

# Osmolyte-induced separation of the mechanical folding phases of ubiquitin

Sergi Garcia-Manyes<sup>1</sup>, Lorna Dougan<sup>2</sup>, and Julio M. Fernández<sup>1</sup>

Department of Biological Sciences, Columbia University, New York, NY 10027

Edited by Alan Fersht, University of Cambridge, Cambridge, United Kingdom, and approved May 8, 2009 (received for review February 25, 2009)

**Solvent molecules play key roles in the conformational dynamics of proteins. Here we use single molecule force-clamp spectroscopy to probe the role played by the stabilizing osmolyte glycerol on the conformational ensembles visited by a single ubiquitin protein folding after mechanical extension. Using a variety of force-pulse protocols, we find that glycerol stabilizes the native state of ubiquitin, making it more resistant to mechanical unfolding. We also find that although glycerol enhanced the hydrophobic collapse of unfolded and highly extended ubiquitins, it had no effect on the resulting collapsed states that are essential precursors of the folded state. These disparate effects of glycerol may be the result of distinct structural roles played by solvent molecules at the transition state of each folding ensemble. Our results open the way for a detailed analysis of the transition state structures that form along the folding trajectory of a mechanically extended protein.**

force-clamp | glycerol | single molecule | solvent

Understanding the mechanisms by which proteins acquire their native topology remains an important question in biology. The acquisition of the native state involves a delicate interplay between the structural search for the native state and desolvation of the solvent exposed protein backbone. Water molecules are inextricably involved in this dynamic conversion between the different protein conformations encountered along the folding pathway (1). Furthermore, water molecules mediate the hydrogen bonding network that determines the final structure of the native state of the protein and the dynamics of contraction of the unfolded polypeptide (2, 3). A comprehensive description of the (un)folding reaction must therefore incorporate the solvating environment that envelops the protein and determines its structure, dynamics, and often function (1, 4). Studies using bulk experimental techniques have provided a wealth of information on the thermodynamic properties of proteins upon modification of the solvent environment (5–8) and molecular mechanisms for osmolyte induced protein stability have been proposed (9). However, direct insight into the role of osmolyte molecules on the structure of the transition state of a protein, which is the main determinant of protein dynamics, can only be gained at the single molecule level.

Single molecule force spectroscopy is a valuable tool to uncover the structure of the transition state of a protein under the application of a stretching force along a well-defined coordinate, the end-to-end length (10, 11). While earlier pioneering experiments allowed location of the mechanical transition state of several proteins (12–14), molecular dynamics simulations complemented those experimental findings by providing atomic-level insight into the key events that determine protein unfolding, namely the dynamic rupture of interstrand hydrogen bonds (15). Most recently, the combination of solvent substitution with force-clamp spectroscopy has proved successful in experimentally providing a molecular perspective on the role of solvent molecules in protein unfolding (16, 17) and in understanding the driving forces involved in the collapse of an extended polypeptide (18). Specifically, the use of solvent substitution provides a useful tool to uncover the molecular architecture of the mechanical transition state of proteins in detail. These experiments demonstrated that solvent bridging plays a pivotal role in the mechanical stability of the native state of the I27

protein under force (16). However, despite this significant progress, these studies have been mostly limited to the study of thermodynamically stable states, namely the unfolded and the folded structures (16, 18), and little is known about the transition state structures of the evolving conformations acquired by the polypeptide as it folds.

A significant advantage of single molecule techniques is their ability to individually monitor the full folding trajectory of a single molecule (19, 20). In the case of mechanical experiments, applying a stretching force to a protein slows down the folding mechanism exponentially, which permits capturing the conformational dynamics followed by a single protein from a totally extended conformation to its natively refolded form (21–23). Recent measurements conducted in constant force mode have allowed us to separate the different stages encompassing the folding reaction of the small protein ubiquitin (22, 24). These experiments permitted identification of 3 distinct conformational ensembles visited during the folding process by probing the mechanical stability of each protein conformation as its structural fingerprint. This ensemble of conformations comprise the folded and unfolded protein forms, along with a recently discovered ensemble of minimum energy collapsed structures (MECS) that are obligate precursors to the native form (25). Such structurally heterogeneous and mechanically labile MECS originate after the protein's hydrophobic collapse, and interconvert into the native state following a 2-state, rate limiting reaction (26–28). Interestingly, this set of mechanically labile structures has been also identified in analogous experiments using ubiquitin monomers, suggesting that they are not a particular feature of polyproteins (25).

