta1}$-d_3) and ($C_{\delta2}$-d_3) labeled leucine isotopomers incorporated at Leu98 ((d_3)Leu98), or (C_β-d_3) alanine incorporated at Ala101 ((d_3)Ala101) (Figure 1). Both d_3 and d_7 variants of leucine were included to explore the potential increases in sensitivity afforded by the additional C−D bonds. Each deuterated variant shows symmetric and asymmetric C−D stretching absorptions between 2100 and 2300 cm^{-1}. The most intense of these absorptions, i.e., the asymmetric stretches for (d_7)Leu68, (d_7)Leu94, (d_3)Leu98, and (d_3)Ala101 and the symmetric stretch for (d_3)Met80, were used to characterize the folding process (Figure 3).

Cyt c provides two axial iron ligands, Met80 and His18 (Figure 1), with the former being of particular importance to both folding and function.[29−31] Previously we demonstrated that cyt c populates a structurally distinct folding intermediate during chemical or alkaline induced denaturation, and that both intermediates involve disruption of the Met80-Fe bond and reorganization of the heme pocket involving the Met80 loop.[4,5] Moreover, a variety of lower resolution techniques have suggested that a similar intermediate might be populated

dx.doi.org/10.1021/jp303521t | J. Phys. Chem. B 2012, 116, 6397−6403

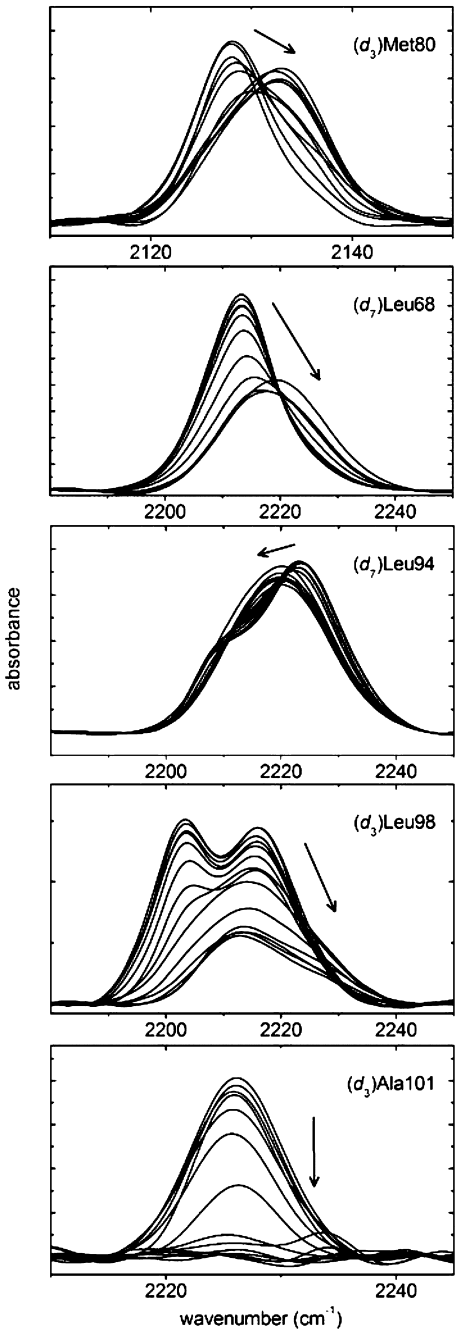


Figure 3. Background-corrected FT IR spectra for the different deuterium-labeled proteins between 25 and 90 °C. The arrows indicate the direction of the spectral change from low to high temperature.

during thermal unfolding.[12,17,19,20,23,32] In agreement with our previous work,[4,6] we found that two Gaussian functions, centered at 2128 cm^{-1} and 2136 cm^{-1}, are required to fit the symmetric stretch absorption of (d_3)Met80 at 25 °C (Figure 4A). The Gaussian function centered at 2128 cm^{-1} was assigned as folded (d_3)Met80, and the Gaussian function centered at 2136 cm^{-1} was assigned to the above-mentioned folding intermediate. Consistent with these assignments, the intensity of the absorption at 2136 cm^{-1} first increases and then decreases with increasing temperature. Above 45 °C, the relative amplitudes of the absorptions at 2128 cm^{-1} and 2136 cm^{-1} recede and are replaced by an absorption centered at 2134 cm^{-1}, whose relative amplitude increases asymptotically with increasing temperature. This absorption is identical to that of the free amino acid under the same conditions and is thus assigned to a solvent exposed state. From this data, we conclude that, during thermal unfolding, cyt *c* populates an intermediate involving Met80 dissociation from the heme center, similar to the intermediate populated during chemical denaturation.

