Fast Chain Contraction during Protein Folding: "Foldability" and Collapse Dynamics

Linlin Qiu, Cherian Zachariah,* and Stephen J. Hagen†

University of Florida, Physics Department, P.O. Box 118440, Gainesville, Florida 32611-8440, USA

(Received 14 November 2002; published 25 April 2003)

Theory indicates that at least some proteins will undergo a rapid and unimpeded collapse, like a disorganized hydrophobic chain, prior to folding. Yet experiments continue to find signs of an organized, or barrier-limited, collapse in even the fastest ($\sim\mu$s) folding proteins. Does the kinetic barrier represent a signature of the equilibrium "foldability" of these molecules? We have measured the rate of chain contraction in two nonfolding analogs of a very fast-collapsing protein. We find that these chains contract on the same time scale ($\sim 10^{-5}$s) as the natural protein, and both pass over an energetic barrier at least as large as that encountered by the protein. The equilibrium foldability of the native structure therefore does not alone determine the dynamics of collapse; even the disordered chains contract $\sim$1000$\times$ slower than expected for an ideal chain.

DOI: 10.1103/PhysRevLett.90.168103 PACS numbers: 87.15.He, 87.15.Cc, 87.64.Ni

The initial contraction of a polypeptide chain from an expanded to a more compact form constitutes the earliest stage of protein folding. Like the rate of folding, the rate of this initial hydrophobic collapse can vary over orders of magnitude, depending on the amino acid sequence of the chain. Even for chains of comparable length ($\sim 10^2$ residues), collapse may occur within tens of microseconds or within $\sim$0.6–10 ms, or even as slowly as $\simeq$ 8 s [1,2] after the initiation of folding. Simulations of folding in model chains suggest two extreme cases for this initial phase [3]: a chain may acquire nativelike (i.e., folded) structure as it contracts, in a so-called "specific" collapse, or it may contract rapidly and "nonspecifically"—without forming nativelike contacts—and then reorganize into its folded form. Distinguishing between specific and nonspecific collapse has thus become an important and controversial element in the interpretation of dynamical phenomena observed in the earliest stages of protein folding. The very slowest collapse processes presumably fit the specific model. The most rapid ($\sim\mu$s) collapse processes (as in, e.g., cytochrome *c*), by some interpretations, then represent the opposite case, a free nonspecific or "polymeric" contraction, triggered by transfer of the disordered chain from good-solvent to poor-solvent conditions [4].

However, the experimental data for even the fastest-collapsing proteins also show some evidence for concurrent structure formation [1,5]. Several authors argue that these rapid contractions coincide with the formation of significant nativelike structure, and that collapse therefore requires passage over an entropic barrier; the resulting compact configuration then represents the first intermediate state on the path to the folded form [1,5]. Observation of exponential collapse kinetics, nonglobular Kratky plots in x-ray scattering, far-UV circular dichroism, or proton exchange protection at early times in such proteins all imply to some extent that these proteins collapse through the entropic bottleneck of structure formation, or equivalently that a truly disordered collapse would proceed differently, i.e., more rapidly. This raises the question: If even the fastest events carry signatures of barrier-crossing kinetics, do all "foldable" amino acid sequences invariably collapse in a specific process? More generally, does the existence of a stable native structure for an amino acid sequence dictate the dynamics of the initial collapse?

We have used laser temperature-jump spectroscopy to measure the rate of expansion and contraction in two nonfolding, disordered analogs of a fast-collapsing and fast-folding protein. We studied the polypeptide chains that comprise residues 1-65 and 1-80, respectively, of horse cytochrome *c*. These fragments, obtained by chemical cleavage of the intact 104-residue protein [6], lack the stabilizing *C*-terminal α helix (residues 88-100) of the protein and consequently exhibit little secondary or tertiary structure at equilibrium: their circular dichroism spectra more nearly resemble thermally denatured cytochrome *c* than they do the native protein [Fig. 1(a)] [4,6], they show little evidence of stable H bonding, and they lack those residues that do form such bonds early in the folding of the intact protein. They do, however, adopt a partially compact configuration, as evidenced by a weak fluorescence (see below) and an intrinsic viscosity ($[\eta] =$ 6.6 ml/g for 1-65 at *p*H 7, 25 °C) that lies intermediate between the value for a folded globule (2.7 ml/g) and a typical random coil ($\simeq$11.2 ml/g for 65 residues) [6].