The capability of force-clamp spectroscopy to drive a single protein to different regions of the energy landscape provides the unprecedented opportunity to probe the effect of solvents on each distinct conformational ensemble visited by a single ubiquitin protein along its folding pathway (18, 22, 25). Using a variety of force-pulse protocols, here we show that glycerol, a protecting osmolyte (7), readily separates the distinct phases of the mechanical folding of ubiquitin by stabilizing the native state. In contrast to the glycerol sensitivity of the extended and native state of ubiquitin, we measure that the mechanical resistance of an ensemble of collapsed states is largely insensitive to the solvent environment. Our experiments reveal the role of glycerol in the different phases of the folding reaction while providing plausible molecular interpretations of the mechanical transition state of each particular protein conformation.

## Results

**Broad Separation of Folding Phases in Glycerol Solutions.** To capture and fully separate the conformational dynamics sampled by a single

Author contributions: S.G.-M., L.D., and J.M.F. designed research; S.G.-M. and L.D. performed research; S.G.-M. and L.D. analyzed data; S.G.-M., L.D., and J.M.F. wrote the paper.

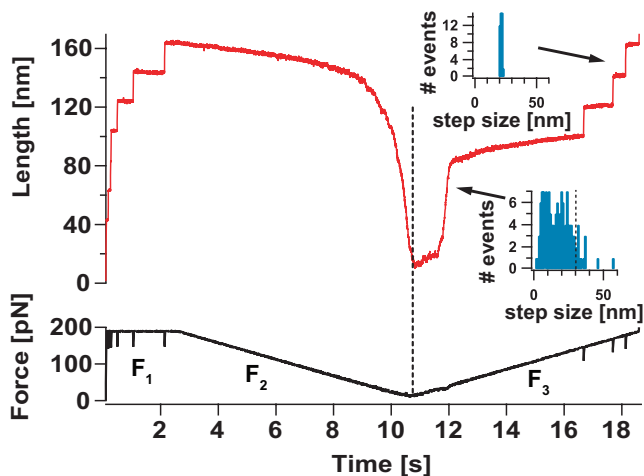
The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

<sup>1</sup>To whom correspondence may be addressed. E-mail: sergi@biology.columbia.edu or jfernandez@columbia.edu.

<sup>2</sup>Present address: School of Physics and Astronomy, University of Leeds, Leeds LS2 9J, United Kingdom.

This article contains supporting information online at [www.pnas.org/cgi/content/full/0902090106/DCSupplemental](http://www.pnas.org/cgi/content/full/0902090106/DCSupplemental).



**Fig. 1.** Wide separation of the folding ensembles in glycerol solutions. We use a force-ramp protocol in 30% glycerol solution to completely separate the 3 conformational ensembles of a folding polypeptide, demonstrating that they are structurally distinct. The ubiquitin polypeptide is first unfolded under a constant force of 190 pN ( $F_1$  regime) during 2.5 s, resulting in a well-defined series of step increases in length of 20 nm that mark the unfolding and extension of the individual modules in the chain. The force is then relaxed linearly down to 10 pN, while monitoring the contraction of the end-to-end length of the highly extended ubiquitin polypeptide as it undergoes hydrophobic collapse ( $F_2$ ). Upon reaching 10 pN, the force is linearly ramped back up again to 190 pN ( $F_3$ ). When the force value reaches approximately 50–60 pN, the protein undergoes a heterogeneous stepwise extension (*Bottom Inset*) marking the unraveling of the ensemble of collapsed conformations. At a much higher force of approximately 150 pN we observe the appearance of the 20 nm steps (*Top Inset*) that mark the native ensemble.

