

Achieving Secondary Structural Resolution in Kinetic Measurements of Protein Folding: A Case Study of the Folding Mechanism of Trp-cage**

Robert M. Culik, Arnaldo L. Serrano, Michelle R. Bunagan,* and Feng Gai*

Protein folding kinetics are often measured by monitoring the change of a single spectroscopic signal, such as the fluorescence of an intrinsic fluorophore or the absorbance at a single frequency within an electronic or vibrational band of the protein backbone. While such an experimental strategy is easy to implement, the use of a single spectroscopic signal can leave important folding events undetected and overlooked. Herein, we demonstrate, using the miniprotein Trp-cage as an example, that the structural resolution of protein folding kinetics can be significantly improved when a multi-probe and multi-frequency approach is used, thus allowing a more complete understanding of the folding mechanism.

Trp-cage is a 20-residue miniprotein designed by Andersen and co-workers.^[1] Among the many Trp-cage variants (the name and sequence of the Trp-cage peptides studied here are listed in Table S1 in the Supporting Information), TC5b is the most studied, both experimentally and computationally. As shown (Figure 1), the folded structure of Trp-cage consists of three secondary structural elements: an α -helix from residues 2–8, a 3_{10} -helix consisting of residues 12–14, and a polyproline region spanning residues 17–19, which together generate a hydrophobic cage housing the sole tryptophan residue of the peptide. Because of its small size and fast folding rate, Trp-cage has been an extremely popular model for computational studies of protein folding dynamics.^[2–42] However, experimental investigations of the folding kinetics and mechanism of Trp-cage remain scarce. Using a temperature-jump (*T*-jump) fluorescence technique, Hagen and co-workers^[43] showed that TC5b folds in about 4 μ s at room temperature, while an infrared (IR) *T*-jump study by Bunagan et al. indicated that the P12W mutant of TC5b, or Trp²-cage, folds

even faster.^[44] In both cases, single-exponential relaxation kinetics were observed, suggesting that folding proceeds in a two-state manner. On the other hand, equilibrium unfolding studies provided evidence suggesting the existence of folding intermediates corresponding to a compact denatured state^[45,46] and a partially folded state with maximal thermal stability of 20 °C.^[47] Moreover, a large number of different folding pathways have been observed in computer simulations, including, for instance, the formation of an early intermediate in which the hydrophobic core is bisected by the D9–R16 salt bridge,^[48] and the concurrent formation of the α -helix and the hydrophobic core,^[19,27,28] among others.

Generating a conclusive experimental verification of these previous simulation results experimentalists face a great challenge, because the kinetic techniques commonly used in protein folding studies offer relatively low structural resolution. To overcome this limitation and to provide new insights into the folding mechanism of Trp-cage, we seek to use a multi-probe approach to dissect the folding kinetics of individual local structural elements of the native fold. To this end, we measure *T*-jump-induced conformational relaxation kinetics^[49] at well-chosen frequencies in the amide I' region of the protein which report the absorbance changes of the α -helix, the 3_{10} -helix, the unfolded structural ensemble, as well as the asparagine (Asp) side chain. Separation of the α -helix IR signal from those arising from other structural motifs is facilitated by using the following Trp-cage sequence: DA*Y-A*QWLKDGPPSSGRPPPS (hereafter referred to as ¹³C-TC10b), where A* represents ¹³C=O-labeled alanine, amide I' frequency of which is known to be red-shifted by about 40 cm⁻¹ from that of the unlabeled helical amides.^[50–52] Andersen and co-workers have shown that this sequence,

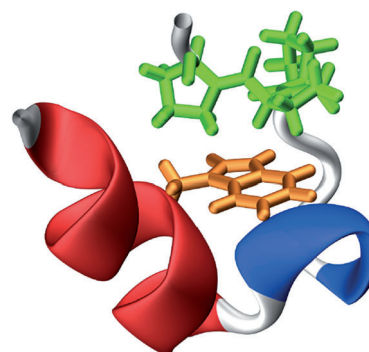


Figure 1. Structure of the Trp-cage (taken from protein data bank (PDB) file 1L2Y), showing the α -helix (red), the 3_{10} -helix (blue), the polyproline region (green), and the sole tryptophan (orange).

[*] A. L. Serrano,^[†] Prof. Dr. F. Gai
 Department of Chemistry, University of Pennsylvania
 231 S. 34 Street, Philadelphia, PA 19104 (USA)
 E-mail: gai@sas.upenn.edu

Prof. Dr. M. R. Bunagan
 Department of Chemistry, College of New Jersey
 2000 Pennington Road, Ewing, NJ 08628 (USA)
 E-mail: bunagan@tcnj.edu

R. M. Culik^[†]
 Department of Biochemistry and Molecular Biophysics
 University of Pennsylvania (USA)

[†] These authors contributed equally.