Tryptophan (Trp) fluorescence data show that these fragments expand gradually with rising T or increased denaturant. In compact configurations of these chains, the heme group (covalently attached at residues 14, 17, and 18) quenches the fluorescence of the single Trp residue at position 59 through the distance-dependent Förster (dipole-dipole) mechanism. Figure 1(b) shows that the emission of Trp-59 rises $\simeq$ 35% upon the addition of 0.5 M GdnHCl, although the characteristic Förster distance ($R_0 \approx$ 3.4 nm) remains virtually constant (since the heme absorption and Trp emission spectra and the solvent refractive index scarcely change); the increasing emission

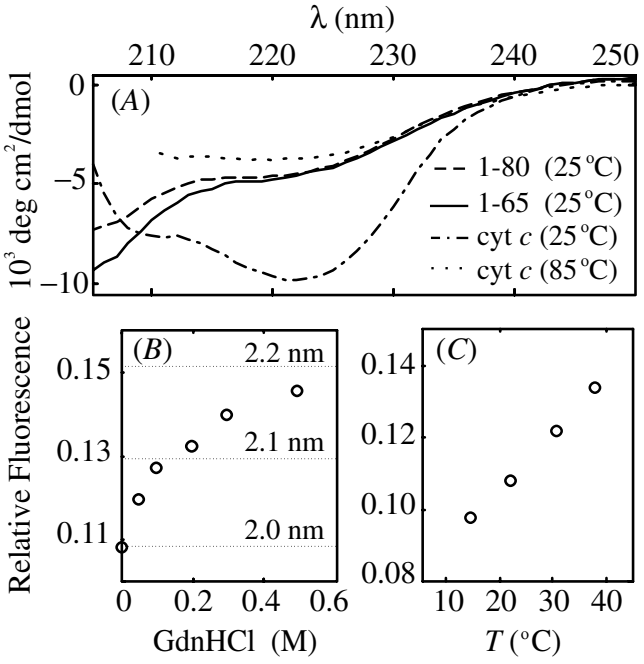

FIG. 1. (a) Equilibrium far-UV circular dichroism of the 1-65 and 1-80 fragments of the 104-residue horse cytochrome *c* at $T = 25°C$ (in 0.1 M phosphate, *p*H 7), showing secondary-structure content substantially less than that of the intact folded protein, but comparable to that of the thermally denatured protein at 85°C [6]; (b) Trp fluorescence of 1-65 fragment, relative to a reference peptide representing the Trp-containing region (residues 55–63, sequence *KGITWKEET*) of intact cytochrome *c*. Rising emission at higher denaturant (GdnHCl) levels indicates reduced Förster quenching of Trp-59 fluorescence by the heme, as a result of chain expansion. Dotted lines show average fluorescence of a Gaussian ensemble of random coils having the indicated mean heme-Trp distance. (c) *T* dependence of Trp fluorescence of 1-65 fragment in 0.1 M phosphate, *p*H 7, relative to the reference peptide.

indicates a growing heme-Trp separation and thus an expansion of the disordered coil. With rising temperature, the fluorescence of the fragments falls, but it falls more slowly than does that of a reference Trp peptide [Fig. 1(c)], indicating that gradual swelling of the chains reduces the Förster quenching of Trp-59, and thus partially compensates for the intrinsic, negative *T* dependence of Trp fluorescence. Our *T*-jump apparatus time-resolves this thermally induced expansion.

