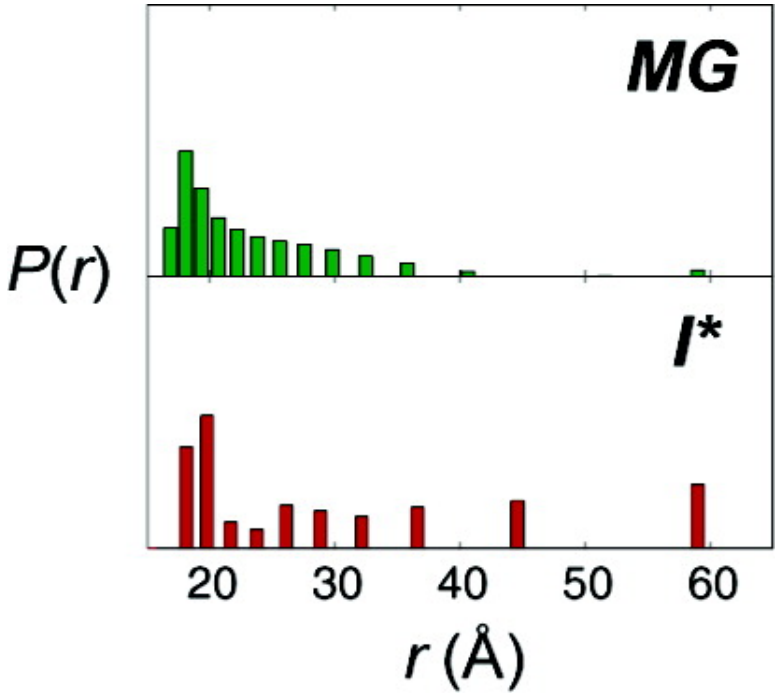
Communication

Nature of the Cytochrome *c* Molten Globule

Ekaterina V. Pletneva, Harry B. Gray, and Jay R. Winkler

J. Am. Chem. Soc., **2005**, 127 (44), 15370-15371• DOI: 10.1021/ja0555318 • Publication Date (Web): 13 October 2005



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 10/13/2005

Nature of the Cytochrome *c* Molten Globule

Ekaterina V. Pletneva, Harry B. Gray,* and Jay R. Winkler*

Beckman Institute, California Institute of Technology, Pasadena, California 91125

Received August 12, 2005; E-mail: hbgray@caltech.edu; winklerj@caltech.edu

Molten globules are compact proteins with fluctuating tertiary conformations that contain substantial amounts of secondary structure.[1] Cytochrome *c* (cyt *c*) undergoes a transition from an acid-denatured form to a molten globule, or A state, at high salt concentrations.[2] The hydrodynamic dimensions of this state are only slightly greater than those of the native protein.[3−6] Studies have indicated that the A state has native-like helix content and possesses some native tertiary contacts.[7] Although its NMR spectrum does not exhibit the chemical shift dispersion indicative of a static conformation,[8] hydrogen exchange NMR experiments have revealed considerable protection of amide protons in *N*-, *C*-, and 60s helices.[9] Native-like tertiary interactions between the *N*- and *C*-helices also have been observed in the cyt *c* A state.[9,10]

We are employing fluorescence energy transfer (FET) kinetics to probe the structural features and heterogeneity of dynamic protein structures.[11,12] Here we focus on the properties of the dynamic A state of five dansyl-labeled (Dns) Cys mutants of *Saccharomyces cerevisiae* iso-1 cyt *c* (Figure 1a). All five variants exhibit structures and stability similar to those of the wild-type protein.[12] The covalently bound Fe(III) heme group quenches Dns fluorescence by energy transfer; the critical length, R_0, of this donor (***D***)−acceptor (***A***) pair is 36 Å for the acid-unfolded state and 39 Å for the native and A states.[12] In the absence of specific information about its structure and heterogeneity, we have chosen model-independent approaches to analyze A-state FET kinetics. The Dns fluorescence decay is described by a discrete Laplace transform; inversion of this series provides a distribution function ($P(k)$) describing the probability that an exponential term with rate constant k_i contributes to the observed Dns fluorescence decay. Inversion of the transform using a constrained linear least-squares method (LSQNONNEG algorithm in MATLAB) produced the narrowest $P(k)$ distributions. In a second approach, the entropy of the $P(k)$ distributions ($S = -\sum_k P(k)\ln[P(k)]$) was used as a regularization parameter, and the $P(k)$ distributions were determined by simultaneous minimization of χ^2 and maximization of the entropy S. The maximum-entropy (ME) regularization method broadens $P(k)$ distributions but does not substantially change the positions of the local maxima (Supporting Information). The two methods of analysis yielded the two limits (narrowest and widest) for the $P(k)$ distributions consistent with our data. The $P(k)$ distributions are readily transformed to distance distributions ($P(r)$).[12]

