Signatures of hydrophobic collapse in extended proteins captured with force spectroscopy

Kirstin A. Walther*†, Frauke Gräter*‡, Lorna Dougan*, Carmen L. Badilla*, Bruce J. Berne‡§, and Julio M. Fernandez*§

Departments of *Biological Sciences, †Physics, and ‡Chemistry, Columbia University, New York, NY 10027

Contributed by Bruce J. Berne, March 13, 2007 (sent for review January 23, 2007)

We unfold and extend single proteins at a high force and then linearly relax the force to probe their collapse mechanisms. We observe a large variability in the extent of their recoil. Although chain entropy makes a small contribution, we show that the observed variability results from hydrophobic interactions with randomly varying magnitude from protein to protein. This collapse mechanism is common to highly extended proteins, including nonfolding elastomeric proteins like PEVK from titin. Our observations explain the puzzling differences between the folding behavior of highly extended proteins, from those folding after chemical or thermal denaturation. Probing the collapse of highly extended proteins with force spectroscopy allows separation of the different driving forces in protein folding.

atomic force microscopy | molecular dynamics | protein folding | single molecule

Proteins can reversibly fold from a random coil conformation into a well defined native state, a process constituting a major research area in biology. Traditionally, experiments involved varying the ambient environment, such as changing the temperature or pressure, or using denaturing chemicals. Protein folding probed under these conditions has revealed two-state folding for many small proteins (1–3). Based on such experiments, Wolynes and colleagues proposed that the energy landscape of a collapsing polypeptide is funnel-shaped under folding conditions (4–6). In this scenario, the protein's energy decreases as it forms favorable interactions, thus driving it toward the native state. In the classic folding experiments, as well as in the theoretical models, proteins start in the denatured state from collapsed random coil conformations that are only a few Ångström larger than their native state (7, 8). In these conformations, the side chains of the collapsed polypeptide are in close proximity to each other. It is widely accepted that, under such conditions, protein folding is driven mostly by hydrophobic interactions that are finely balanced by entropy (9–11). However, given that the denatured state in these experiments is not well defined, it has proved difficult to separate the hydrophobic, electrostatic, and entropic contributions to protein collapse and folding. We use single-molecule force-clamp spectroscopy to bring proteins to an extended conformation of >80% of their contour length, where the side chains are separated and exposed to the solvent, and native contact formation is rare (12–14). Thus, proteins are driven to the outer regions of the folding landscape, which have not been explored so far. Studying the collapse of such extended proteins greatly simplifies the folding dynamics, permitting a more direct identification of the major driving forces (11).

Highly extended proteins have been routinely described as entropic chains using models of polymer elasticity such as the worm-like chain (WLC) model (15) or the freely rotating chain model (16). In this simplified picture, the collapse of a protein from an extended state is driven by entropy, with the force vs. length relationship depending solely on the contour length and the stiffness of the protein chain (17). Furthermore, equilibration of an entropic chain to a given length is thought to occur on time scales much faster than those accessible in our experiments (18, 19). A unique collapse behavior can thus be expected for the entropic collapse of proteins of the same chain length.

We study here the mechanism of the initial collapse of unfolded and extended proteins using the force-quench and -ramp techniques (20). These experiments and comparison to entropic chain behavior from molecular simulations reveal signatures of hydrophobic collapse, which we find to be common to all proteins we studied.

Results and Discussion

We stretched single polyubiquitin proteins at an initial high force of 100 pN, causing the probabilistic unfolding of the individual ubiquitins in the chain, where each unfolding event is marked by a step increase in the protein's end-to-end length by 19.2 ± 0.5 nm (Fig. 1 and ref. 20). The resulting unfolding staircase serves both as a fingerprint for identifying single molecules as well as an indication that the polyprotein is fully unfolded (20). Subsequently, the force was quenched to a low value of 10 pN to monitor the collapse of the extended chain. Fig. 1 *A* and *B* shows the response of seven different polyubiquitin molecules to the exact same force protocol (Fig. 1*C*). Upon quenching the force, we observed a rapid initial collapse, which was previously attributed to entropic recoil (20). In some cases, this was followed by further contraction to a small size and protein folding (Fig. 1*A*). However, after the initial collapse, most proteins remained elongated and failed to fold (Fig. 1*B*). Despite the use of an identical force protocol in all cases, the magnitude of the initial collapse was surprisingly heterogeneous. This heterogeneity is particularly apparent when comparing the initial collapse of proteins that have very similar extended lengths (Fig. 1 *A* and *B*). Furthermore, the extent of the initial collapse was larger in proteins that folded (0.42 ± 0.22 of their length at 100 pN; Fig. 1 *Inset*, top histogram) than in those that failed to fold on the experimental time scale of several seconds (0.72 ± 0.12, lower histogram). These results are in disagreement with the assumption that entropic recoil is the sole mechanism driving the initial collapse of an extended protein. None of the current models of polymer elasticity can account for the observed variability in the initial collapse. To examine the nature of the forces driving protein collapse, we modified our force–clamp protocol by linearly decreasing the force to 10 pN over 4 sec, instead of the step-like quench (Fig. 2 *A–C*). The force-ramp protocol permits the observation of the full force-length relationship of the extended protein, rather than only the two force values measured with force quench (100 and 10 pN).