[**] We thank the National Institutes of Health (GM-065978, RR01348, and GM-008275) for funding. R.M.C. acknowledges a training grant in structural biology.

Supporting information for this article is available on the WWW under <http://dx.doi.org/10.1002/anie.201104085>.

which is referred to as TC10b in their study, yields a more stable Trp-cage fold and is therefore a better model for both experimental and computational studies.^[53] In addition, we employ several well-chosen mutations and ϕ -value analysis^[54] to determine the structural elements formed in the folding transition state.

As shown (see Figures S1 and S2 and Table S2 in the Supporting Information), the thermal unfolding properties of the Trp-cage variants studied here, determined by circular dichroism (CD) spectroscopy, are in quantitative agreement with those reported in the literature.^[44,53] For example, the thermal melting temperature (i.e., T_m) of ^{13}C -TC10b is determined to be $(55.0 \pm 1.0)^\circ\text{C}$ and matches well with a T_m of 56°C reported by Andersen and co-workers for TC10b.^[53]

In comparison with the FTIR difference spectrum of TC5b (Figure S3 in the Supporting Information), the FTIR difference spectrum of ^{13}C -TC10b (Figure 2) indicates that the negative spectral feature at around 1615 cm^{-1} is due to the ^{13}C -labeled alanine (Ala) residues, thus uniquely reporting the thermal melting of the α -helical segment within the protein. The negative peak at around 1646 cm^{-1} arises from the loss of unlabeled helical amides. The apparent blue shift and lower intensity of the unlabeled helical amide I band in the difference spectrum, relative to that observed for unlabeled Trp-cage, is due to spectral overlapping with the amide I band of the $^{13}\text{C}=\text{O}$ units in the thermally denatured state.^[52] On the other hand, the positive spectral feature arises from $^{12}\text{C}=\text{O}$ units in the thermally unfolded state of ^{13}C -TC10b. In addition, the negative feature at around 1586 cm^{-1} is due to the absorbance change of the deprotonated Asp side chain, that is, $\nu_{\text{as}}(\text{COO}^-)$,^[55] in response to protein unfolding. Since the salt bridge formed between the side chains of residues D9 and R16 is a key structural determinant of the Trp-cage stability and fold,^[56] we believe that this spectral feature provides an excellent IR marker for probing the global folding/unfolding kinetics of the cage structure.^[57]

As shown (Figure 3), the T -jump-induced conformational relaxation kinetics probed at both 1580 and 1612 cm^{-1} can be adequately described by a single-exponential function and the corresponding rate constants, as indicated (Figure 4), are indistinguishable from each other within the limit of experimental errors. Interestingly, however, when probed at

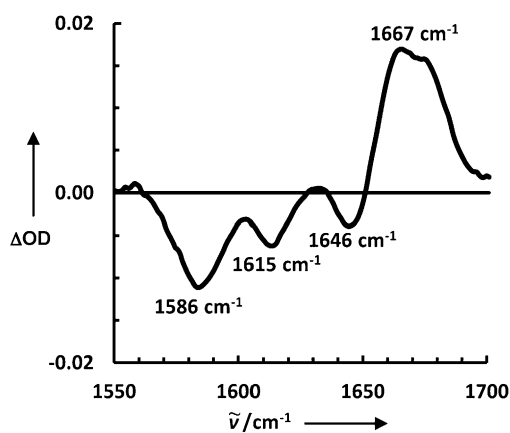


Figure 2. A representative FTIR difference spectrum of ^{13}C -TC10b between 65.0 and 25.0°C (OD = optical density).

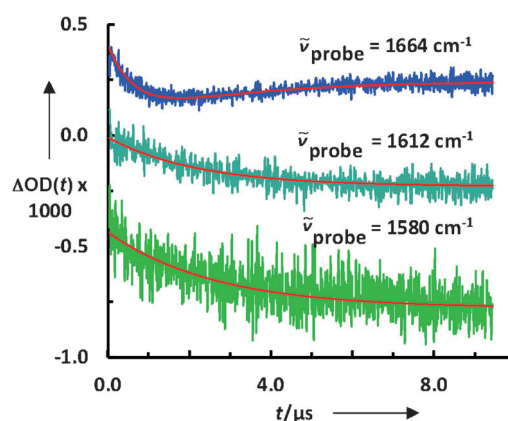


Figure 3. Representative T -jump-induced conformational relaxation traces of ^{13}C -TC10b in response to a T jump from 5 to 10°C , probed at different frequencies as indicated. The smooth lines are the corresponding fits of these data to either a single-exponential (for 1580 and 1612 cm^{-1}) or a double-exponential function (for 1664 cm^{-1}) and the resulting rate constants (k) are given in Figure 4. For easy comparison, these data have been offset.