The spectra of (d_7)Leu68, (d_7)Leu94, and (d_3)Leu98 were next characterized (Figure 4B–D). For (d_3)Leu98, the spectrum at 25 °C was well fit with two Gaussian functions with center frequencies at 2202 cm^{-1} and 2216 cm^{-1}, respectively. These have been assigned previously as overlapping asymmetric absorptions of the diastereotopic $C_{\delta1}$ and $C_{\delta2}$ methyl groups.[7] At high temperature, the spectrum was well fit with two Gaussian functions with center frequencies of 2211 cm^{-1} and 2222 cm^{-1}, similar to the GdnHCl-denatured state. For (d_7)Leu68 and (d_7)Leu94, the absorptions at ∼2220 cm^{-1} were assigned as the asymmetric absorptions of the diastereotopic $C_{\delta1}$ and $C_{\delta2}$ methyl groups based on the previous characterization of the corresponding (d_3)Leu labeled proteins.[7,33] Fitting the asymmetric stretching absorption of (d_7)Leu68 at 25 °C required two Gaussian functions, both centered at 2213 cm^{-1}. The high-temperature spectrum also required two Gaussian functions, with center frequencies of 2215 cm^{-1} and 2225 cm^{-1}. Fitting the asymmetric stretch vibration of (d_7)Leu94 at 25 °C again required two Gaussian functions, centered at 2225 cm^{-1} and 2208 cm^{-1}; however, fitting the high temperature spectrum required only a single Gaussian centered at 2220 cm^{-1}. While it is not surprising that the low-temperature spectra of (d_7)Leu68 and (d_7)Leu94 differ significantly due to distinct microenvironments in the folded protein, the differences observed for the high temperature state are more surprising and suggest that cyt *c* retains residual structure. In addition, the signal intensity of the d_7 variants is significantly increased relative to the d_3 variant, due to the coincidence of the diastereomeric CD_3 absorptions,[33] resulting in less error and a more accurate characterization of the unfolding transitions (see below).

Finally, fitting the asymmetric CD_3 stretching vibration of (d_3)Ala101 at 25 °C required only a single Gaussian with a center frequency of 2226 cm^{-1} (Figure 4E). As at the other sites examined, the intensity of this absorption decreased with increasing temperature, but in this case, it was not replaced with other absorptions. This suggests that, at higher temperatures, the absorption of this residue is heterogeneously broadened beyond our limit of detection of ∼0.05 mOD. Thus, unlike residues comprising more central parts of the protein, the C-terminal end of the C-terminal helix appears highly disordered in the state populated at high temperature.

To obtain the temperature-dependent fractional concentrations of the folded and high temperature states, spectra at intermediate temperatures were fit to a superposition of the low- and high-temperature spectra (Figure 5). As expected, for (d_3)Met80, fitting required that we allow the amplitude of the Gaussian function assigned to the intermediate to vary independently. For each of the other four isotopically labeled proteins, superposition of the corresponding low- and high-temperature spectra yielded fits of satisfying quality at all temperatures. This indicates that the corresponding residues transition between microenvironments without populating any intermediates.