Author contributions: K.A.W., F.G., and L.D. contributed equally to this work; K.A.W., F.G., L.D., and J.M.F. designed research; K.A.W., F.G., and L.D. performed research; K.A.W., F.G., L.D., C.L.B., and B.J.B. contributed new reagents/analytic tools; K.A.W. and L.D. analyzed data; and K.A.W., F.G., L.D., and J.M.F. wrote the paper.

The authors declare no conflict of interest.

Abbreviations: WLC, worm-like chain; MD, molecular dynamics.

§To whom correspondence may be addressed. E-mail: bb8@columbia.edu or jfernandez@columbia.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0702179104/DC1.

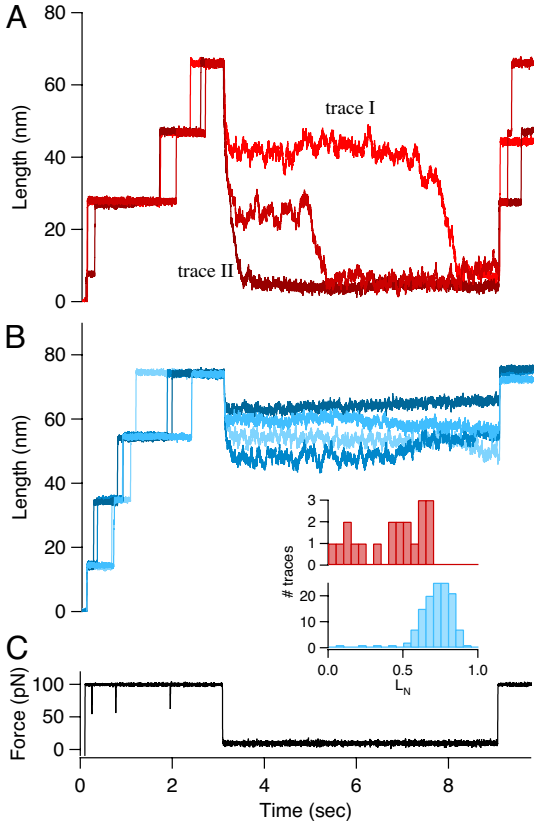

Fig. 1. Probing protein collapse in force-clamp spectroscopy. A single polyubiquitin protein is stretched at an initial high force of 100 pN, causing unfolding of individual modules marked by a step increase in length of 19.2 ± 0.5 nm. Subsequently, the force is quenched to 10 pN to monitor the collapse of the extended chain. To probe refolding, the protein is stretched again at 100 pN. (*A*) If ubiquitin modules have folded, the protein's length will decrease by a multiple of ≈20 nm as compared with its length before the force quench. In addition, a second series of unfolding events will be observed. (*B*) Most proteins failed to fold and, upon restoring the force to 100 pN, their length returned to the same value as before the quench. (*C*) The same force protocol was used for all force-quench experiments. (*Inset*) The magnitude of the initial collapse was very different for each polyprotein. Proteins that folded collapsed to 0.42 ± 0.22 of their length at 100 pN (*Upper*), whereas those that failed to fold only showed a reduction in length to 0.72 ± 0.12 (*Lower*).

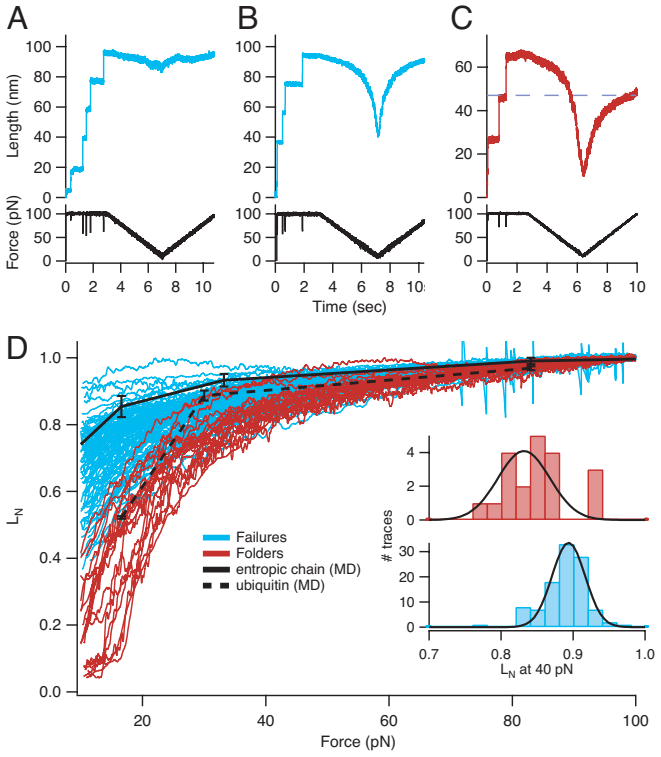

Fig. 2. To examine the nature of the forces driving protein collapse, we linearly decreased the force instead of quenching. In this modified force protocol, the protein was again unfolded at a high force of 100 pN. Subsequently, the force was ramped from 100 pN down to 10 pN in 4 sec and back up to 100 pN to probe refolding. In some cases, while the force was being relaxed, the protein collapsed very little (*A*), whereas in others, the same reduction in force caused a large contraction of the extended protein (*B*). (*C*) Protein folding was indicated by a reduction in length of 19.2 ± 0.5 nm upon restoring the force to 100 pN (see dashed line). (*D*) To compare all recordings, the length during the ramp was normalized by its value for the extended conformation at 100 pN. This normalized length, L_N, is shown as a function of force during the ramp down to 10 pN (folders in red, failures in blue). The force-length relationship of a purely entropic chain obtained with MD simulations is shown as the black curve and agrees well with those proteins in the experiment that failed to collapse significantly. If enthalpic interactions are included in the MD simulation, the protein contracts to an even shorter length (dashed line). (*Inset*) Histograms of L_N at 40 pN for folders (*Upper*) and failures (*Lower*). At this force, the distinct separation between folders (L_N = 0.85 ± 0.42) and failures (L_N = 0.89 ± 0.03) cannot be solely attributed to the reduction in length caused by folding.

