Direct Observation of the Three-State Folding of a Single Protein Molecule

Ciro Cecconi,1,2* Elizabeth A. Shank,1* Carlos Bustamante,1,2,3† Susan Marqusee1†

We used force-measuring optical tweezers to induce complete mechanical unfolding and refolding of individual Escherichia coli ribonuclease H (RNase H) proteins. The protein unfolds in a two-state manner and refolds through an intermediate that correlates with the transient molten globule–like intermediate observed in bulk studies. This intermediate displays unusual mechanical compliance and unfolds at substantially lower forces than the native state. In a narrow range of forces, the molecule hops between the unfolded and intermediate states in real time. Occasionally, hopping was observed to stop as the molecule crossed the folding barrier directly from the intermediate, demonstrating that the intermediate is on-pathway. These studies allow us to map the energy landscape of RNase H.

Protein folding remains a major unsolved challenge for modern molecular biology. Theoretical studies emphasize the potential heterogeneous nature of the process; however, traditional bulk biochemical experiments often mask this complexity in their inherent ensemble averaging. For instance, many proteins are observed to populate partially structured con-formational basins with associated changes in solution properties. Such “solution basins” can be probed by a variety of bulk techniques. The central portion of the poly

*Department of Molecular and Cell Biology and Institute for Quantitative Biology, 1Department of Physics, Howard Hughes Medical Institute, University of California, Berkeley, CA 94720, USA.
†These authors contributed equally to this work.
‡To whom correspondence should be addressed. E-mail: carlos@alice.berkeley.edu; marqusee@berkeley.edu

16. Additional results, materials, and methods are available as supporting material on Science Online.
36. Additional results, materials, and methods are available as supporting material on Science Online.
and kinetic data makes RNase H an ideal model system to explore single-molecule protein folding trajectories and the nature of protein folding intermediates.

We used two 500-base pair double-stranded DNA molecules to tether individual RNase H molecules between two polystyrene spheres ~2 μm in diameter (Fig. 1A) (19). These “molecular handles,” attached to the protein through unique engineered cysteine residues, function as spacers to prevent nonspecific bead-bead interactions and permit manipulation of the ~3 nm diameter protein. RNase H is folded and retains activity when attached to DNA handles (fig. S1). By moving the bead on the pipette relative to the bead in the trap, each RNase H molecule was stretched and relaxed multiple times, generating force-extension curves (Fig. 1B) (20). We observed sudden changes in extension of the molecule (transitions) during both stretching and relaxation, corresponding to the unfolding and refolding of the protein. These transitions were not observed during control experiments with DNA handles alone (yellow trace, Fig. 1B).

Unfolding transitions occurred at two distinct forces, ~19 pN or ~5.5 pN, whereas the refolding transitions exhibited a single narrow distribution centered at ~5.5 pN, coincident with the lower of the two unfolding transition forces (fig. S2).

The increment in contour length (ΔLc) corresponding to each transition was estimated by fitting the force-extension curves with the worm-like chain model (21). The high-force transitions at ~19 pN were consistent with the complete unfolding of RNase H (ΔLc of 50 ± 5 nm versus 50.4 nm calculated for the full-length protein). In contrast, the ΔLc obtained from the low-force unfolding (39 ± 6 nm) and refolding transitions (40 ± 10 nm) imply that the associated conformational changes involve a smaller portion of the polypeptide chain. Indeed, close inspection of the force-extension curves reveals that the refolding transition does not restore the original length of the molecule, leaving a gap between the stretching and relaxation curves at ~5.5 pN (Fig. 1B) and supporting the idea of partial refolding. Together, these data indicate that the native RNase H structure completely unfolds at ~19 pN (N, native → U, unfolded), and then, upon return to ~5.5 pN, partially refolds into an intermediate structure (U → I). The refolding intermediate unfolds again at ~5.5 pN in the next stretching cycle, unless the relaxed protein is incubated for a time sufficient to fully refold (I → N) before being stretched again.