dx.doi.org/10.1021/jp303521t | *J. Phys. Chem. B* 2012, 116, 6397–6403

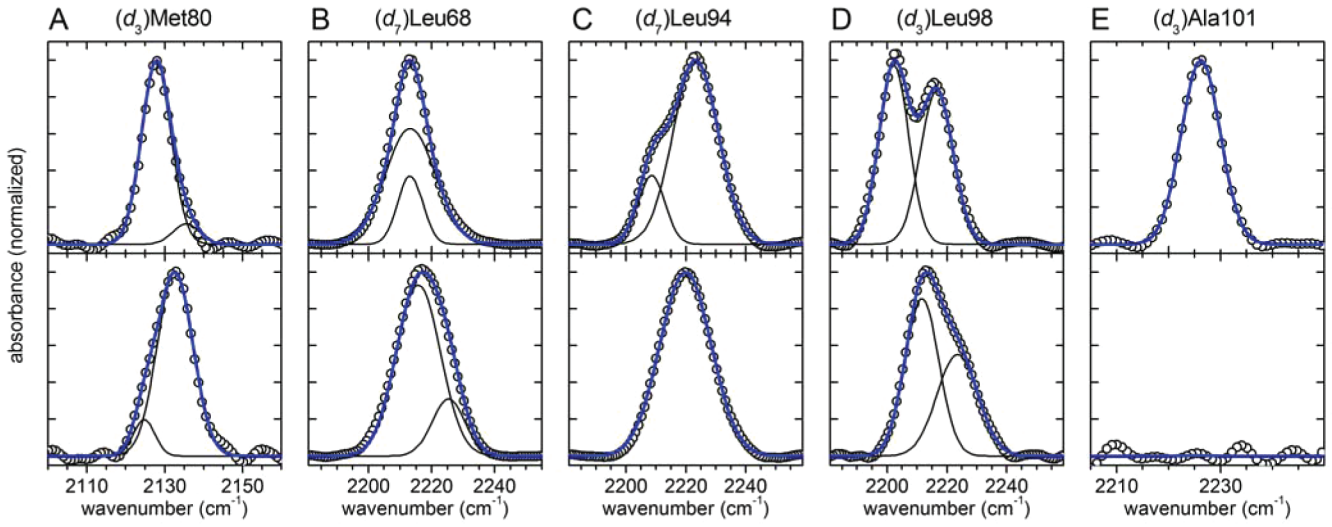


Figure 4. (A–E) FT IR spectra at 25 °C (upper panel) and high temperature (lower panel) for the different deuterium-labeled proteins (open circles, experimental data; black lines, Gaussian functions used to deconvolute the spectra; blue lines, sum of all Gaussians).

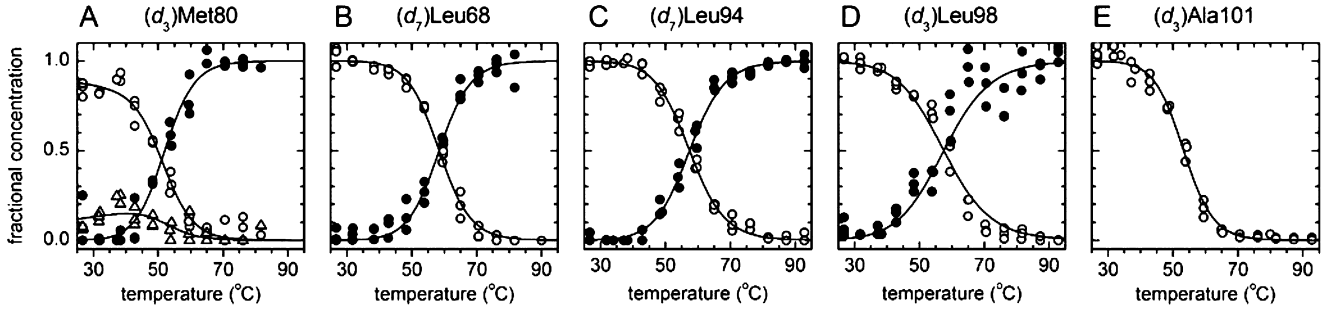


Figure 5. (A–E) Temperature-dependent fractional concentrations for the folded (open circles), intermediate (open triangles), and high-temperature (closed circles) states. Lines are fits to two- or three-state models as described in detail in the text.