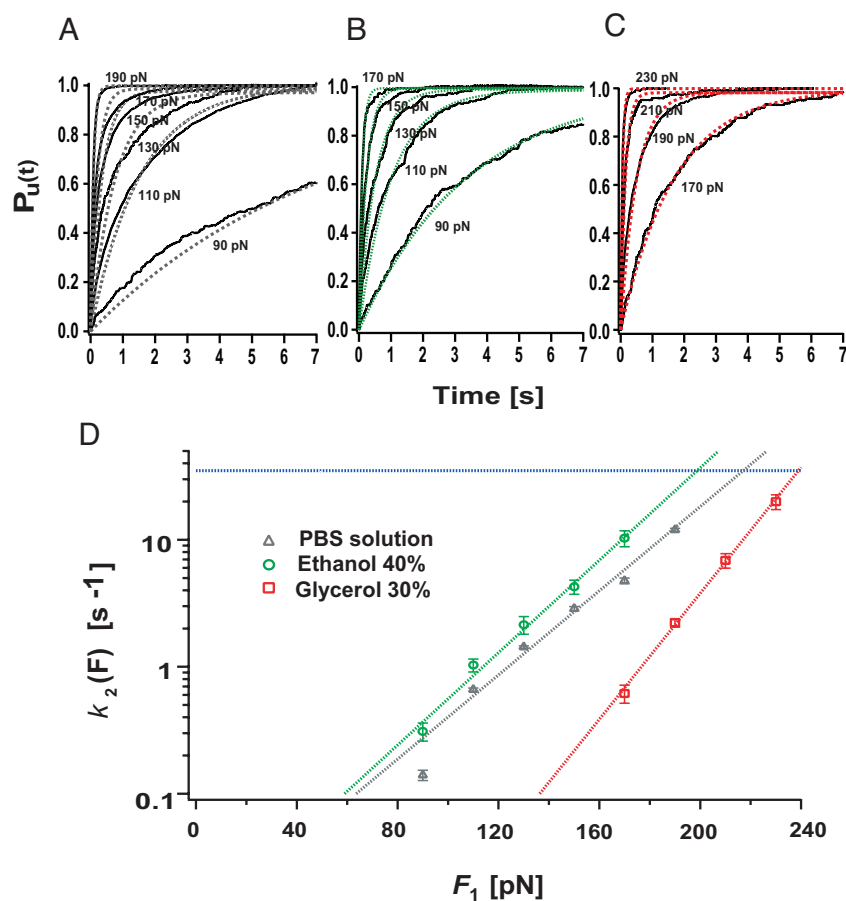
protein during its folding trajectory, we first apply a force-ramp protocol. This protocol permits the observation of the full force-length relationship of an extended polyprotein during protein collapse and folding. Fig. 1 shows a single ubiquitin polypeptide in 30% vol/vol aqueous glycerol being first unfolded under a constant high force of 190 pN to a highly extended conformation where all of the residues are exposed to the solvent. Under force-clamp conditions, stretching a ubiquitin polypeptide results in a well-defined series of step increases in length of approximately 20 nm, marking the unfolding and extension of the individual modules in the chain ( $F_1$  regime). The force is then relaxed linearly down to 10 pN, while monitoring the end-to-end length of the protein as it collapses ( $F_2$ ). Upon reaching 10 pN the force is ramped back up again to 190 pN ( $F_3$ ). As the force begins to rise, the collapsed polypeptide immediately shows resistance to extension. When the force reaches approximately 50–60 pN, we observe a rapid stepwise extension of the collapsed conformations. The step size distribution for the unraveling of such collapsed structures in glycerol is broad (*Bottom Inset*) and similar to that measured in PBS solution (25). Interestingly, the majority ( $\approx 90\%$ ) of the measured step sizes lies below 29 nm (dashed line in the histogram *Inset*), indicating that the distribution of step sizes corresponding to the collapsed conformations is largely contained within the contour length of a single ubiquitin monomer. This result suggests that just after hydrophobic collapse, the polyprotein chain rapidly segregates into its individual ubiquitin monomers. A small number of larger steps ( $\approx 10\%$ ) were also observed, because of either lack of sufficient time resolution or conformations that involved interactions between neighboring ubiquitin domains (29–31). Although multiple domains in a polyubiquitin chain are known to interact with each other forming ubiquitin dimers (31) and tetramers (29), such inherently flexible complexes are likely to be the result of lattice interactions rather than stable conformations (29). Thus, it is not likely that intra-chain binding plays a role in the collapsing dynamics that we observe. At