Fig. 2 shows the results of applying this force-ramp protocol to 126 different, fully extended polyubiquitin molecules. We again observe a surprising degree of heterogeneity in the responses. In some cases, the protein collapsed very little during the ramp of the force down to 10 pN (Fig. 2*A*). In others, a large contraction of the extended protein was observed (Fig. 2*B*). Proteins were sometimes even observed to fold during the force-ramp protocol (Fig. 2*C*). These folding events were detected as a decrease in length by multiples of ≈20 nm after restoration of the force to 100 pN (Fig. 2*C*, dotted line). To compare all these recordings, we normalized their length by the value measured in the initial extended conformation at 100 pN. The normalized length, L_N, is shown in Fig. 2*D* as a function of the force during the ramp down to 10 pN. Proteins that failed to fold during the ramp (blue traces, n = 105) showed large variations in their collapse ranging from L_N = 0.4–0.95 at 10 pN. It is striking to observe the lack of recoil in many of these extended proteins. By contrast, proteins that folded (red traces, n = 21) collapsed much further than the failures, down to L_N = 0.05–0.4 at 10 pN. This distinct separation between failures and folders cannot be solely attributed to the reduction in length caused by folding, because it sets in already before folding, at high pulling forces. This is apparent from the histogram of L_N measured at 40 pN in Fig. 2*D Inset*. The bar centered at L_N = 0.93 in the top histogram constitutes proteins that remained very elongated down to low forces ≈40 pN but then suddenly collapsed and folded [supporting information (SI) Fig. 6]. The heterogeneity in the collapse behavior shown in Fig. 2*D* can also be observed in a single extended protein that was relaxed in two consecutive ramp cycles (SI Fig. 7). During the first ramp down, the protein collapsed to L_N = 0.2 at 10 pN, but in the second ramp, only to L_N = 0.8. These results confirm that the heterogeneity in Fig. 2*D* is a general property of protein collapse from an extended conformation.

The heterogeneity in the response of an extended protein to a relaxing force indicates that other factors in addition to entropy must contribute to the collapse. To obtain a quantitative estimate of the contribution made by entropy to the collapse of an extended ubiquitin protein, we used molecular dynamics (MD) simulation techniques. In these simulations, all noncovalent

BIOPHYSICS

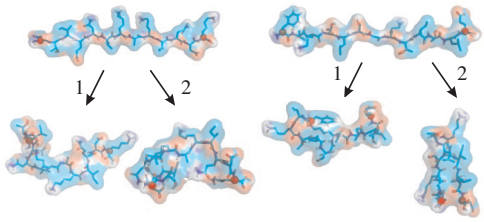

Fig. 3. The heterogeneity in protein collapse is probed by MD simulations for ubiquitin fragments. Two fragments of ubiquitin, residues 21–33 of the α-helix (*Left*, helix) and residues 4–16 of the N-terminal β-hairpin (*Right*, turn) were stretched at 500 pN (*Upper*). Force quench to 16.6 pN resulted in an ensemble of collapsed conformations with different lengths, of which selected snapshots are shown (*Lower*).

interactions were turned off, and only entropy remained as the driving force (for details, see *Materials and Methods*). A similar approach has been used to sample the conformational space of unfolded polypeptides (21, 22). Our simulations start by unfolding and extending a ubiquitin protein at a constant force of 500 pN. Then, all noncovalent interactions, except steric repulsion, were turned off, and the pulling force was relaxed to a constant value. The protein was equilibrated until its length no longer changed, except for fluctuations due to thermal motion. By repeating this procedure for several quench forces in the range of 0–167 pN, we obtained the force-length relationship for ubiquitin as a purely entropic chain (Fig. 2*D*, black solid line)[¶]. Comparing the simulation with the experimental data shows that those proteins that failed to collapse significantly are well described by a purely entropic chain. These results suggest that additional driving forces must have played a major role for proteins that were observed to collapse to even shorter lengths. Indeed, when the full force field is restored, the ubiquitin protein collapses to a much greater extent in the 20- to 30-ns time scale of the simulation (Fig. 2*D*, dashed line). Finer details of the collapse over longer time scales can be obtained by simplifying these simulations by using short segments of ubiquitin (Fig. 3 and SI Fig. 8). Repeated MD simulations with these short fragments showed that, whereas the entropic collapse was always similar, the extent of the collapse varied greatly in the presence of the full field (SI Fig. 8). This heterogeneity resulted from large differences in the backbone conformation and side-chain packing of the collapsing structures and hence in the hydrophobic and electrostatic interactions formed (Fig. 3). Fully extending a protein at a high force drives all dihedral angles toward $\pm$ 180° (SI Fig. 9). Thus, in an extended nearly linear conformation, all native contacts are lost, and the side chains are forced into the solvent. They are separated from each other by distances ranging from $\approx$3 Å for neighboring residues, up to the fully extended length of the protein (Fig. 3 and SI Figs. 10 and 11). We compared the magnitude of attractive forces acting in the extended protein on these length scales and find hydrophobic forces, as estimated from the potential of mean force between hydrophobic solutes in water (23, 24), to be >10-fold larger than corresponding electrostatic interactions (SI Fig. 11). Hence, it is likely that the main driving force for the collapse of an extended protein is hydrophobic.

