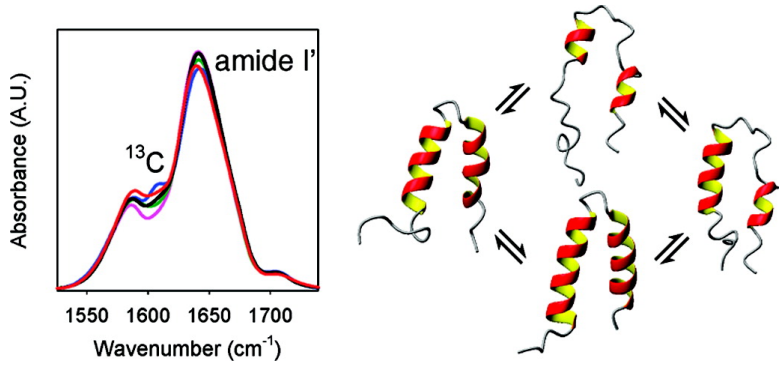
Communication

Site-Specific Unfolding Thermodynamics of a Helix-Turn-Helix Protein

Krista E. Amunson, Loren Ackels, and Jan Kubelka

J. Am. Chem. Soc., **2008**, 130 (26), 8146-8147 • DOI: 10.1021/ja802185e • Publication Date (Web): 05 June 2008



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

J|A|C|S COMMUNICATIONS

Published on Web 06/05/2008

Site-Specific Unfolding Thermodynamics of a Helix-Turn-Helix Protein

Krista E. Amunson, Loren Ackels, and Jan Kubelka*

Department of Chemistry, University of Wyoming, 1000 East University Avenue, Laramie, Wyoming 82071

Received March 24, 2008; E-mail: jkubelka@uwyo.edu

Understanding protein folding has been one of the principal challenges for modern biophysics. The obstacles for elucidating folding mechanisms arise from the complex protein structural transitions, which are difficult to characterize using standard experimental methods. Isotopically edited infrared (IR) spectroscopy has emerged as a promising experimental technique that combines site-specific resolution with an inherently fast time-scale.[1] Substituting a ^{13}C on the C═O of selected amino acids shifts the amide I (predominantly a C═O stretch) ~40 cm^{-1} lower in wavenumber, where it can be resolved from the rest of the (^{12}C) amide modes. However, thus far the isotopically edited IR has been applied predominantly to studies of short peptides.[2−4] In the first application to a protein, Brewer et al.[5] probed the folding of the 36-residue villin headpiece subdomain at one site using a single $^{13}C═^{18}O$ labeled amino acid. A systematic study taking a full advantage of isotopically edited IR to monitor folding in multiple specific sites, has not yet been reported.

In this Communication we present the first application of ^{13}C isotopically edited IR to probe the folding of a protein in multiple, independent sites. We investigate a 40-residue helix-turn-helix (hth) subdomain from the P22 viral scaffolding protein.[6] The hth are the simplest α-helical motifs, which combine both secondary and tertiary interactions and are important as building blocks of larger protein domains, often autonomously stable.[7] Unlike an equivalent motif in β-sheet proteins, the β-hairpin, whose folding has been extensively studied,[4,8] the folding of hth motifs has received much less attention.[7,9] We synthesized the protein, along with four ^{13}C isotopically labeled variants (Figure 1a), using standard FMOC solid phase methodology and investigated the thermal denaturation using circular dichroism (CD) and Fourier-transform IR (FTIR) spectroscopies (Supporting Information).

The temperature dependent CD spectra of a P22 subdomain (0−85 °C) are characteristic of the unfolding of a highly α-helical protein (Figure 1b). The sequences corresponding to each of the individual α-helices were also synthesized and CD was used to test for any residual helical structure. The helix 2 sequence is completely unstructured (Figure 1b inset), for the helix 1 fragment the CD suggests some residual structure; however, the CD signal is smaller than that of the protein at 85 °C and also virtually temperature independent (Supporting Information). The helical motif is therefore stabilized predominantly by the tertiary context.