a much higher force, in excess of 150 pN, we begin to observe the unfolding of the stabilized native ubiquitin proteins that have had time to refold within the short time that the protein was held in its collapsed length, exhibiting a narrow distribution of step sizes centered at  $\approx 20$  nm (*Top Inset*). This experiment clearly demonstrates the 3 distinct phases of a folding polypeptide separated by the different effects of glycerol, namely the native folded protein, the extended protein and the collapsed protein. It is important to note that the 3 phases are observed in the same polypeptide molecule as it progresses toward the native state. This presents a unique opportunity to separate the timescales of the distinct folding ensembles within a single molecule experiment. Independently studying the effect of force and solvent substitution on the native, unfolded, and collapsed states of ubiquitin provides an ideal experimental platform to compare their distinct properties.

### Probing the Physicochemical Properties of the Native Conformations.

We begin by measuring the properties of the native conformations of the ubiquitin protein by measuring their unfolding force dependency under different solvent environments. In our experiments, we probe the mechanical properties of the native state of mechanically stable proteins by studying their unfolding kinetics under the effect of a constant stretching force (Fig. 2A) (10, 32). To gain insight into the molecular structure of the mechanical transition state of unfolding, we measure the force dependency of the unfolding rate of single polyproteins using the Bell model (10, 16, 24, 33), which allows us to estimate the distance to the transition state,  $\Delta x$ , and the height of the energy barrier,  $\Delta G$ , that define the unfolding reaction as described in the [supporting information \(SI\) Text](#). In combination with solvent substitution, these experiments have revealed the molecular architecture of the unfolding transition state of the protein I27, demonstrating that solvent bridging in the transition state region determines  $\Delta x$ . In particular, recent experiments showed that upon replacement of water by increasing amounts of the larger molecule glycerol, the distance to the unfolding transition state increases from  $\Delta x = 2.5$  Å, the size of a water molecule, up to  $\Delta x = 4.4$  Å, consistent with the size of a glycerol molecule. Furthermore, the replacement of water with glycerol greatly increases  $\Delta G$ , thereby mechanically stabilizing the native state (16). Whether this trend represents a general mechanism for the transition state architecture of other mechanically stable proteins such as ubiquitin remains experimentally unexplored.

Fitting the Bell equation to the data Fig. 2D (gray dashed line) corresponding to the force-dependent unfolding rate of ubiquitin in standard PBS buffered aqueous solution (Fig. 2C, gray triangles) resulted in  $k_{0,2} = 8.8 \times 10^{-3} \pm 5.0 \times 10^{-4} \text{ s}^{-1}$  and  $\Delta x_2 = 1.6 \pm 0.1$  Å, in close agreement with earlier experiments (10). From the measured value of  $k_{0,2}$  and assuming a prefactor  $A = 10^6 \text{ s}^{-1}$  (34) we readily estimated the value of  $\Delta G_2 = 11.0 \text{ kcal mol}^{-1}$  (25). To investigate the effect of solvent substitution on the mechanical stability of the native state of ubiquitin, we have now extended our measurements to study its unfolding force-dependency in solutions containing 40% vol/vol ethanol and 30% vol/vol glycerol. Fig. 2B and C show the averaged unfolding trajectories and their corresponding exponential fits obtained at different pulling forces in 40% vol/vol ethanol and 30% vol/vol glycerol, respectively. The unfolding force dependency obtained from these data are shown in Fig. 2D (40% vol/vol ethanol, green circles; 30% vol/vol glycerol, red squares). From the data, it is apparent that glycerol, a stabilizing osmolyte (9), makes the native state of ubiquitin significantly more resistant to mechanical unfolding. This effect is not solely because of viscosity, but rather reflects an increase in the activation barrier height (35). Arrhenius fits to the data (Fig. 2D, dashed lines) show that glycerol causes an increase in the activation energy barrier to unfold by  $\Delta\Delta G_2 \approx 3.2 \text{ kcal/mol}$ . By contrast, substitution with a 40% vol/vol ethanol solution has only a minor effect ( $\Delta\Delta G_2 \approx 0.05 \text{ kcal/mol}$ ). Furthermore, while the measured distance to the transition state increases when replacing PBS solution with 30% vol/vol