Previous studies have shown that the range and strength of hydrophobic interactions can be influenced by solvent polarity (25–27), although the molecular origin of these observations remains unclear (28). To examine the role of hydrophobic

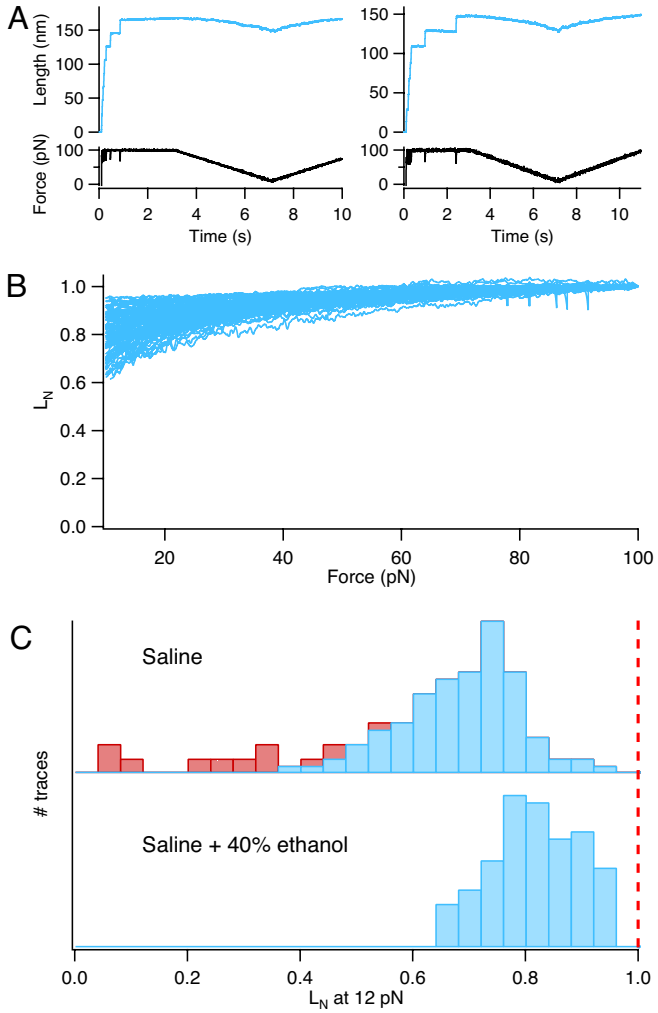

Fig. 4. Force-ramp experiments with ubiquitin in 40% vol/vol ethanol. (*A*) Two representative examples of the observed collapse of extended ubiquitin polyproteins in ethanol. In this experiment, all proteins failed to fold on the time scale of the experiment. (*B*) L_N as a function of the force during the ramp down to 10 pN (compare with Fig. 2*D*). Proteins remained more elongated in the presence of ethanol. (*C*) Histograms of L_N at 12 pN in the absence (*Upper*, $L_N = 0.62 \pm 0.19$; 105 failures, blue bars and 21 folders, red bars) and presence (*Lower*, $L_N = 0.82 \pm 0.08$; $n = 109$) of ethanol. The observed shift and narrowing of the distribution in the presence of ethanol are very significant, approaching the maximum possible value of $L_N = 1$ (red dashed line).

interactions in protein collapse, we repeated our force-ramp experiments with ubiquitin in a solution containing 40% vol/vol ethanol (Fig. 4*A*). This solution reduces the strength of the hydrophobic interactions (26, 27). Fig. 4*A* shows two typical traces obtained in these experiments. Although the unfolding rate and extension of a polyubiquitin protein changed only slightly in the presence of ethanol (SI Fig. 12), it is immediately apparent that the extent of the collapse was greatly reduced. Furthermore, no proteins were observed to fold in the time scale of the force-ramp experiments. However, force-quench experiments, where folding is more likely to occur than in the force-ramp experiments because of the constant low quench force, showed that folding was still possible in the ethanol solution (see SI Fig. 13). Fig. 4*B* shows the normalized length measured from 109 different molecules relaxed in the ethanol solution. Although we still observe a large degree of heterogeneity in the collapse, the range is visibly narrower than that observed in the standard saline, and the proteins remained more elongated

[¶]Enthalpic contributions to the collapse from covalent interactions, such as the dihedral potential, were found to be insignificant in this force range. Solvent effects on the chain entropy due to volume exclusion can also be expected to be negligible.