Methods.—We prepared the 1-65 and 1-80 peptides by cyanogen bromide cleavage of horse ferricytochrome *c* (Sigma Chemical Co.) at the unoxidized Met-65 or Met-80 as described previously [6], purified the peptides by reverse phase high performance liquid chromatography on a C18 column, and used mass spectrometry to verify the molecular weight of the fragments. The laser temperature-jump method applies an infrared laser pulse to heat the protein sample quickly, and then time resolves the Trp fluorescence changes that follow this perturbation. Our *T*-jump apparatus uses a 1 m, 650 psi H_2 cell to Raman shift the 5–7 ns pulses of a Nd:YAG laser (1.064 μm) to 1.89 μm; these pulses, when formed into pairs of counterpropagating pulses and focused onto the aqueous sample solution flowing within a silica capillary (100 μm inner diameter), generate a temperature jump of (5–20) °C within 20–30 ns. After a time delay *t*, a 5 ns pulse from the 266 nm 4th harmonic of a Nd:YAG laser excites the Trp fluorescence of the sample. A microscope objective (10X, NA 0.28) positioned at 90° projects the emitted light onto a photomultiplier. The sample temperature remains uniform and elevated for ~1 ms and returns to its pretrigger value by ~400 ms. A syringe pump continuously replaces the sample in the capillary. Measurements at 50–100 values of the delay *t* in random order build up a complete kinetic trace. We use a fluorescence standard of known *T* dependence to measure the *T*-jump amplitude.

Results.—Figure 2 shows that the rapid temperature jump triggers an abrupt drop in the fluorescence signal (at $t = 0$) owing to the intrinsic negative *T* dependence of the Trp fluorescence. As expected from equilibrium fluorescence data, however, the emission from the 1-65 and

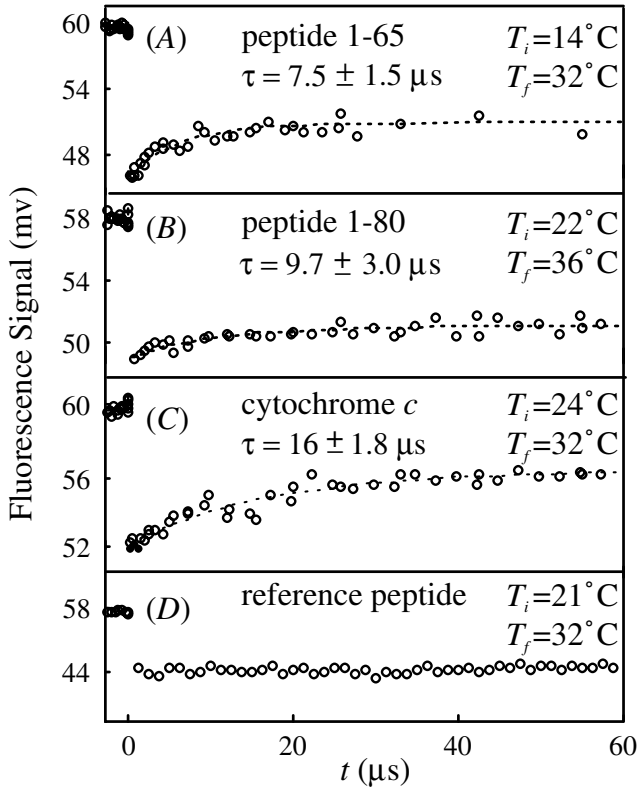

FIG. 2. Relaxation processes observed in tryptophan fluorescence following nanosecond *T* jump at $t = 0$, for (a) 1-65 and (b) 1-80 fragments of cytochrome *c* in 0.1 M phosphate, *p*H 7; (c) intact cytochrome *c* in 1.5 M GdnHCl, 0.1 M phosphate, *p*H 7; (d) tryptophan-containing reference peptide (*KGITWKEET*) in 0.1 M phosphate *p*H 7.

1-80 peptides then grows in a small-amplitude relaxation as the partially denatured coils expand to a new equilibrium size. Fitting this relaxation to an exponential decay gives a time constant $\tau \simeq 25$ μs at $T = 22\,^{\circ}$C and $\simeq 8$ μs at $T = 32\,^{\circ}$C in the 1-65 peptide, and a slightly slower τ in 1-80. A control experiment on a reference Trp peptide, which corresponds to residues 55-63 of the cytochrome *c* sequence, finds no such compensating relaxation (Fig. 2). Furthermore, the addition of as little as 0.5 M GdnHCl suppresses the relaxation amplitude in the 1-65 and 1-80 peptides to nearly unobservable small values.