The IR amide I′ spectra (in D_2O) at 0 °C of the unlabeled P22 subdomain (U) and its four ^{13}C labeled variants: on the helix 1 N-terminus (h1N) and C-terminus (h1C), on the turn (t) and helix 2 (h2), are shown in Figure 1c. The ^{13}C isotopically shifted amide I′ component is clearly seen near 1600 cm^{-1}. The temperature-dependent spectral data were analyzed using singular value decomposition (SVD) and global fitting. The SVD of each of the six (CD, unlabeled, and four labeled FTIR) spectral sets yielded at least three significant components (Supporting Information): a minimum of three independent states is therefore required to account for the data. A three-state thermodynamic model, with fully folded

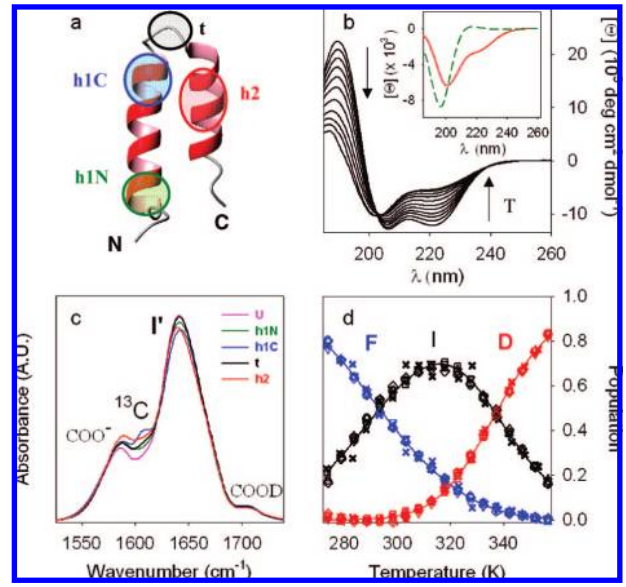


Figure 1. (a) The hth subdomain of P22 viral protein (sequence NH_2−ITGDVSAANK^{10} DAIRKQMDAA^{20} ASKGDVETYR^{30} KLKAK-LKGIR -COOH) with indicated ^{13}C isotopically labeled positions: on helix 1 N-terminus (denoted h1N: A7, A8), helix 1 C-terminus (h1C: A19, A20, A21), turn (**t**: G24, V26), and helix 2 (h2: L32, A34, L36). (b) Temperature dependent CD spectra. Inset: CD of helix 1 (red solid line) and helix 2 (green dashed line) fragments. (c) Amide I′ IR spectra h1N (green), h1C (blue), t (black), and h2 (red). Amide I′ of the unlabeled protein (U, pink) is shown for comparison. (d) Analysis of the experimental data with a three state model: F ⇌ I ⇌ D described by two equilibrium constants: $K_1 = [I]/[F] = \exp[-(\Delta H_1/R)(1/T - 1/T_{m1})]$, $K_2 = [D]/[I] = \exp[-(\Delta H_2/R)(1/T - 1/T_{m2})]$. Solid lines are fitted populations (F, blue line; I, black line; D, red line), symbols correspond to the experimental data: CD (crosses) and amide I′ IR for U (circles), h1N (squares), h1C (triangles up), t (triangles down), and h2 (diamonds). The thermodynamic parameters of the fit are $T_{m1} = (293 \pm 1)$ K, $\Delta H_1 = (10 \pm 1)$ kcal·mol^{-1}, $T_{m2} = (337 \pm 1)$ K, $\Delta H_2 = (20 \pm 1)$ kcal·mol^{-1}.

(F), denatured (D) and intermediate (I) states, fits all the experimental data (Figure 1d) as expected, since all the data sets report on the overall unfolding on the protein. We note that our analysis assumes temperature independent IR spectra for the folding substates. The inherent temperature dependence of the IR is not well understood, with the possible exception of unstructured peptides,[10] and cannot be accounted for without additional assumptions and empirical parameters. On the other hand, a commonly used first approximation of temperature-independent IR is justified by the same thermodynamics obtained from the SVD of the FTIR and the CD data (Figure 1d), the latter being much less temperature sensitive (e.g., CD of the unstructured fragments: Figure S2 in Supporting Information).

To obtain site-specific unfolding, we analyzed the difference spectra (labeled minus unlabeled) of the SVD eigenvectors (basis spectra) for the three substates, shown in Figure 2. Qualitatively distinct contributions to the individual thermodynamic states are

10.1021/ja802185e CCC: $40.75

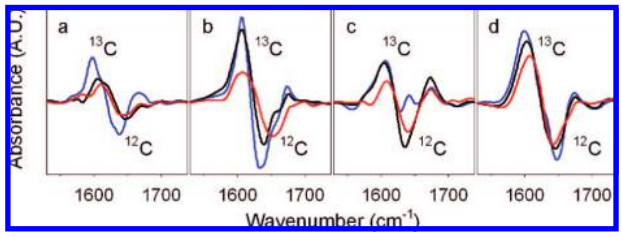


Figure 2. Difference amide I′ spectra for the three thermodynamic states: F (blue), I (black), and D (red) between the ^{13}C labeled and the unlabeled P22 subdomain: (a) h1N, (b) h1C, (c) t, and (d) h2.