1664 cm^{-1} , a frequency where both the 3_{10} -helix and disordered conformation are known to absorb,^[58] the T -jump-induced conformational relaxation kinetics can only be fitted by two exponential functions with amplitudes of opposite sign (Figure 3). As indicated (Figure 4), the rate constant of the positive (and slower) kinetic phase is also identical, within experimental uncertainty, to those rates measured at 1580 and 1612 cm^{-1} . Therefore, we attribute this kinetic phase to the global folding–unfolding transition of the Trp-cage structure. Consequently, we assign the fast phase, the amplitude of which decreases with time, to the local unfolding of the 3_{10} -helix.

The assignment of the fast kinetic phase observed at 1664 cm^{-1} to T -jump-induced conformational relaxation of the 3_{10} -helix is consistent with several lines of evidence. First, it has been shown that 3_{10} -helices absorb in the region of 1660 cm^{-1} .^[58,59] Second, the full amplitude of this phase decreases with increasing final temperature (for the same T -

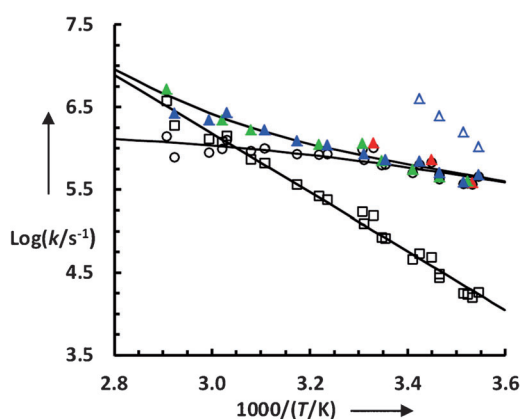


Figure 4. Conformational relaxation rate constants (k , filled symbols) of ^{13}C -TC10b obtained with a probing frequency of 1580 (red), 1612 (green), and 1664 cm^{-1} (blue), respectively. The blue empty triangles represent the relaxation rates of the fast kinetic phase observed at 1664 cm^{-1} . The black empty symbols represent the global folding (circle) and unfolding (square) rates of the protein.