The relaxation time [7] depends strongly on the final (post-T-jump) temperature T (Fig. 3). Over the experimental range $T = 17$–$38\,^{\circ}$C, τ varies in a roughly Arrhenius manner with activation energy $E_a \simeq 57 \pm 7$ kJ/mol for 1-65, and $E_a \simeq 68 \pm 13$ kJ/mol for 1-80. Although τ does vary with the final temperature, we find it insensitive to the initial temperature (i.e., to the size of the T jump) for $\Delta T \simeq 5 - 20\,^{\circ}$C. Thus we find that, under a particular solvent condition, these partially compact but poorly structured chains respond to solvent perturbations by changing their average dimension with a characteristic relaxation time $\tau \sim 10^{-5}$ s.

We compare this behavior with the collapse of the intact protein. Denatured cytochrome *c* exhibits a similar relaxation on a time scale ~ 60 μs at 22 °C (at *p*H 7, without denaturant), and $\sim 2\times$ faster in the presence of 1.5 M GdnHCl denaturant [1]. In a T-jump experiment this appears as a fast but fairly weak increase in Trp fluorescence of the few partially denatured molecules present at equilibrium (Fig. 2), followed by large-scale unfolding on millisecond time scales. The rate of this relaxation also varies with temperature, with $E_a \sim$ 30–40 kJ/mol (Fig. 3 and Ref. [1]). Thus we find similar dynamics in the fast reconfigurations of the folding and nonfolding variants of the protein: cytochrome *c* and its fragments both respond to solvent changes through a fast relaxation on a time scale $\sim 10^{-5}$ s, with an apparent activation energy substantially larger than $k_B T$ (≈ 2.5 kJ/mol). If the formation of nascent structure slows the expansion and collapse of cytochrome *c* relative to that of its nonfolding analogs, this has a fairly weak effect on the dynamics of collapse.

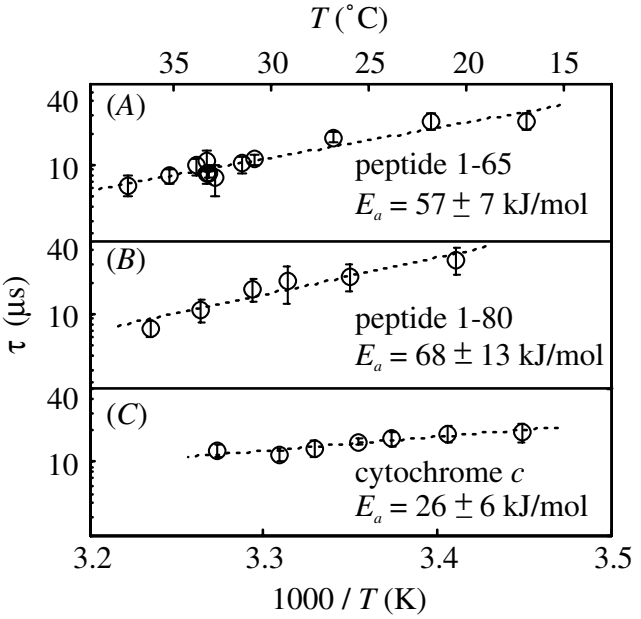

FIG. 3. Relaxation time vs T, and Arrhenius fit (broken line) observed after laser T jump in (a) 1-65 and (b) 1-80 fragments of cytochrome *c* in *p*H 7.0 buffer, and (c) wild-type cytochrome *c* in 1.5 M denaturant (GdnHCl), *p*H 7.0. The relaxation time for intact protein slows ~2-fold in the absence of denaturant.