We have extracted distributions of ***D***−***A*** distances from measurements of FET kinetics with the acid-unfolded protein (pH 2) and under conditions favoring the molten globule (pH 2 and high $[SO_4^{2-}]$)[13] (Figure 1b and Supporting Information). The conversion to the A state was confirmed by both far-UV CD and visible absorption spectroscopy.[11] The acid-unfolded protein is a heterogeneous ensemble of unfolded structures. Comparison of these $P(r)$ distributions to those of the GuHCl-denatured protein[12] suggests that the acid-unfolded state is more compact, consistent with SAXS measurements.[3] As the salt concentration increases, we observe a decrease in the mean and root-mean-squared distances, as well as in the variances of the $P(r)$ distributions. The trend is the same for all variants, with the exception of Dns4cyt. For the latter protein, the $P(r)$ distributions are very similar in the native, molten globule, and unfolded states, suggesting that the Dns label always remains close to the heme.[12] FET kinetics are a sensitive indicator of conformational heterogeneity: at lower salt concentrations, subpopulations of compact and extended structures are clearly evident (Figure 1b). The $P(r)$ distributions confirm that the nature of the polypeptide ensemble strongly depends on the salt concentration.

At high salt concentrations, our experiments with five Dns variants consistently show that the ***D***−***A*** distances in the molten globule are essentially the same as in the native state (Figure 2 and Supporting Information); these results are consistent with our preliminary study of the native protein labeled at position 102.[11] The equilibrium $P(r)$ distributions can be transformed into potential energy functions $U(r)$[14] that underscore the striking resemblance of the two states. The similarity of the two potentials accounts for the extremely fast conversion of the molten globule to the native configuration.[8] Taken together with reports from other laboratories,[6,9,10] the combined data suggest that the cyt *c* molten globule is a highly structured state with a substantial number of native interactions.

That is not to say that there are not subtle structural differences between cyt *c* native and A states. In general, the widths of the $P(r)$ distributions are greater for the molten globule, reflecting a larger number of energetically accessible conformations. In addition to polypeptide flexibility, equilibration among different ligation states of the heme also may contribute to the observed conformational diversity.[15] Amide hydrogen exchange studies indicate that the molten globule contains hydrogen-bonded structures in all three major helices, while the loop regions remain disordered.[9] With five labeling sites extensively sampling cyt *c* conformations (Figure 1a), we also find increased structural heterogeneity in the loops. Even at 1 M Na_2SO_4, the Dns50cyt variant shows a more prominent subpopulation of slightly extended structures than the other variants (Figure 2). The region around this site is apparently susceptible to large-scale conformational fluctuations. Disruption of tertiary hydrogen bonding contacts between Tyr48 and the heme propionate group, observed in resonance Raman experiments,[16] is consistent with this proposal. Our broadened distance distributions directly reflect the greater conformational diversity of the A state. NMR spectra suggest that conformational interconversions in the molten globule occur on the millisecond time scale or faster.[2,8]

Molten globules are commonly believed to represent intermediates in globular protein folding.[1] We have compared the structural features of the equilibrium A state to the "burst-phase" intermediate formed during refolding of five Dns-labeled cyt *c* variants (Figure 2 and Supporting Information). Measurements of FET kinetics produced snapshots of the $P(r)$ distributions of the transient species.[17] As our data show, the polypeptide ensemble of the folding protein is much more heterogeneous than a uniformly collapsed molten globule: there are distinct populations of disparate structures,