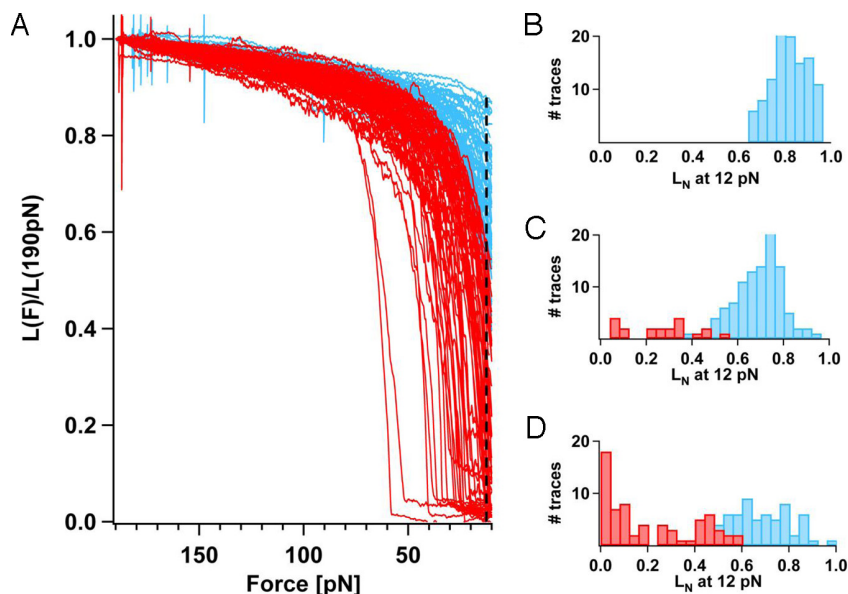
**Fig. 2.** The mechanical stability of the native state of ubiquitin is strongly dependent on the solvent environment. Applying a constant stretching force to a ubiquitin polyprotein results in length versus time trajectories showing the characteristic staircase of unfolding events, where each step of 20 nm marks the individual unfolding of a ubiquitin module in the chain. The average time-course of unfolding was obtained by summation and normalization of  $n > 50$  recordings. (A–C) Multiple trace averages of unfolding events for ubiquitin in (A) PBS aqueous solution, data from (25), (B) ethanol solution (40% vol/vol), and (C) glycerol solution (30% vol/vol) at different constant pulling forces. We obtain the average rate of unfolding of ubiquitin by fitting a single exponential to the average time-course of unfolding at each particular force. (D) Semilogarithmic plot of the rate of unfolding of ubiquitin,  $k_2$ , as a function of the pulling force in: PBS aqueous solution, gray triangles; 40% ethanol, green circles; 30% glycerol, red squares, obtained from the data shown in (A–C). The dashed line in each case represents the fit of the Arrhenius/Bell term,  $k_2(F) = k_{0,2} \exp(F\Delta x_2/kT)$ , to the experimental data. From this fit we obtain the values of the rate constant in the absence of force,  $k_{0,2}$ , and the distance to the transition state,  $\Delta x_2$ , for each different solvent environment. Assuming a prefactor  $A = 10^6 \text{ s}^{-1}$  we can estimate the height of the energy barrier of unfolding,  $\Delta G_2$ , by using the relationship  $k_{0,2} = A \exp(-\Delta G_2/kT)$ . The obtained values are shown in Table S1. These results show that both  $\Delta G_2$  and  $\Delta x_2$  are strongly dependent on the solvent environment. Blue horizontal lines in D represent the limit in the rate resolution for experiments conducted with cantilevers with a spring constant of 15 pN/nm.

glycerol from  $\Delta x_{2, \text{water}} = 1.6 \pm 0.1 \text{ \AA}$  to  $\Delta x_{2, \text{gly}} = 2.3 \pm 0.1 \text{ \AA}$ , the value remains almost unaltered upon the addition of 40% vol/vol ethanol,  $\Delta x_{2, \text{ethanol}} = 1.7 \pm 0.1 \text{ \AA}$ . The values of the parameters defining the energy landscape under all tested chemical environments, namely the height of the activation energy barrier,  $\Delta G_2$ , and the position of the transition state,  $\Delta x_2$ , are shown in Table S1.