What sets the time scale for expansion and collapse of these chains? The kinetics and dynamics of transitions between expanded and compact configurations of heteropolymers such as proteins remain poorly understood. Although a number of theoretical models have addressed homopolymer collapse dynamics, these primarily describe the collapse of long flexible chains; they predict collapse to occur by a sequential aggregation on many length scales [8]. The 1-65 and 1-80 polypeptides, which represent only 14–18 persistence lengths of polypeptide, presumably lack the flexibility to undergo this progressive collapse. Furthermore, such models typically predict the collapse time to increase for deeper quenches: the simple de Gennes model [8], for example, predicts the collapse time $\tau \simeq \tau_R \Delta T / T_\Theta$ associated with a T quench to scale with the quench depth ΔT and the Rouse time τ_R. Our data, and rapid-mixing studies [1], give no indication that pretrigger solvent composition or temperature affect collapse kinetics in cytochrome *c*.

By contrast, the Zimm model gives a characteristic relaxation time for an ideal or noninteracting chain, following a perturbation [9]. The slowest conformational relaxations occur on a time scale $\tau_z \simeq \eta_0 R^3 / k_B T$, which scales with the solvent viscosity η_0 and the rms separation $R = \sqrt{\langle r^2 \rangle}$ of the chain end points. Eliminating R in favor of the experimentally determined $[\eta]$ and the molecular weight M of the fragments, this becomes $\tau_z \simeq \eta_0 M [\eta] / N_A k_B T$ or $\tau_z \simeq 21$ ns for 1-65 and $\tau_z \simeq 24$ ns for 1-80 in water at 22 °C.

This prediction falls far short ($\sim 1000\times$) of the measured relaxation times, and it fails to account for the strong T dependence observed in τ. Although fully extended polypeptides may relax on the nanosecond ideal-chain time scale, τ_z (as perhaps occurs in MD studies or in elemental helix formation [10]), the more complicated process of reorganizing a polypeptide into a compact configuration evidently entails a far slower time scale. Energetic and entropic costs for burying hydrophobic surface may play a role, even in the nonfolding peptides. Furthermore, ideal-chain models do not take account of potential energy barriers to bond rotation in

the chain backbone, the energetics of contact between sequence-distant residues, accessibility of void space in the solvent, and other such forces affecting the chain [11]. This suggests that most denatured polypeptides probably cannot explore their compact configurations at rates much faster than those observed here; the very fastest (i.e., lowest barrier) folding phenomena observed in complete proteins in fact play out on the $\sim 10^{-5}$ s time scale [10,12].

Our results show that the equilibrium stability of the native state does not control the collapse/expansion dynamics of a polypeptide: even destabilized, nonfoldable chains can encounter substantial barriers to expansion and contraction. This finding sheds a useful light on the controversy over the specific vs nonspecific character of fast polypeptide collapse. Several lines of evidence now show that proteins such as cytochrome *c* pass over a free-energy barrier while rapidly collapsing to a compact form. If classifying this process as specific or nonspecific rests on the question of whether nativelike intrachain contacts form, i.e., if it implicitly assumes that nonfolding molecules cannot possibly collapse specifically, then the classification does not allow for all the possible scenarios, including barrier-limited collapse of molecules that lack any stable native structure. Understanding the physical origin of these dynamics will involve more than just categorizing them as specific or nonspecific.

Along these lines, several authors have proposed that the compaction of any polypeptide may favor the development of weak secondary structure, and that the extent of the "burst phase" contractions seen in proteins correlates with the increase of secondary structure content [13]. Thus the formation of small amounts of secondary structure may generate a free-energy obstacle to any compaction, in either folding or nonfolding chains; in this case, suitable time-resolved studies will find some evidence of structure formation during even the fastest protein collapse events.

The authors gratefully acknowledge funding support from the National Science Foundation, MCB No. 0077907.

*Department of Biochemistry and Molecular Biology, University of Florida, Gainesville, FL.
†Email address: sjhagen@ufl.edu

[1] M. C. R. Shastry and H. Roder, Nat. Struct. Biol. **5**, 385 (1998); S. Hagen and W. Eaton, J. Mol. Biol. **301**, 1037 (2000).

[2] K. W. Plaxco *et al.*, Nat. Struct. Biol. **6**, 554 (1999); S. H. Park, M. C. Shastry, and H. Roder, Nat. Struct. Biol. **6**, 943 (1999); L. Pollack *et al.*, Proc. Natl. Acad. Sci. U.S.A. **96**, 10115 (1999); L. Pollack *et al.*, Phys. Rev. Lett. **86**, 4962 (2001).