10.1021/ja0555318 CCC: $30.25

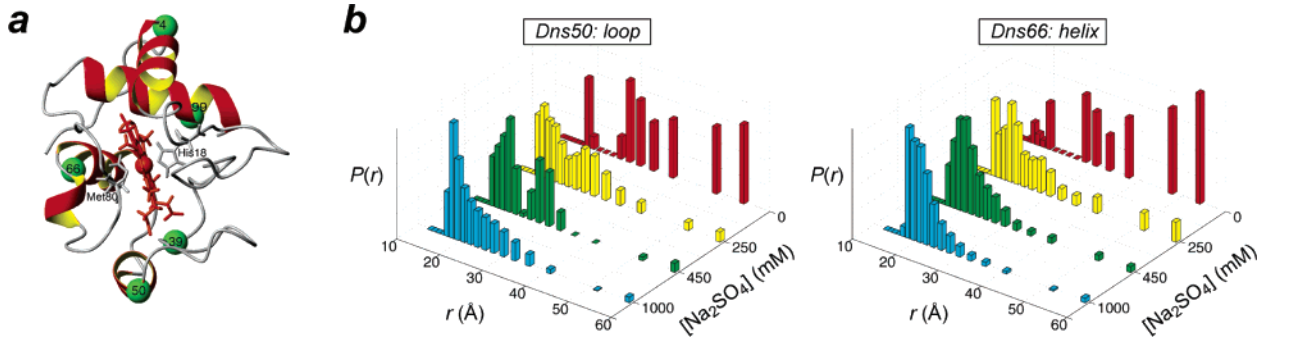


Figure 1. (a) Positions of the labeling sites in *Saccharomyces cerevisiae* iso-1 cyt *c*. (b) Distributions of ***D***−***A*** distances, $P(r)$, in Dns50cyt and Dns66cyt at pH 2.0, 18 °C, and different concentrations of Na_2SO_4 from ME analyses of the FET kinetics. The analyses yielded upper limits for the distribution widths consistent with our data. At distances longer than 1.5 R_0, energy transfer rate constants and ***D***−***A*** distances cannot be determined reliably; the structures with $r \geq 1.5\ R_0$ are represented with a single bar.

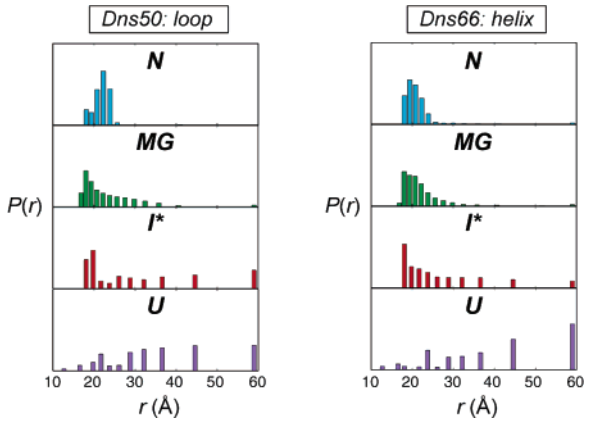


Figure 2. Distributions of ***D***−***A*** distances, $P(r)$ (ME analysis), in Dns50cyt and Dns66cyt for the native state at pH 7.0 (***N***); the molten globule, at pH 2.0 and 1.0 M Na_2SO_4 (***MG***); the transient species formed 1 ms after stopped-flow triggered refolding (0.2 M GuHCl, 0.15 M imidazole, pH 7.0) (***I****); and the unfolded state at 2 M GuHCl, 0.15 M imidazole, pH 7.0 (***U***).

some compact and others more extended. Moreover, local interactions appear to be different as well. Site-specific EPR spin-labeling studies of cyt *c* refolding have revealed early immobilization of a spin-label at residue 50.[18] By contrast, we find that this site is very mobile in the molten globule. Clearly, the equilibrium molten globule is not a suitable model for the early intermediates in cyt *c* refolding. The conformational properties of this partially unfolded state may be more closely related to those of late intermediates[5,19] or even the native state. Rather than providing a model for folding intermediates, the cyt *c* molten globule may serve as a valuable template for understanding the molecular determinants of cooperativity.