Keywords: IR spectroscopy · kinetics · protein folding · proteins

- [1] J. W. Neidigh, R. M. Fesinmeyer, N. H. Andersen, *Nat. Struct. Biol.* **2002**, *9*, 425–430.
- [2] C. D. Snow, B. Zagrovic, V. S. Pande, *J. Am. Chem. Soc.* **2002**, *124*, 14548–14549.
- [3] C. Simmerling, B. Strockbine, A. E. Roitberg, *J. Am. Chem. Soc.* **2002**, *124*, 11258–11259.
- [4] S. Chowdhury, M. C. Lee, G. M. Xiong, Y. Duan, *J. Mol. Biol.* **2003**, *327*, 711–717.
- [5] J. W. Pitera, W. Swope, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 7587–7592.
- [6] G. V. Nikiforovich, N. H. Andersen, R. M. Fesinmeyer, C. Frieden, *Proteins Struct. Funct. Genet.* **2003**, *52*, 292–302.
- [7] A. Schug, T. Herges, W. Wenzel, *Phys. Rev. Lett.* **2003**, *91*, 158102.
- [8] P. Carnevali, G. Toth, G. Toubassi, S. N. Meshkat, *J. Am. Chem. Soc.* **2003**, *125*, 14244–14245.
- [9] S. Chowdhury, M. C. Lee, Y. Duan, *J. Phys. Chem. B* **2004**, *108*, 13855–13865.
- [10] P. J. Steinbach, *Proteins Struct. Funct. Genet.* **2004**, *57*, 665–677.
- [11] M. Ota, M. Ikeguchi, A. Kidera, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 17658–17663.
- [12] A. Linhananta, J. Boer, I. MacKay, *J. Chem. Phys.* **2005**, *122*, 114901–114915.
- [13] A. S. N. Seshasayee, *Theor. Biol. Med. Modell.* **2005**, *2*, 7–11.
- [14] F. Ding, S. V. Buldyrev, N. V. Dokholyan, *Biophys. J.* **2005**, *88*, 147–155.
- [15] A. Irbäck, S. Mohanty, *Biophys. J.* **2005**, *88*, 1560–1569.
- [16] J. L. Alonso, P. Echenique, *Biophys. Chem.* **2005**, *115*, 159–168.
- [17] J. Chen, W. Im, C. L. Brooks, *J. Am. Chem. Soc.* **2006**, *128*, 3728–3736.
- [18] L. X. Zhan, J. Z. Y. Chen, W. K. Liu, *Proteins Struct. Funct. Genet.* **2007**, *66*, 436–443.
- [19] D. Paschek, H. Nymeyer, A. E. Garcia, *J. Struct. Biol.* **2007**, *157*, 524–533.
- [20] L. Yang, M. P. Grubb, Y. Q. Gao, *J. Chem. Phys.* **2007**, *126*, 125102.
- [21] D. A. C. Beck, G. W. N. White, V. Daggett, *J. Struct. Biol.* **2007**, *157*, 514–523.
- [22] S. Piana, A. Laio, *J. Phys. Chem. B* **2007**, *111*, 4553–4559.
- [23] J. Copps, R. F. Murphy, S. Lovas, *Biopolymers* **2007**, *88*, 427–437.
- [24] A. Kentsis, T. Gindin, M. Mezei, R. Osman, *PLoS ONE* **2007**, *2*, e446.
- [25] Z. H. Hu, Y. H. Tang, H. F. Wang, X. Zhang, M. Lei, *Arch. Biochem. Biophys.* **2008**, *475*, 140–147.
- [26] P. Hudáky, P. Straner, V. Farkas, G. Varadi, G. Toth, A. Perczel, *Biochemistry* **2008**, *47*, 1007–1016.
- [27] W. X. Xu, Y. G. Mu, *Biophys. Chem.* **2008**, *137*, 116–125.
- [28] D. Paschek, S. Hempel, A. E. Garcia, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 17754–17759.
- [29] S. Wu, P. I. Zhuravlev, G. A. Papoian, *Biophys. J.* **2008**, *95*, 5524–5532.
- [30] X. Q. Yao, Z. S. She, *Biochem. Biophys. Res. Commun.* **2008**, *373*, 64–68.
- [31] S. Kannan, M. Zacharias, *Proteins Struct. Funct. Genet.* **2009**, *76*, 448–460.
- [32] J. Cerný, J. Vondrasek, P. Hobza, *J. Phys. Chem. B* **2009**, *113*, 5657–5660.
- [33] Y. Chebaro, X. Dong, R. Laghaei, P. Derreumaux, N. Mousseau, *J. Phys. Chem. B* **2009**, *113*, 267–274.
- [34] D. Matthes, B. L. de Groot, *Biophys. J.* **2009**, *97*, 599–608.
- [35] F. Marinelli, F. Pietrucci, A. Laio, S. Piana, *PLoS Comput. Biol.* **2009**, *5*, e1000452.
- [36] Z. Gattin, S. Riniker, P. J. Hore, K. H. Mok, W. F. van Gunsteren, *Protein Sci.* **2009**, *18*, 2090–2099.
- [37] M. Gao, H. Q. Zhu, X. Q. Yao, Z. S. She, *Biochem. Biophys. Res. Commun.* **2010**, *392*, 95–99.
- [38] C. Velez-Vega, E. E. Borrero, F. A. Escobedo, *J. Chem. Phys.* **2010**, *133*, 105103.
- [39] N. J. Bruce, R. A. Bryce, *J. Chem. Theory Comput.* **2010**, *6*, 1925–1930.