(compare with Fig. 2*D*). The histograms of L_N at 12 pN shown in Fig. 4*C* show a clear shift of the distribution in the presence of ethanol. The top histogram shows that in standard saline the extent of the collapse was broadly distributed, with $L_N = 0.62 \pm 0.19$ (105 failures, blue bars; and 21 folders, red bars). The lower histogram shows that in the ethanol solution the proteins remained significantly more extended with $L_N = 0.82 \pm 0.08$ ($n = 109$, only failures were observed, blue bars). The observed shift and narrowing of the distribution are very significant, because it approaches the maximum possible value of $L_N = 1$. These experiments strongly support the view that hydrophobic forces play a major role in the collapse of an extended protein.

Our experiments show that extended proteins have much lower conformational entropy than is generally assumed in the conventional random-coil picture of a denatured protein. A fully extended ubiquitin can do a large amount of work, ≈170 kT during collapse at an equilibrium force of 35 pN (20). From our MD simulations, we can estimate the contribution made by chain entropy. The simulated entropic collapse shown in Fig. 2*D* (solid line) is well described by the WLC with a persistence length of 1.2 nm. By integrating this WLC between 100 and 35 pN, we find that chain entropy contributes only ≈24 kT to the overall change in free energy during the quench. Hence, the remainder of the work, ≈150 kT, must be mainly due to hydrophobic interactions. In our force-clamp experiments, a protein is unfolded and extended, forcing the unraveling of all hydrophobic side chains and fully exposing them to the solvent. This process must then be reversed on decreasing the force. Given a hydrophobic solvation free energy of $\Delta G_h = 150$ kT and a change in hydrophobic surface area of $\Delta A \sim 33$ nm^2 during collapse, we obtain a surface tension of $\gamma \sim 30$ cal/mol/Å^2 (see *SI Text*), which is in good agreement with the values commonly used to estimate the hydrophobic effect in protein folding (29, 30).

Force-quench experiments such as the ones shown in Fig. 1 showed a puzzling departure from the conventional views of protein folding. The complex and heterogeneous collapse observed after a force quench cannot be explained by a purely entropic mechanism nor by conventional two-state protein-folding theories. Clearly, a better understanding of the forces driving the collapse of a highly extended protein was needed to explain these phenomena. Our experiments confirm that hydrophobic interactions are the principal driving force for protein folding (11) even from highly extended conformations. Our simulations showed that the magnitude of the hydrophobic collapse varied greatly depending on the region of backbone and side chain dihedral space sampled by the collapsing protein. Thus, the extent of the collapse observed should vary depending on the initial conditions and the path followed during any particular trajectory. These considerations readily explain both the variability observed during the force quench experiments of Fig. 1, as well as the wide distribution of normalized length values measured in the force-ramp experiment of Figs. 2 and 3. Furthermore, they also explain why the extended polyprotein does not have any memory of the individual ubiquitin proteins in the chain (13).

A collapse driven mostly by hydrophobic forces should be a very general property of an extended protein and should not be very dependent on amino acid sequence. To test this hypothesis, we have completed a series of force-quench and -ramp experiments with another three very different proteins: the I27 Ig module of human cardiac titin (31), a PEVK protein from human titin (31), and protein L (32). In all three cases, the collapse measured from an extended state showed a similar range and heterogeneity (SI Fig. 14), as was observed for ubiquitin (Fig. 2*D*). In these experiments, we also observed folding events in I27 and protein L. The elastomeric PEVK protein does not fold. These experiments confirm the view that the collapse behavior observed in ubiquitin is a general property of extended proteins. This common collapse behavior suggests an intrinsic degree of hydrophobicity as an evolutionary constraint for soluble proteins, as inferred previously from protein sequence analysis (9–11, 33). By comparison the collapse and folding of RNA hairpins is dominated by entropic barriers arising from the conformational state of particular regions in the hairpin (34, 35).

Given that most of the collapse of an extended protein is driven by hydrophobic forces, the use of the WLC model of polymer elasticity can only be considered phenomenological. From our experiments, we have found that unfolded proteins can be far stiffer than assumed before. For example, the trace shown in Fig. 2*A*, as well as the traces where $L_N \sim 0.9$, can be fit by a WLC with a persistence length of $P > 1$ nm. These traces most likely represent an extended state where the hydrophobic forces are diminished because of a particular combination of dihedrals, and only entropic collapse is possible. A persistence length of $P > 1$ nm is in agreement with the persistence length of approximately three residues determined on the basis of conformational flexibility of the protein backbone as deduced from NMR relaxation data (36). High-persistence-length values were also recently observed in chemically denatured proteins using FRET techniques (37). Such high-persistence lengths have also been observed in PEVK proteins (38–40) and considered to be an indication that, in addition to entropy, intramolecular interactions also play a role in the force-extension behavior of this natively disordered muscle protein (41, 42). Thus, it is now clear that the apparent persistence length of ≈0.4 nm, resulting from WLC fits to protein unfolding data from atomic force microscopy experiments (43), reflected a phenomenological stiffness, comprising effects due to both chain entropy and hydrophobic collapse. A new model of protein elasticity needs to be developed that properly accounts for both the entropic and the much larger hydrophobic collapse mechanisms. This development is also necessary to rederive improved scaling laws for the various stages of protein collapse (18). However, these efforts are beyond the scope of this work.