Natively unstructured proteins are central to many cellular processes and may be responsible for some human diseases.[1,20] Studies of unfolded, partially folded, and molten globule states are important for understanding the forces that determine different protein folds as well as the origin of their malignant behavior. De novo designed protein models have predetermined secondary structure and mimic many functional properties of natural proteins, but often lack rigid frameworks.[21] The structural features of these artificial proteins are reminiscent of those of molten globules: despite many secondary and tertiary contacts, the potential energy surface is broad, giving rise to substantial conformational heterogeneity. More work needs to be done to understand fully the nature of key interactions that preferentially stabilize only a limited set of molecular conformations and thus lead to a stable three-dimensional protein structure.

Acknowledgment. This work was supported by NIH GM068461 (J.R.W.) and the Ellison Medical Foundation (Senior Scholar Award in Aging to H.B.G.).

Supporting Information Available: Comparison of $P(k)$ distributions obtained with LSQNONNEG and ME analyses; Na_2SO_4-induced changes in $P(r)$ distributions; potential energy functions $U(r)$ for the native and molten globule states; and comparison of $P(r)$ distributions for the native state, molten globule, the transient folding species, and GuHCl-unfolded state in Dns4cyt, Dns39cyt, and Dns99cyt (5 pages, print/PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Ptitsyn, O. B. *Adv. Protein Chem.* **1995**, *47*, 83−229.
(2) Ohgushi, M.; Wada, A. *FEBS Lett.* **1983**, *164*, 21−24.
(3) Kataoka, M.; Hagihara, Y.; Mihara, K.; Goto, Y. *J. Mol. Biol.* **1993**, *229*, 591−596.
(4) Wilkins, D. K.; Grimshaw, S. B.; Receveur, V.; Dobson, C. M.; Jones, J. A.; Smith, L. J. *Biochemistry* **1999**, *38*, 16424−16431.
(5) Akiyama, S.; Takahashi, S.; Kimura, T.; Ishimori, K.; Morishima, I.; Nishikawa, Y.; Fujisawa, T. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 1329−1334.
(6) Cinelli, S.; Spinozzi, F.; Itri, R.; Finet, S.; Carsughi, F.; Onori, G.; Mariani, P. *Biophys. J.* **2001**, *81*, 3522−3533.
(7) Goto, Y.; Calciano, L. J.; Fink, A. L. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 573−577.
(8) Sosnick, T. R.; Mayne, L.; Hiller, R.; Englander, S. W. *Nat. Struct. Biol.* **1994**, *1*, 149−156.
(9) Jeng, M.-F.; Englander, W.; Elöve, G. A.; Wand, A. J.; Roder, H. *Biochemistry* **1990**, *29*, 10433−10436.
(10) Marmorino, J. L.; Lehti, M.; Pielak, G. J. *J. Mol. Biol.* **1998**, *275*, 379−388.
(11) Lyubovitsky, J. G.; Gray, H. B.; Winkler, J. R. *J. Am. Chem. Soc.* **2002**, *124*, 14840−14841.
(12) Pletneva, E. V.; Gray, H. B.; Winkler, J. R. *J. Mol. Biol.* **2005**, *345*, 855−867.
(13) Marmorino, J. L.; Pielak, G. J. *Biochemistry* **1995**, *34*, 3140−3143.
(14) Jia, Y. W.; Talaga, D. S.; Lau, W. L.; Lu, H. S. M.; DeGrado, W. F.; Hochstrasser, R. M. *Chem. Phys.* **1999**, *247*, 69−83.
(15) Santucci, R.; Bongiovanni, C.; Mei, G.; Ferri, T.; Polizio, F.; Desideri, A. *Biochemistry* **2000**, *39*, 12632−12638.
(16) Jordan, T.; Eads, J. C.; Spiro, T. G. *Protein Sci.* **1995**, *4*, 716−728.
(17) Lyubovitsky, J. G.; Gray, H. B.; Winkler, J. R. *J. Am. Chem. Soc.* **2002**, *124*, 5481−5485.
(18) DeWeerd, K.; Grigoryants, V.; Sun, Y.; Fetrow, J. S.; Scholes, C. P. *Biochemistry* **2001**, *40*, 15846−15855.
(19) Colón, W.; Roder, H. *Nat. Struct. Biol.* **1996**, *3*, 1019−1025.
(20) Wright, P. E.; Dyson, H. J. *J. Mol. Biol.* **1999**, *293*, 321−331.
(21) Lieberman, M.; Tabet, M.; Sasaki, T. *J. Am. Chem. Soc.* **1994**, *116*, 5035−5044.

JA0555318