**The Extent of Protein Collapse from the Unfolded State Is Significantly Enhanced by Glycerol.** In our single-molecule assay, the complete unfolding reaction leads the protein to a totally extended conformation, close to its full contour length, where all amino acids are exposed to the solvent environment and native contact formation is rare (18, 36). Studying the collapse of such extended proteins at the single molecule level permits a direct determination of the major driving forces governing the process. In particular, the extent of protein collapse was notably altered by changing the polarity of the solvent environment, which modifies the strength of hydrophobic interactions between the amino acids in the extended polyprotein chain. These experiments strongly supported the view that hydrophobic forces play a major role in the collapse of an extended protein (18). To further understand the mechanical and physicochemical properties of the unfolded state of ubiquitin, we use the force-ramp protocol depicted in Fig. 1. The protein is first unfolded under constant force to a highly extended conformation. Then, the force is linearly decreased down to 10 pN, triggering the contraction of the protein. This protocol permits the observation of the full force-length relationship of the extended protein as it collapses. Fig. 3A shows the collapsing trajectories of 136 different fully extended polyubiquitin molecules such as the one observed in Fig. 1 in a 30% vol/vol glycerol solution. Remarkably, we observe a high degree of heterogeneity in the responses. Such heterogeneity can be ex-

plained in terms of the different strength of the hydrophobic driving forces that act on each particular collapse trajectory, which depends on the initial combination of dihedral angles sampled by the protein once it is unfolded and stretched at 190 pN (18). In some cases, the protein collapses very little during the ramp down of the force down to 10 pN ( $F_2$  region in Fig. 1). In others, a large contraction of the protein was observed. Proteins were often observed to fold during the force-ramp protocol, as in the example shown in Fig. 1. These folding events were detected as an increase in the length by multiples of approximately 20 nm after restoration of the force to 190 pN. To compare all of the recordings, we normalized their length by the value measured in the initial extended conformation at 190 pN. The normalized length,  $L(F)/L(190 \text{ pN})$ , is shown in Fig. 3A as a function of the force during the ramp down in force to 10 pN. Proteins that failed to fold during the ramp down process (blue traces in Fig. 3A,  $n = 65$ ) showed large variations in their extent of collapse. Proteins that folded (red traces in Fig. 3A,  $n = 71$ ) collapsed much further than the failures. This is apparent from the histogram of  $L(F)/L(190 \text{ pN})$  measured at 12 pN in Fig. 3D where proteins which failed to fold are shown in blue and proteins that successfully folded are shown in red. In the presence of glycerol the extent of collapse is greatly enhanced, causing a much larger number of molecules to fold (52%) during the ramp-down of the force (red traces in Fig. 3A). Strikingly, the large hydrophobic collapse measured in 30% vol/vol glycerol solutions contrasts with the much reduced collapse observed in PBS solution and 40% vol/vol ethanol solution. In particular, 40% vol/vol ethanol greatly reduces the extent of collapse and prevents folding during the force ramp down (18). This is readily apparent when comparing the extent of collapse measured at 12 pN. Fig. 3B–D show  $L(F)/L(190)$  at 12 pN for proteins exposed to 40% vol/vol ethanol (Fig. 3B), PBS solution (Fig. 3C), and 30% vol/vol glycerol (Fig. 3D).