[3] D. Thirumalai, J. Phys. I **5**, 1457 (1995); A. Gutin, V. Abkevich, and E. Shakhnovich, Biochemistry **34**, 3066 (1995); J. Chahine *et al.*, Phys. Rev. Lett. **88**, 168101 (2002).

[4] V. R. Agashe, M. C. R. Shastry, and J. B. Udgaonkar, Nature (London) **377**, 754 (1995); T. R. Sosnick, M. D. Shtilerman, L. Mayne, and S. W. Englander, Proc. Natl. Acad. Sci. U.S.A. **94**, 8545 (1997); P. X. Qi, T. R. Sosnick, and S. W. Englander, Nat. Struct. Biol. **5**, 882 (1998); B. A. Krantz *et al.*, J. Mol. Biol. **324**, 359 (2002).

[5] B. Nölting *et al.*, Proc. Natl. Acad. Sci. U.S.A. **94**, 826 (1997); J. M. Sauder and H. Roder, Fold. Des. **3**, 293 (1998); S. Akiyama *et al.*, Proc. Natl. Acad. Sci. U.S.A. **99**, 1329 (2002); S. Akiyama, S. Takahashi, K. Ishimori, and I. Morishima, Nat. Struct. Biol. **7**, 514 (2000).

[6] J. Babul, E. B. McGowan, and E. Stellwagen, Arch. Biochem. Biophys. **148**, 141 (1972); C. Bryant and E. Stellwagen, J. Biol. Chem. **260**, 332 (1985).

[7] In the absence of a free-energy barrier to expansion, the observed relaxation time τ characterizes *either* expansion or collapse following a small perturbation [S. J. Hagen, Proteins Struct. Funct. Genet. **50**, 1 (2003)]. In the case of a barrier-limited process, with separate expansion and contraction time scales, $\tau^{-1} = \tau_{\mathrm{exp}}^{-1} + \tau_{\mathrm{contr}}^{-1}$ [*Protein Folding*, edited by T. E. Creighton (W. H. Freeman and Co., New York, 1992)]. Then $\tau \rightarrow \tau_{\mathrm{contr}}$ under solvent conditions that favor contraction, as in the present study.

[8] P. G. de Gennes, J. Phys. (Paris) Lett. **46**, L639 (1985); A. Halperin and P. M. Goldbart, Phys. Rev. E **61**, 565 (2000); C. F. Abrams, N.-K. Lee, and S. P. Obukhov, Europhys. Lett. **59**, 391 (2002).

[9] M. Doi and S. F. Edwards, *The Theory of Polymer Dynamics* (Oxford, New York, 1986).

[10] W. A. Eaton *et al.*, Annu. Rev. Biophys. Biomol. Struct. **29**, 327 (2000).

[11] C. W. Manke and M. C. Williams, Macromolecules **18**, 2045 (1985); J. J. Portman, S. Takada, and P. G. Wolynes, J. Chem. Phys. **114**, 5082 (2001); S. J. Hagen, C. W. Carswell, and E. M. Sjolander, J. Mol. Biol. **305**, 1161 (2001).

[12] R. M. Ballew, J. Sabelko, and M. Gruebele, Proc. Natl. Acad. Sci. U.S.A. **93**, 5759 (1996); L. Qiu, S. A. Pabit, A. E. Roitberg, and S. J. Hagen, J. Am. Chem. Soc. **124**, 12 952 (2002).

[13] N. D. Socci, W. S. Bialek, and J. N. Onuchic, Phys. Rev. E **49**, 3440 (1994); D. P. Yee, H. S. Chan, T. F. Havel, and K. A. Dill, J. Mol. Biol. **241**, 557 (1994); N. G. Hunt, L. M. Gregoret, and F. E. Cohen, J. Mol. Biol. **241**, 214 (1994); A. Maritan, C. Micheletti, and J. R. Banavar, Phys. Rev. Lett. **84**, 3009 (2000); J. R. Banavar *et al.*, Proteins **47**, 315 (2002); V. N. Uversky and A. L. Fink, FEBS Lett. **515**, 79 (2002).