We therefore monitored whether the protein had fully refolded into its native structure by examining subsequent stretching curves for the presence of a high-force unfolding transition. The probability that the intermediate folds into the native state between successive cycles while held at force F for a period of time t is given by

\[ P_f(F, t) = 1 - \exp\left(-t k_{obs}(I→N) \exp[-(F \Delta x_{I→N}^T)/k_B T]\right) \]

(22), where \( \Delta x_{I→N}^T \) is the distance from the intermediate to the transition state, \( k_{obs}(I→N) = k_m k_{obs}(I→N)^{1/N} \) [where \( k_m \) reflects any possible contribution of the instrument to the overall refolding rate (12) and \( k_{obs}(I→N) \) is the rate of the I → N transition at zero force], and \( k_B T \) is the product of the Boltzmann constant and absolute temperature. The probability of the I → N transition was determined as a function of force and incubation time, and the data were fit with a linearized form of Eq. 1 (Fig. 2A), yielding \( k_{obs}(I→N) = 0.17 \pm 0.03 \text{ s}^{-1} \) and \( \Delta x_{I→N}^T = 1.5 \pm 0.3 \text{ nm} \). Bulk studies on RNase H have also revealed a transient folding intermediate with a similar refolding rate (16).

The hysteresis observed between the unfolding of the native state (N → U) and the refolding transitions suggests that the rate of pulling in these experiments is faster than the rate at which these states can equilibrate under these conditions. By pulling the molecule at different loading rates, the observed unfolding rate of the molecule at zero force is found to vary by a factor of 13 pN s^{-1} (16).
be \( k_{\text{obs}}(N\rightarrow U) = 3 \times 10^{-4} \text{ s}^{-1} \), and the distance between the native state and the first transition state is \( \Delta x_{N\rightarrow U} = 2.0 \pm 0.1 \text{ nm} \) (Fig. 2B) (23) when analyzed in a manner analogous to that used in previous AFM studies (24). This value of \( k_{\text{obs}}(N\rightarrow U) \) corresponds well with the \( k_{\text{obs}}(N\rightarrow U) \) value, 1.7 (±0.04) \( \times 10^{-5} \text{ s}^{-1} \) (16).

In contrast to the hysteresis observed for the N \( \rightarrow \) U transition, the unfolding and refolding transitions of the mechanical refolding intermediate (I \( \rightarrow \) U and U \( \rightarrow \) I) coincide (Fig. 1B), indicating that this process occurs reversibly under the experimental conditions used here (Fig. S3). Consistent with this observation, the force-extension curves occasion-ally displayed rapid fluctuations in extension near 5.5 pN rather than a single sharp transition (Fig. S4). We examined this behavior further by relaxing the unfolded protein and holding the polypeptide at a fixed force with the use of the instrument’s force-feedback mode (12). When held at a force near 5.5 pN, the RNase H molecule displayed bistability: The molecule “hopped” between the intermediate and the denatured form of the protein in real time, with the molecular extension shifting rapidly by 15 ± 4 nm (Fig. 3A). Changing the set point of the force altered this equilibrium between the extended unfold state and the compact intermediate structure (Fig. 3A). Such hopping has also been seen for a simple RNA hairpin using a similar approach (12), and for a chemically destabilized two-state (U \( \Leftrightarrow \) N) protein monitored for short time periods by fluorescence resonance energy transfer (25). We now show the extended real-time equilibrium behavior of a complex globular protein.