The force-spectroscopy experiments described here allowed us to probe regions of the folding free-energy landscape hitherto unvisited by other chemical, thermal, or low force mechanical denaturation experiments. From our experiments, we can now estimate the changes in free energy for a collapsing protein. Fig. 5*A* (upper curve) shows the calculated contribution made by entropic work for a protein collapsing at 10 pN. The free energy change shows a minimum at $L_N \sim 0.8$, in agreement with the extent of collapse of a purely entropic chain (Fig. 2*D*). The magnitude and position of this minimum solely depend on the pulling force during collapse. Although entropy makes a contribution, we have found that proteins in highly extended states are driven to collapse mostly by hydrophobic forces. Fig. 5*A* (lower curves) shows a hypothetical representation of the potential of mean force of hydrophobic collapse in an extended polypeptide chain. Such a potential results from a convolution of the pairwise distances between hydrophobic side chains and the potential of mean force measured for two hydrophobic solutes, such as those depicted in SI Fig. 10. The pairwise distances are random variables, constrained by the dihedral space available at any given end-to-end length of the protein. In the extreme case of a protein stretched to its contour length, L_C, only one dihedral conformation is possible (± 180°; see SI Fig. 9). However, shorter lengths can accommodate a rapidly increasing number of dihedral conformations, all resulting in the exact same end-to-end distances but with very different hydrophobic driving forces. Thus, for a collapsing protein, the magnitude of the hydrophobic driving force is determined by its trajectory through dihedral space (Fig. 5*A*, lower curves). To represent these considerations graphically, we construct a simplified potential energy landscape for the collapse of a highly extended protein (Fig. 5 *B* and *C*). The radial distance from the native state describes the protein's end-to-end length, and the angle parameter represents an arbitrary dihedral

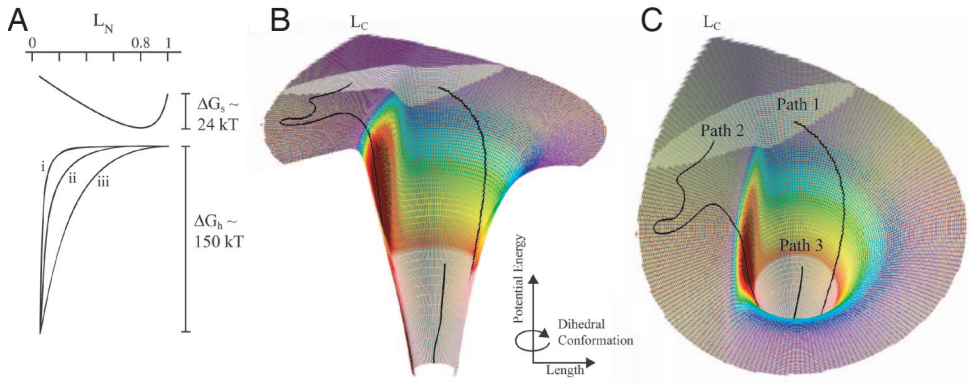

Fig. 5. Energy landscape for the collapse of extended proteins. (*A*) Change in free energy due to entropy, ΔG_s (*Upper*) and hydrophobic forces, ΔG_h (*Lower*) for the collapse of an extended protein. ΔG_s is shown for a collapsing protein under a stretching force of 10 pN. The potential energy for the hydrophobic collapse of an extended polypeptide chain results from a convolution of the pairwise distances between hydrophobic side chains and the potential of mean force measured for two hydrophobic solutes such as those depicted in SI Fig. 10. Three examples for ΔG_h are shown that illustrate varying hydrophobic strengths that arise from different combinations of pairwise distances and thus conformations in dihedral space during the collapse. (*B* and *C*) In our landscape, the radial distance describes the protein's end-to-end length, the angle represents an arbitrary dihedral space coordinate, and the potential energy is shown on the *z* axis. As proteins are stretched to higher forces, the dihedral space becomes severely limited as the protein is stretched close to its contour length, as indicated by the asymmetry in the outer edge of the landscape. At a given unfolding force, the unfolded protein ensemble comprises substates on the energy landscape with different lengths and allowed dihedral space (white areas). Upon quenching the force, highly extended proteins begin to collapse from a well defined end-to-end length, albeit from very different starting conformations in dihedral space (see path 1 vs. path 2).

space coordinate. The *z* axis represents a hypothetical potential of mean force for hydrophobic collapse, which decreases as the protein moves toward its native state. Given that the observed collapse is a property of extended polypeptides, regardless of whether they fold (see above), we exclude the bottom of the funnel, where two-state folding is typically observed. For clarity, we idealize our energy landscape to be very smooth, although it is clear that any such landscape will have a degree of roughness (6). The funnel depicted in Fig. 5 *B* and *C* readily explains the diversity in the observed collapse after a force quench (Fig. 1). Extension at a high force drives a protein toward L_C (Fig. 5 *B* and *C*). Upon quenching the force, the protein begins to collapse from a well defined end-to-end length, albeit from very different starting conformations in dihedral space (see path 1 vs. path 2). Thereafter, the extent and temporal evolution of the collapse will be largely determined by the strength of the hydrophobic interactions. For example, in path 2, the protein meanders on a relatively flat energy landscape and eventually finds the steep parts of the funnel leading toward collapse. Thus, proteins following path 2 will be observed to collapse over long time scales (trace I, Fig. 1*A*) or fail in the time scale of the experiment (Fig. 1*B*). By contrast, path 1 begins already on a steep slope driving a near two-state collapse behavior (trace II, Fig. 1*A*). Proteins that begin their collapse from much smaller extensions (path 3, Fig. 5 *B* and *C*) sense only the steep parts of the funnel, where hydrophobic collapse is strongest, and do not have access to the flat regions of the energy landscape. Under these conditions, the protein exhibits two-state behavior, as demonstrated in optical tweezers experiments with RNase H (44). The same holds for thermal or chemical denaturation experiments (1–3), where the end-to-end length in the denatured state is only slightly larger than that of the folded protein. Thus, the results presented here now explain the apparent discrepancy in the folding behavior of proteins whose denatured conformations differ greatly in length.