Using the van't Hoff equation, we determined the residue-specific midpoint temperature, T_m, as well as the corresponding $\Delta H°$ and $\Delta S°$ from sigmoidal fits to the temperature-dependent fractional amplitudes evoking a three-state equilibrium (folded, intermediate, and high temperature states) for (d_3)Met80, and a two-state equilibrium (folded and high temperature states) for the remaining four proteins (Table 1). The residue-specific data reveal a detailed picture of the unfolding transitions (Figure 6). At 25 °C, because of the presence of 1 M GdnHCl, cyt *c* already populates an intermediate (Figure 6, I_1) in which Met80 is dissociated from the heme cofactor. The lability of the Fe–S bond is well-known, and thought to be of functional importance.[34−36] As the temperature is increased, the population of this intermediate increases slightly and reaches a maximum around 40 °C after which Met80 transitions to a solvent-exposed environment with the population of a second intermediate (Figure 6, I_2), which occurs with a T_m of 51 ± 1 °C. At a slightly higher temperature (T_m = 53.0 ± 0.9 °C), the terminus of the C-terminal helix appears to fray, with Ala101 losing its folded structure. Similar fraying of the C-terminal helix of cyt *c* has also been detected by H/D exchange.[37] We assign the frayed structure as a substate of I_2 as it appears to involve more localized structural changes. At ~58 °C, the protein transitions to a third (un)folding intermediate (Figure 6, I_3) where the protein core loses its native structure, as measured at Leu68 in the 60s loop and Leu94 and Leu98 in the C-terminal helix. The T_m of this transition is in good agreement with that determined via circular dichroism, which suggests that the transition involves a

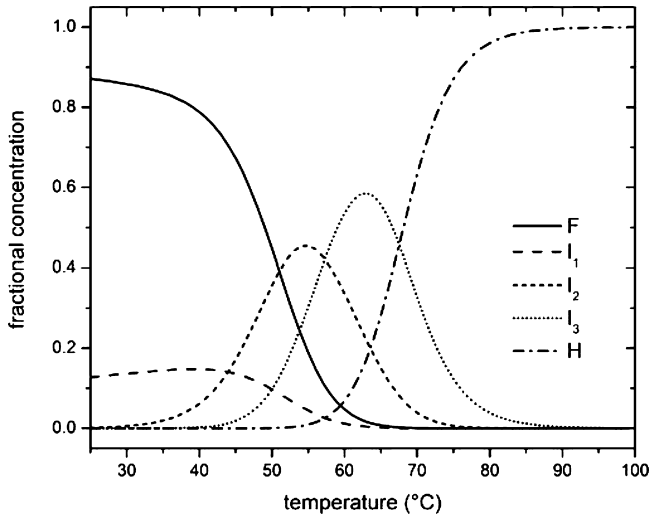


Figure 6. Temperature-dependent fractional concentration for the thermodynamic states identified based on the site-specific folding data. The following parameters were used: $\Delta H°_{F\rightarrow I1}$ = 3.1 kcal/mol, $T_{m,F\rightarrow I1}$ = 200 °C; $\Delta H°_{I1\rightarrow I2}$ = 47 kcal/mol, $T_{m,I1\rightarrow I2}$ = 51 °C; $\Delta H°_{I2\rightarrow I3}$ = 38 kcal/mol, $T_{m,I2\rightarrow I3}$ = 58 °C; and $\Delta H°_{I3\rightarrow H}$ = 62 kcal/mol, $T_{m,I3\rightarrow H}$ = 65 °C.

dx.doi.org/10.1021/jp303521t | J. Phys. Chem. B 2012, 116, 6397−6403

significant portion of the protein. However, we note that Trp59 fluorescence remains quenched, suggesting that I_3 retains a compact structure. Furthermore, the $\Delta H°$ and $\Delta S°$ values suggest that the forces underlying the transition at Leu68 are similar to those underlying the loss of bulk secondary structure, while those for Leu94 and Leu98 appear to be somewhat different. Interestingly, in the absence of GdnHCl, evidence for a cooperative transition between a state resembling I_2 and the unfolded state, without the population of an intermediate, suggests that GdnHCl may stabilize I_3.[18] Finally, at ~65 °C, the protein transitions from a collapsed to an at least partially extended state (Figure 6, H), as marked by recovery of the Trp59 fluorescence.