evident: for example, in h1N (Figure 2a) the difference spectrum changes significantly between the F and I states, but little from I to D. It is also evident that the ^{13}C and ^{12}C signals have different contributions to the spectral changes. We focus on the ^{13}C signal since it reflects the conformation of local labeled segments, while variations in the ^{12}C part due to altered vibrational coupling by the ^{13}C labels may be more complex.[3,11]

To separate the contribution of the ^{13}C isotopically labeled groups, the amide I′ was decomposed into the ^{13}C and ^{12}C bands using a target transformation procedure.[12] For each labeled protein, the "target" ^{13}C bands (Figure 2, positive differences) were best approximated by the linear combination of basis functions obtained by a SVD of the labeled together with the unlabeled IR spectra (Supporting Information). This procedure allows for the analysis of the partially resolved ^{13}C bands without empirical band fitting. The corresponding transformation of the temperature-dependent SVD weights yields the thermodynamics of unfolding of the ^{13}C-labeled protein segment (Figure 3). For all isotopically labeled P22 variants, the local unfolding of the ^{13}C labeled regions can be described by a two-state thermodynamic model (Supporting Information), as suggested by similar ^{13}C band contributions to two of the three global states (Figure 2.)

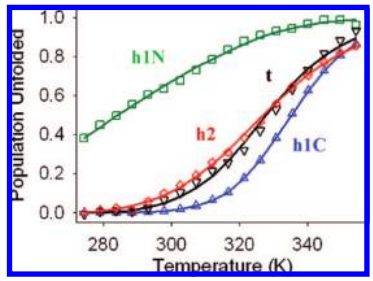


Figure 3. Site-specific thermal unfolding of the P22 subdomain. Fits of the two-state model F ⇌ D; $K = [D]/[F] = \exp[-(\Delta H/R)(1/T - 1/T_m)]$ (solid lines) to the ^{13}C amide I′ component for h1N (green: $T_m = (285 \pm 2)$ K, $\Delta H = (7 \pm 1)$ kcal·mol^{-1}), h1C (blue: $T_m = (334 \pm 1)$ K, $\Delta H = (22 \pm 2)$ kcal·mol^{-1}), t (black: $T_m = (328 \pm 1)$ K, $\Delta H = (19 \pm 2)$ kcal·mol^{-1}), and h2 (red: $T_m = (327 \pm 1)$ K, $\Delta H = (15 \pm 1)$ kcal·mol^{-1}). Symbols (same as in Figure 1) correspond to the experimental data points.

As the results in Figure 3 show, the isotopically labeled regions exhibit characteristic unfolding thermodynamics, which also differ from the overall unfolding (Figure 1). Similar behavior was detected for a "globally downhill" folding small helical protein using NMR.[13] The P22 subdomain unfolds from the N-terminus of helix 1, implying that the global, partially helical intermediate (Figure 1d) corresponds to the unfolded N-terminus. This is consistent with the 3-D structure of the P22 subdomain (Figure 1a),[6] in which the terminus of the longer helix 1 is not in tertiary contact with the second α-helix. The labeled part of helix 2 starts unfolding next through a gradual transition, while a sharper unfolding of the turn segment precedes the C-terminus of helix 1. Both α-helices retain some helical population, which is consistent with the incomplete global unfolding (Figure 1d). A negligible amount of residual structure in the isolated α-helical fragments (Figure 1) implies that the hydrophobic contacts near the turn stabilize the neighboring α-helical segments.

The application of ^{13}C isotopically labeled IR to proteins is not without difficulties. A relatively small fraction of ^{13}C labels and overlap with the carboxylic side-chain vibrations can obscure the isotopic amide I′ signals. However, our results demonstrate that it is possible to detect 2−3 ^{13}C amino acids in a 40-residue protein sequence. The local structural transitions can be resolved by a statistical analysis of the data sets and provide insights into the details of the folding mechanism that cannot be obtained from the overall "average" signals. A description in terms of a few thermodynamic states, which accounts for the overall unfolding (Figure 1d), is not sufficient for characterizing the site-specific unfolding data. Instead, more detailed, residue-level modeling will be necessary.[14] A better understanding of the temperature dependence of the amide I′ IR spectra will also be needed to properly account for the nonstructural temperature effects.[10]

Acknowledgment. This work was supported in part by Faculty Grant-in-Aid and Basic Research Grant programs of the University of Wyoming.