Materials and Methods

Protein Engineering. Polyproteins of ubiquitin, I27 and the PEVK-I27 chimera were cloned and expressed as described elsewhere (20, 39, 45). The PEVK-I27 construct is based on the human titin I27 and PEVK (exon 161) sequences. The core sequence for the chimera consists of three tandem repeats of the PEVK exon followed by an I27 module. This core sequence is repeated four times in the final polyprotein chimera. The plasmid containing the B1 Ig-binding domain of peptostreptococcal protein L (32) was a generous gift from David Baker (University of Washington, Seattle, WA). The polyprotein $(\text{protein L})_8$ was constructed by using published methods (46).

Single-Molecule Force-Clamp Spectroscopy. We used a custom-made atomic force microscope under force-clamp conditions (47). Each cantilever (Si_3N_4 from Veeco, Santa Barbara, CA) was calibrated in solution by applying the equipartition theorem (48). The spring constant was typically found to be ≈14 pN nm^{-1}. Experiments were carried out either in Hepes or PBS buffer, both at pH 7. Ethanol was obtained from Sigma-Aldrich (St. Louis, MO). Aqueous solutions of 40% vol/vol ethanol were carefully prepared to ensure the same salt concentration and pH.

Single chains of proteins were picked up by pushing the cantilever onto the surface at 500 pN for 1–3 sec. In our experiments, we first stretched a single polyprotein at a constant force, resulting in a staircase such as shown in Figs. 1 and 2, where each step corresponds to the unfolding of a single module in the chain. Because molecules were picked up at random positions, the number of steps varied between 1 and 8 (for polyprotein L and polyI27) or 9 (for polyubiquitin). After a few seconds, the force was either quenched or linearly decreased. Varying the duration of the ramp between 1 and 4 sec (tested on ubiquitin) had no effect on our results. Finally, the force was again increased to observe whether the protein modules had refolded.

Data Analysis. All length recordings were corrected for the cantilever displacement. Failures were identified as curves that did not contain any steps during the force ramp and that displayed the same end-to-end length before and after the decrease in force. Curves were classified as successes if, after the decrease in force, the length had decreased by a multiple of ≈20 (ubiquitin), 24 (I27) or 16 nm (protein L).

MD Simulations. All simulations were carried out by using the MD software package GROMACS 3.1.4 (49). Simulation details for the equilibration of ubiquitin [Protein Data Bank ID code 1UBQ (50)] have been described (51). The GROMOS96 force field (52) for the

protein and the single-point-charge water model (53) were used. Starting from the equilibrated system, ubiquitin termini were subjected to a constant force of 500 pN, resulting in unfolding and stretching of the protein. In a series of subsequent force-clamp MD simulations of stretched ubiquitin and shorter protein fragments thereof, the force was quenched to 16.6–83 pN. Theses force-clamp simulations were repeated for ubiquitin as an entropic chain, i.e., in the absence of nonbonded interactions and hydrophobic forces, by using the GROMOS96 and the OPLS/AA force field for comparison (54). All simulations details are given in the *SI Text*.

We thank Ronan Zangi, Jasna Brujic, and Helmut Grubmüller for fruitful discussions. This work was supported by National Institutes of Health HL66030 and HL61228 (to J.M.F.) and National Science Foundation Grant CHE-03-16896 (to B.J.B.). F.G. is supported by a Columbia University grant to J.M.F. and B.J.B.