- [40] M. S. Lee, M. A. Olson, *J. Chem. Theory Comput.* **2010**, *6*, 2477–2487.
- [41] R. Day, D. Paschek, A. E. Garcia, *Proteins Struct. Funct. Genet.* **2010**, *78*, 1889–1899.
- [42] W. Zheng, E. Gallicchio, N. Deng, M. Andrec, R. M. Levy, *J. Phys. Chem. B* **2011**, *115*, 1512–1523.
- [43] L. L. Qiu, S. A. Pabit, A. E. Roitberg, S. J. Hagen, *J. Am. Chem. Soc.* **2002**, *124*, 12952–12953.
- [44] M. R. Bunagan, X. Yang, J. G. Saven, F. Gai, *J. Phys. Chem. B* **2006**, *110*, 3759–3763.
- [45] H. Neuweiler, S. Doose, M. Sauer, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 16650–16655.
- [46] K. H. Mok, L. T. Kuhn, M. Goetz, I. J. Day, J. C. Lin, N. H. Andersen, P. J. Hore, *Nature* **2007**, *447*, 106–109.
- [47] Z. Ahmed, I. A. Beta, A. V. Mikhonin, S. A. Asher, *J. Am. Chem. Soc.* **2005**, *127*, 10943–10950.
- [48] R. H. Zhou, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 13280–13285.
- [49] C. Y. Huang, Z. Getahun, Y. J. Zhu, J. W. Klemke, W. F. Degrado, F. Gai, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 2788–2793.
- [50] L. Tadesse, R. Nazarboghi, L. Walters, *J. Am. Chem. Soc.* **1991**, *113*, 7036–7037.
- [51] S. M. Decatur, J. Antonic, *J. Am. Chem. Soc.* **1999**, *121*, 11914–11915.
- [52] C. Y. Huang, Z. Getahun, T. Wang, W. F. Degrado, F. Gai, *J. Am. Chem. Soc.* **2001**, *123*, 12111–12112.
- [53] B. Barua, J. C. Lin, V. D. Williams, P. Kummner, J. W. Neidigh, N. H. Andersen, *Protein Eng. Des. Sel.* **2008**, *21*, 171–185.
- [54] A. R. Fersht, A. Matouschek, L. Serrano, *J. Mol. Biol.* **1992**, *224*, 771–782.
- [55] A. Barth, C. Zscherp, *Q. Rev. Biophys.* **2002**, *35*, 369–430.
- [56] D. V. Williams, A. Byrne, J. Stewart, N. H. Andersen, *Biochemistry* **2011**, *50*, 1143–1152.
- [57] S. Bagchi, C. Falvo, S. Mukamel, R. M. Hochstrasser, *J. Phys. Chem. B* **2009**, *113*, 11260–11273.
- [58] D. F. Kennedy, M. Crisma, C. Toniolo, D. Chapman, *Biochemistry* **1991**, *30*, 6541–6548.
- [59] R. A. G. D. Silva, S. C. Yasui, J. Kubelka, F. Formaggio, M. Crisma, C. Toniolo, T. A. Keiderling, *Biopolymers* **2002**, *65*, 229–243.
- [60] S. Williams, T. P. Causgrove, R. Gilmanishin, K. S. Fang, R. H. Callender, W. H. Woodruff, R. B. Dyer, *Biochemistry* **1996**, *35*, 619–697.
- [61] P. A. Thompson, W. A. Eaton, J. Hofrichter, *Biochemistry* **1997**, *36*, 9200–9210.
- [62] C. Y. Huang, J. W. Klemke, Z. Getahun, W. F. DeGrado, F. Gai, *J. Am. Chem. Soc.* **2001**, *123*, 9235–9238.
- [63] R. H. Zhou, *J. Mol. Graphics Modell.* **2004**, *22*, 451–463.
- [64] J. Juraszek, P. G. Bolhuis, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 15859–15864.
- [65] Y. Xu, R. Oyola, F. Gai, *J. Am. Chem. Soc.* **2003**, *125*, 15388–15394.
- [66] A. L. Serrano, M. J. Tucker, F. Gai, *J. Phys. Chem. B* **2011**, *115*, 7472–7478.
- [67] S. Mukherjee, P. Chowdhury, M. R. Bunagan, F. Gai, *J. Phys. Chem. B* **2008**, *112*, 9146–9150.
- [68] H. Maity, M. Maity, M. M. G. Krishna, L. Mayne, S. W. Englander, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 4741–4746.

Communications

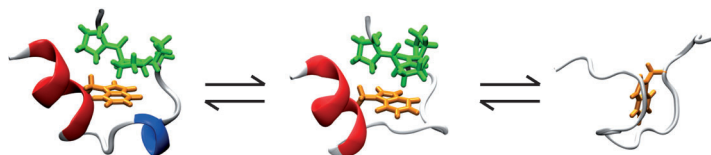
VIP

Protein Folding

R. M. Culik, A. L. Serrano,

M. R. Bunagan,* F. Gai* — ■■■■-■■■■

Achieving Secondary Structural
Resolution in Kinetic Measurements of
Protein Folding: A Case Study of the
Folding Mechanism of Trp-cage



A new twist: A multi-probe and multi-frequency approach is shown for dissecting the folding dynamics of individual protein structural elements. In response to a temperature jump the 3_{10} -helix (blue

in the picture) of the miniprotein Trp-cage unfolds before the global unfolding of the protein, whereas the formation of the cage structure depends on the folding of the α -helix (red).