The force-dependent rates of unfolding and refolding of the intermediate were determined from the lifetimes of the extended (U) and compact (I) states seen in the hopping experiments. The position of the transition state between I and U was then estimated by fitting the rates to the Arrhenius-like equation

\[
k_{\text{I}\rightarrow \text{U}} = k_0 \exp(-\Delta G_{\text{I}\rightarrow \text{U}}/k_B T)
\]

where \( k_0 \) is the unfolding rate at zero force, and \( \Delta G_{\text{I}\rightarrow \text{U}} \) is the distance from I to the transition state along the reaction coordinate (12). A similar analysis holds true for the reverse U \( \rightarrow \) I reaction. The slopes of plots of \( \ln k \) versus force yielded \( \Delta G_{\text{I}\rightarrow \text{U}} = 5 \pm 1 \text{ nm} \) and \( \Delta G_{\text{U}\rightarrow \text{I}} = 6 \pm 1 \text{ nm} \). These values are substantially larger than those found for the native-state unfolding (\( \Delta G_{\text{N}\rightarrow \text{U}} \)) of several other proteins using the AFM (26). Our results on RNase H therefore convey a picture of the intermediate as a pliable structure that can deform elastically a great amount before the reaction is committed to unfolding. Such large transition-state distances have been observed in other biomolecules and were interpreted as reflecting the mechanical behavior of structures lacking the nonlocal specific contacts associated with tertiary interactions (6, 12). Thus, the large distance from the intermediate to its transition state and the low forces required to unfold it suggest that this structure is only able to form weak, possibly transient tertiary interactions, and that it therefore resembles a molten globule structure (27).

We evaluated the thermodynamics of the U \( \Leftrightarrow \) I transition (\( \Delta G_{\text{U}\rightarrow \text{I}} \)) with three independent methods (Fig. 3B). We measured the area under the refolding plateau in the force-extension curves, calculated the force-dependent equilibrium constant between U and I, and analyzed the transitions with the statistics of a two-state system. After correcting for the entropy loss due to tethering and stretching the U state (calculated to be 5.1 ± 0.6 kcal/mol (19)), we obtained \( \Delta G_{\text{U}\rightarrow \text{I}} = 4 \pm 3 \text{ kcal/mol} \), \( 4 \pm 2 \text{ kcal/mol} \), and \( 3.8 \pm 0.8 \text{ kcal/mol} \), respectively. The agreement between these free energy values validates the analyses. The remarkable similarity between these values and that observed in ensemble experiments (\( \Delta G_{\text{U}\rightarrow \text{I}} \)) strengthens the conclusion that the intermediate detected in our single-molecule mechanical manipulations correlates with that sampled in solution.

To further probe this relationship, we performed similar optical tweezer experiments using a variant of RNase H (I53D) that displays two-state folding in solution (28). In this case we saw no evidence of an intermediate (Fig. 3C) (29). These results strongly suggest that the intermediate observed in our single-molecule experiments is similar to that detected by ensemble methods, and hence the new features found in this study are relevant to the folding process in solution.

In the ensemble refolding experiments of RNase H, the formation of I occurs very
Rapidly (within the 12-ms dead time of the stopped-flow instrument) and is therefore not observed directly (16). It has thus been impossible to determine unambiguously whether this kinetic intermediate is a distinct thermodynamic state or simply the result of a shift in the population of a continuum of unfolded states. Both models have been suggested from experimental and theoretical studies (4). In our studies, we used force to modulate the equilibrium between the U and I states until the interconversion between these two forms became observable. Our data clearly indicate that the transition between the I and U states is first-order (Fig. 3A) (fig. S5), allowing us to directly characterize the kinetic, thermodynamic, and mechanical properties of this molten globule-like state. The forces holding together the intermediate are small but still substantial (~5.5 pN), amounting to about one-third of the forces that stabilize the fully folded state at the lower force.

Sometimes during constant force experiments, the hopping corresponding to the U ⇔ I transition spontaneously ceased. An additional compaction always preceded the termination of hopping (Fig. 3D). The size of this compaction, as estimated from the worm-like chain model, corresponds well to that expected for the I → N transition at the given force. Indeed, stretching the molecule after hopping ceased invariably resulted in a high-force unfolding transition (~19 pN), as expected for the unfolding of the native state. Thus, it was possible to observe the I → N transition directly. In 78% of the traces (n = 18) in which hopping stopped, the transition to the native state clearly took place from the folded intermediate structure, as in Fig. 3D. In the rest of the cases, the time spent in I before the transition to N may have been too short to be resolved in our experiments.