1. Khorasanizadeh S, Peters ID, Roder H (1996) *Nat Struct Biol* 3:193–205.
2. Gladwin ST, Evans PA (1996) *Folding Des* 1:407–417.
3. Krantz BA, Sosnick TR (2000) *Biochemistry* 39:11696–11701.
4. Frauenfelder H, Sligar SG, Wolynes PG (1991) *Science* 254:1598–1603.
5. Leopold PE, Montal M, Onuchic JN (1992) *Proc Natl Acad Sci USA* 89:8721–8725.
6. Bryngelson JD, Onuchic JN, Socci ND, Wolynes PG (1995) *Proteins Struct Funct Genet* 21:167–195.
7. Shortle D, Ackerman MS (2001) *Science* 293:487–489.
8. Alexandrescu AT, Abeygunawardana C, Shortle D (1994) *Biochemistry* 33:1063–1072.
9. Kauzmann W (1959) *Adv Protein Chem* 14:1–63.
10. Tanford C (1980) *The Hydrophobic Effect: Formation of Micelles and Biological Membranes* (Wiley, New York).
11. Dill KA (1990) *Biochemistry* 29:7133–7155.
12. Li MS, Hu CK, Klimov DK, Thirumalai D (2006) *Proc Natl Acad Sci USA* 103:93–98.
13. Best RB, Hummer G (2005) *Science* 308.
14. Cieplak M, Szymczak P (2006) *J Chem Phys* 124.
15. Bustamante C, Marko JF, Siggia ED, Smith S (1994) *Science* 265:1599–1600.
16. Livadaru L, Netz RR, Kreuzer HJ (2003) *Macromolecules* 36:3732–3744.
17. Rief M, Gautel M, Oesterhelt F, Fernandez JM, Gaub HE (1997) *Science* 276:1109–1112.
18. Thirumalai D (1995) *J Physique I* 5:1457–1467.
19. Buguin A, BrochardWyart F, deGennes PG (1996) *C R Acad Sci Ser Fasc B* 322:741–746.
20. Fernandez JM, Li H (2004) *Science* 303:1674–1678.
21. Pappu RV, Srinivasan R, Rose GD (2000) *Proc Natl Acad Sci USA* 97:12565–12570.
22. Tran HT, Pappu RV (2006) *Biophys J* 91:1868–1886.
23. Ghosh T, Garcia AE, Garde S (2001) *J Am Chem Soc* 123:10997–11003.
24. Pangali C, Rao M, Berne BJ (1979) *J Chem Phys* 71:2975–2981.
25. Parker JL, Claesson PM (1992) *Langmuir* 8:757–759.
26. Kokkoli E, Zukoski CF (1999) *J Colloid Interface Sci* 209:60–65.
27. Tanford C (1962) *J Am Chem Soc* 84:4240.
28. Gee ML, Israelachvili JN (1990) *J Chem Soc Faraday Trans* 86:4049–4058.
29. Nicholls A, Sharp KA, Honig B (1991) *Proteins Struct Funct Genet* 11:281–296.
30. Sharp KA, Nicholls A, Fine RF, Honig B (1991) *Science* 252:106–109.
31. Labeit S, Kolmerer B (1995) *Science* 270:293–296.
32. Kim DE, Fisher C, Baker D (2000) *J Mol Biol* 298:971–984.
33. Schwartz R, King J (2006) *Protein Sci* 15:102–112.
34. Liphardt J, Onoa B, Smith SB, Tinoco I, Bustamante C (2001) *Science* 292:733–737.
35. Hyeon CB, Thirumalai D (2006) *Biophys J* 90:3410–3427.
36. Danielsson J, Andersson A, Jarvet J, Graslund A (2006) *Magn Reson Chem* 44:S114–S121.
37. Hoffmann A, Kane A, Nettels D, Hertzog DE, Baumgartel P, Lengefeld J, Reichardt G, Horsley DA, Seckler R, Bakajin O, *et al*. (2007) *Proc Natl Acad Sci USA* 104:105–110.
38. Li HB, Oberhauser AF, Redick SD, Carrion-Vazquez M, Erickson HP, Fernandez JM (2001) *Proc Natl Acad Sci USA* 98:10682–10686.
39. Sarkar A, Caamano S, Fernandez JM (2005) *J Biol Chem* 280:6261–6264.
40. Watanabe K, Nair P, Labeit D, Kellermayer MSZ, Greaser M, Labeit S, Granzier H (2002) *J Biol Chem* 277:11549–11558.
41. Linke WA, Ivemeyer M, Mundel P, Stockmeier MR, Kolmerer B (1998) *Proc Natl Acad Sci USA* 95:8052–8057.
42. Forbes JG, Jin AJ, Ma K, Gutierrez-Cruz G, Tsai WL, Wang KA (2005) *J Muscle Res Cell Motil* 26:291–301.
43. Arrondo JLR, Alonso A (2006) in *Springer Series in Biophysics* (Springer, Berlin, Germany), Vol 10, p 280.
44. Cecconi C, Shank EA, Bustamante C, Marqusee S (2005) *Science* 309:2057–2060.
45. Carrion-Vazquez M, Li HB, Lu H, Marszalek PE, Oberhauser AF, Fernandez JM (2003) *Nat Struct Biol* 10:738–743.
46. Carrion-Vazquez M, Oberhauser AF, Fowler SB, Marszalek PE, Broedel SE, Clarke J, Fernandez JM (1999) *Proc Natl Acad Sci USA* 96:3694–3699.
47. Schlierf M, Li H, Fernandez JM (2004) *Proc Natl Acad Sci USA* 101:7299–7304.
48. Florin EL, Rief M, Lehmann H, Ludwig M, Dornmair C, Moy VT, Gaub HE (1995) *Biosens Bioelectron* 10:895–901.
49. Lindahl E, Hess B, van der Spoel D (2001) *J Mol Model* 7:306–317.
50. Vijaykumar S, Bugg CE, Cook WJ (1987) *J Mol Biol* 194:531–544.
51. Gräter F, Grubmuller H (2006) *J Struct Biol* 157:567–569.
52. van Gunsteren WF, Billeter SR, Eising AA, Hünenberger PH, Krüger P, Mark AE, Scott WRP, Tironi IG (1996) *Biomolecular Simulation: The GROMOS96 Manual and User Guide* (Hochschulverlag, Zurich).
53. Berendsen HJC, Postma JPM, van Gunsteren WF, Hermans J (1981) in *Intermolecular Forces* (Reidel, Dordrecht, The Netherlands), pp 331–342.
54. Jorgensen WL, Swenson CJ (1985) *J Am Chem Soc* 107:569–578.