Interestingly, the high temperature spectra of (d_7)Leu68 and (d_7)Leu94 are distinct, suggesting that one or both remain in a distinct microenvironment and thus that state H retains some level of compactness and/or structure. Such residual structure is suggestive of a molten globule. Molten globules are thought to possess significant secondary structure but an incompletely formed and dynamic tertiary structure and are of great interest due to the role they may play during the folding process.[38] The first molten globule identified was in fact formed by cyt *c* (at low pH and high ionic strength).[39,40] It has also been suggested that thermal unfolding produces a state that is similar to that produced by alkaline denaturation,[19] which has been suggested to be a molten globule.[17] Nonetheless, no significant structural detail has been available for these molten globules, which perhaps represent the greatest challenge to characterization using traditional techniques. The current studies reveal that, after the loss of the native core structure, cyt *c* populates a state with a solvent exposed Met80, a non-natively structured core, and a disordered C-terminus. The persistence of Trp59 fluorescence quenching under these conditions is consistent with this model and further supports the compact and molten globule-like nature of the high-temperature species. Moreover, because Trp59 fluorescence is recovered at ~65 °C but the differences in the (d_7)Leu spectra persist at the highest temperature examined (90 °C), the data further suggest that the molten globule does not lose its residual structure via a single transition.

CONCLUSIONS

The IR characterization of site-specifically incorporated C–D bonds provides an unprecedented opportunity to characterize protein dynamics and (un)folding. The examination of additional sites in cyt *c* should allow for the further characterization of the potential molten globule state as well as other aspects of the equilibrium (un)folding mechanism. Most importantly, with the successful application of the C–D technique under time-resolved conditions,[11] it should be possible to compare the details of the equilibrium unfolding mechanism elucidated here with the kinetic mechanism elucidated using T-jump experiments. This comparison will allow for the direct assessment of the contributions of stepwise and funnel mechanisms to the folding process. Such studies are currently in progress.

ASSOCIATED CONTENT

Supporting Information

Synthetic and experimental details and statistical analysis. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*E-mail: jzimm@scripps.edu (J.Z.); floyd@scripps.edu (F.E.R.).

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank Ms. Crystal Moran-Gutierrez and Professor Ashok Deniz for assistance with the Trp fluorescence experiments. Funding for this work was provided by the National Science Foundation (MCB 0346967).