Supporting Information Available: Details of experimental and data analysis procedures. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Decatur, S. M. *Acc. Chem. Res.* **2006**, *39*, 169–175. Dyer, R. B. *Curr. Opin. Struct. Biol.* **2007**, *17*, 38–47.
(2) Tadesse, L.; Nazarbaghi, R.; Walters, L. *J. Am. Chem. Soc.* **1991**, *113*, 7036–7037. Decatur, S. M.; Antonic, J. *J. Am. Chem. Soc.* **1999**, *121*, 11914–11915. Brauner, J. W.; Dugan, C.; Mendelsohn, R. *J. Am. Chem. Soc.* **2000**, *122*, 677–683. Huang, C. Y.; Getahun, Z.; Wang, T.; DeGrado, W. F.; Gai, F. *J. Am. Chem. Soc.* **2001**, *123*, 12111–12112. Huang, R.; Kubelka, J.; Barber-Armstrong, W.; Silva, R.; Decatur, S. M.; Keiderling, T. A. *J. Am. Chem. Soc.* **2004**, *126*, 2346–2354.
(3) Silva, R. A.; G, D.; Kubelka, J.; Decatur, S. M.; Bour, P.; Keiderling, T. A. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 8318–8323.
(4) Setnička, V.; Huang, R.; Thomas, C. L.; Etienne, M. A.; Kubelka, J.; Hammer, R. P.; Keiderling, T. A. *J. Am. Chem. Soc.* **2005**, *127*, 4992–4993.
(5) Brewer, S. H.; Song, B. B.; Raleigh, D. P.; Dyer, R. B. *Biochemistry* **2007**, *46*, 3279–3285.
(6) Sun, Y.; Parker, M. H.; Weigele, P.; Casjens, S.; Prevelige, P. E.; Krishna, N. R. *J. Mol. Biol.* **2000**, *297*, 1195–1202.
(7) Religa, T. L.; Johnson, C. M.; Vu, D. M.; Brewer, S. H.; Dyer, R. B.; Fersht, A. R. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 9272–9277.
(8) Muñoz, V.; Thompson, P. A.; Hofrichter, J. *Nature* **1997**, *390*, 196–199. Dinner, A. R.; Lazaridis, T.; Karplus, M. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 9068–9073. Klimov, D. K.; Thirumalai, D. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 2544–2549. Zagrovic, B.; Sorin, E. J.; Pande, V. *J. Mol. Biol.* **2001**, *313*, 151–169. Hilario, J.; Kubelka, J.; Keiderling, T. A. *J. Am. Chem. Soc.* **2003**, *125*, 7562–7574. Snow, C. D.; Qiu, L. L.; Du, D. G.; Gai, F.; Hagen, S. J.; Pande, V. S. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 4077–4082.
(9) Hoffmann, D.; Knapp, E. W. *J. Phys. Chem. B* **1997**, *101*, 6734–6740. Fezoui, Y.; Braswell, E. H.; Xian, W.; Osterhout, J. J. *Biochemistry* **1999**, *38*, 2796–2804. Du, D. G.; Gai, F. *Biochemistry* **2006**, *45*, 13131–13139. Xian, W. J.; Connolly, P. J.; Oslin, M.; Hausrath, A. C.; Osterhout, J. J. *Protein Sci.* **2006**, *15*, 2062–2070.
(10) Amunson, K. E.; Kubelka, J. *J. Phys. Chem. B* **2007**, *111*, 9993–9998.
(11) Kubelka, J.; Keiderling, T. A. *J. Am. Chem. Soc.* **2001**, *123*, 6142–6150. Turner, D. R.; Kubelka, J. *J. Phys. Chem. B* **2007**, *111*, 1834–1845.
(12) Pelikán, P.; Čeppan, M.; Liška, M. *Applications of Numerical Methods in Molecular Spectroscopy*; CRC Press: Boca Raton, FL, 1994.
(13) Sadqi, M.; Fushman, D.; Muñoz, V. *Nature* **2006**, *442*, 317–321.
(14) Muñoz, V.; Eaton, W. A. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 11311–11316. Galzitskaya, O. V.; Finkelstein, A. V. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 11299–11304. Alm, E.; Baker, D. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 11305–11310.
(15) Henry, E. R.; Eaton, W. A. *Chem. Phys.* **2004**, *307*, 163–185.

JA802185E