These data indicate that the refolding intermediate observed in our experiments exists on-pathway to the folded state. Furthermore, the fact that the U → I transition is invariably present in our force-relaxation curves indicates that the same intermediate is also an obligatory step in the folding trajectory of RNase H (Fig. 4A).

The ability to explore the behavior of single protein molecules with the use of optical tweezers has permitted us to map the energy landscape traversed by the small globular protein RNase H in its transitions to and from its unfolded state (Fig. 4B). Our observation of a folding intermediate that corresponds to the intermediate seen in bulk suggests that both methods probe the same fundamental barriers. The new features revealed in this study therefore enhance our understanding of how proteins fold in solution. The intermediate of RNase H forms a distinct thermodynamic state that, although compact and held together by cohesive interactions, is nonetheless highly deformable. This state appears to be both on-pathway and obligatory to the folding trajectory.

Fig. 4. The energy landscape of RNase H. (A) Schematic representation (37) of the energy landscape of RNase H depicting a reduction in conformational entropy as the protein folds through an on-pathway, obligatory, and productive intermediate species. (B) Free energy reaction profile of RNase H at ~5.5 pN, the force at which ΔG_{U-I} = 0. Relative distances between states correlate with the values obtained experimentally (Fig. 2, A and B, and Fig. 3A), assuming that ΔG_{U-I} obtained at a range of forces between 15 and 20 pN holds for this lower force.

References and Notes
19. See supporting data on Science Online.
20. Single-molecule tethers were recognized by identifying the oversstretching transition due to the DNA handles. Only those fibers that exhibited an overstretching transition of the expected length (~230 nm) at the correct force (~67 pN) for a single fiber containing 1116-base pair DNA were selected for analysis.
23. The ΔF_ΔG_{U-I} and ΔF_ΔG_{N-I} values represent the average width of the potential energy for each state along the mechanical reaction coordinate [6, 26] and thus describe the extent of structural distortion necessary to cross a transition-state barrier. A correlation has been noted between a protein's structure and the position of the unfolding transition state: ΔF_ΔG_{U-I} appears to be large when the stabilizing interactions are predominantly long-range hydrophobic interactions, and small when they consist of clusters of short-range interstrand hydrogen bonds [6, 30]. The value of 2 nm measured here for RNase H is on the higher end of the distribution of transition-state distances obtained for other proteins using the AFM (0.25 to 1.7 nm) [26] and references therein.
29. The IS3D RNase H variant did not exhibit any low-force unfolding transitions; refolding occurred as a gradual compaction rather than as a single sharp transition, and hopping was not observed at any force examined. Refolding to the native state was qualitatively slower for the IS3D variant, which suggests that the formation of the molten globule-like intermediate of RNase H speeds up the attainment of its folded state. This is consistent with bulk kinetic studies that suggest that folding can be accelerated if the interactions that stabilize the intermediate stabilize the transition state structure to an even greater extent (28).
32. We thank R. Dahlquist for support and experimental advice in the early stages of this work, and the Marqusee and Bustamante labs for helpful discussions. Supported by NIH grants GM05945 [S.M.] and GM12543 [C.B.]. U.S. Department of Energy grant AC0376Sf00098 (C.B.), NIH training grant GM008295 (E.A.S.), and a NSF predoctoral fellowship (E.A.S.).

Supporting Online Material
www.sciencemag.org/cgi/content/full/309/5743/2057/DC1
Materials and Methods
SOM Text
Figs. 51 to 56
28 June 2005; accepted 17 August 2005
10.1126/science.1116702