REFERENCES

(1) Pelton, J. T.; McLean, L. R. *Anal. Biochem.* **2000**, *277*, 167–176.
(2) Chin, J. K.; Jimenez, R.; Romesberg, F. E. *J. Am. Chem. Soc.* **2001**, *123*, 2426–2427.
(3) Chin, J. K.; Jimenez, R.; Romesberg, F. *J. Am. Chem. Soc.* **2002**, *124*, 1846–1847.
(4) Sagle, L. B.; Zimmermann, J.; Dawson, P. E.; Romesberg, F. E. *J. Am. Chem. Soc.* **2006**, *128*, 14232–14233.
(5) Weinkam, P.; Romesberg, F. E.; Wolynes, P. G. *Biochemistry* **2009**, *48*, 2394–2402.
(6) Sagle, L. B.; Zimmermann, J.; Dawson, P. E.; Romesberg, F. E. *J. Am. Chem. Soc.* **2004**, *126*, 3384–3385.
(7) Sagle, L. B.; Zimmermann, J.; Matsuda, S.; Dawson, P. E.; Romesberg, F. E. *J. Am. Chem. Soc.* **2006**, *128*, 7909–7915.
(8) Thielges, M. C.; Groff, D.; Cellitti, S.; Schultz, P. G.; Romesberg, F. E. *Angew. Chem., Int. Ed.* **2009**, *48*, 3478–3481.
(9) Thielges, M. C.; Case, D. A.; Romesberg, F. E. *J. Am. Chem. Soc.* **2008**, *130*, 6597–6603.
(10) Cremeens, M. E.; Zimmermann, J.; Yu, W.; Dawson, P. E.; Romesberg, F. E. *J. Am. Chem. Soc.* **2009**, *131*, 5726–5727.
(11) Zimmermann, J.; Thielges, M. C.; Seo, Y. J.; Dawson, P. E.; Romesberg, F. E. *Angew. Chem., Int. Ed.* **2011**, *50*, 8333–8337.
(12) Filosa, A.; English, A. M. *J. Biol. Inorg. Chem.* **2000**, *5*, 448–454.
(13) Moza, B.; Qureshi, S. H.; Ahmad, F. *Biochem. Biophys. Acta* **2003**, *1646*, 49–51.
(14) Pettigrew, G. W.; Aviram, I.; Schejter, A. *Biochem. J.* **1975**, *149*, 155–167.
(15) Osheroff, N.; Borden, D.; Koppenol, W. H.; Margoliash, E. *J. Biol. Chem.* **1980**, *255*, 1689–1697.
(16) Kaminsky, L. S.; Miller, V. J.; Davison, A. J. *Biochemistry* **1973**, *12*, 2215–2221.
(17) Hagarman, A.; Duitch, L.; Schweitzer-Stenner, R. *Biochemistry* **2008**, *47*, 9667–9677.
(18) Schweitzer-Stenner, R.; Hagarman, A.; Verbaro, D.; Soffer, J. B. *Methods Enzymol.* **2009**, *466*, 109–153.
(19) Taler, G.; Schejter, A.; Navon, G.; Vig, I.; Margoliash, E. *Biochemistry* **1995**, *34*, 14209–14212.
(20) Filosa, A.; Ismail, A. A.; English, A. M. *J. Biol. Inorg. Chem.* **1999**, *4*, 717–726.
(21) Heimburg, T.; Marsh, D. *Biophys. J.* **1993**, *65*, 2408–2417.
(22) Eaton, W. A.; Hochstrasser, R. M. *J. Chem. Phys.* **1967**, *46*, 2533–2539.
(23) Filosa, A.; Wang, Y.; Ismail, A. A.; English, A. M. *Biochemistry* **2001**, *40*, 8256–8263.
(24) Santucci, R.; Giartosio, A.; Ascoli, F. *Arch. Biochem. Biophys.* **1989**, *275*, 496–504.
(25) Wallace, C. J.; Clark-Lewis, I. *J. Biol. Chem.* **1992**, *267*, 3852–3861.
(26) Wallace, C. J.; Offord, R. E. *Biophys. J.* **1979**, *179*, 169–182.
(27) Schnolzer, M.; Alewood, P.; Jones, A.; Alewood, D.; Kent, S. B. *Int. J. Pept. Protein Res.* **1992**, *40*, 180–193.
(28) Weinkam, P.; Zimmermann, J.; Sagle, L. B.; Matsuda, S.; Dawson, P. E.; Wolynes, P. G.; Romesberg, F. E. *Biochemistry* **2008**, *47*, 13470–13480.

(29) Raphael, A. L.; Gray, H. B. *J. Am. Chem. Soc.* **1991**, *113*, 1038−1040.
(30) Raphael, A. L.; Gray, H. B. *Proteins* **1989**, *6*, 338−340.
(31) Colon, W.; Wakem, L. P.; Sherman, F.; Roder, H. *Biochemistry* **1997**, *36*, 12535−12541.
(32) Myer, Y. P. *Biochemistry* **1968**, *7*, 765−776.
(33) Zimmermann, J.; Gundogdu, K.; Cremeens, M. E.; Bandaria, J. N.; Hwang, G. T.; Thielges, M. C.; Cheatum, C. M.; Romesberg, F. E. *J. Phys. Chem. B* **2009**, *113*, 7991−7994.
(34) Telford, J. J.; Wittung-Stafshede, P.; Gray, H. B.; Winkler, J. R. *Acc. Chem. Res.* **1998**, *31*, 755−763.
(35) Weinkam, P.; Zong, C.; Wolynes, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 12401−12406.
(36) Hilgen, S. E.; Pielak, G. J. *Protein Eng.* **1991**, *4*, 575−578.
(37) Bai, Y. W.; Sosnick, T. R.; Mayne, L.; Englander, S. W. *Science* **1995**, *269*, 192−197.
(38) Ptitsyn, O. B. *Adv. Protein Chem.* **1995**, *47*, 83−229.
(39) Goto, Y.; Calciano, L. J.; Fink, A. L. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 573−577.
(40) Jeng, M. F.; Englander, S. W.; Elove, G. A.; Wand, A. J.; Roder, H. *Biochemistry* **1990**, *29*, 10433−10437.