**Fig. 3.** The extent of protein collapse from highly extended states is enhanced in glycerol solutions. (A) Plot of 136 individual collapse trajectories in a 30% vol/vol glycerol solution such as the one shown in Fig. 1 ( $F_2$ ). To compare all of the recordings, we normalized their length,  $L(F)$ , by the value measured in the initial extended conformation at 190 pN,  $L(190 \text{ pN})$ . The normalized length,  $L(F)/L(190 \text{ pN})$ , is recorded as the pulling force is linearly decreased from 190 pN down to 10 pN. Proteins that failed to fold during the ramp down process ( $n = 65$ ) are shown in blue. Proteins that folded (which is confirmed during the ramp-up section of the recording as depicted in Fig. 1;  $F_3$  region) are shown in red and collapsed much further than the failures ( $n = 71$ ). The large hydrophobic collapse measured in 30% glycerol solutions (red bars, folders; blue bars, failures) contrasts with the much reduced collapse observed in a 40% ethanol solution. The extent of the collapse in each solvent environment can be compared by measuring the normalized protein length,  $L(F)/L(190)$ , at a force of 12 pN in (A) 40% ethanol, (B) PBS solution, and (C) 30% glycerol.

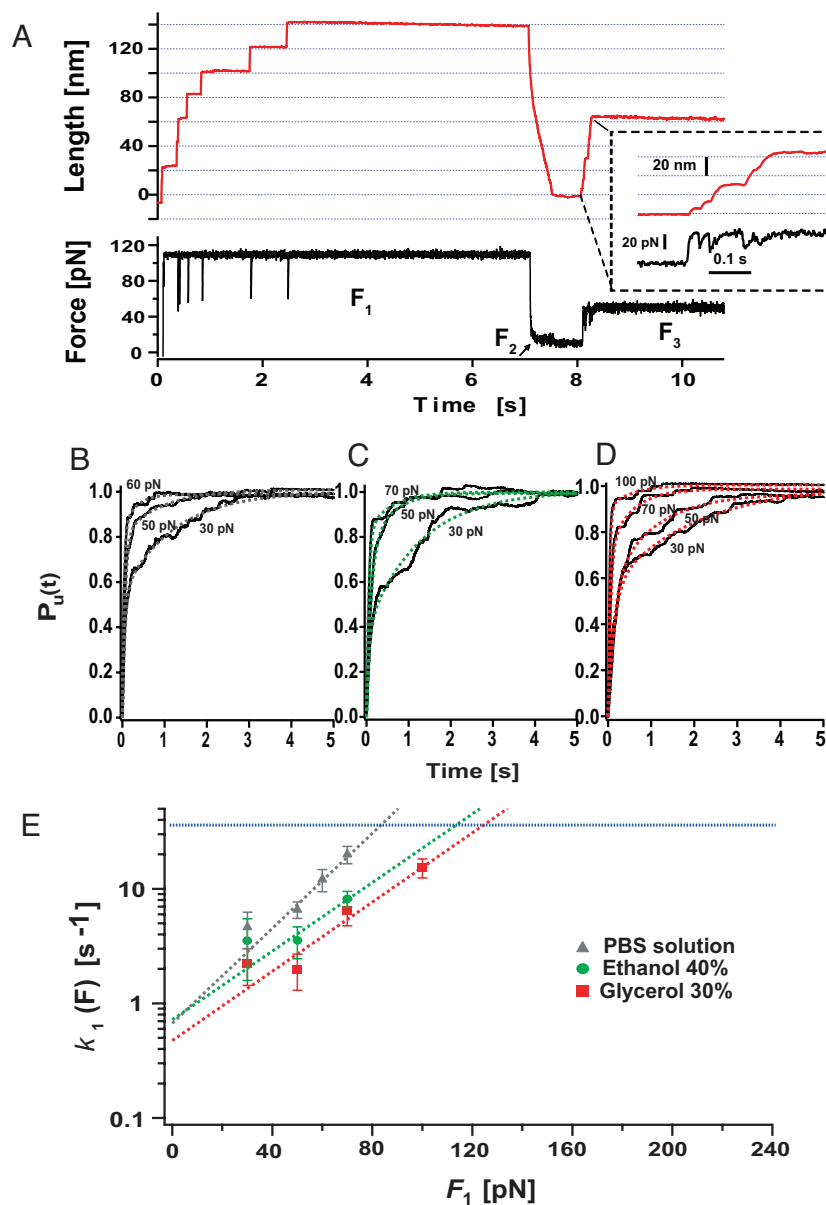
### Probing the Physicochemical Properties of the Ensemble of Collapsed Conformations.

The combination of force clamp and solvent substitution has revealed the important and distinct effect that the solvent environment has on the mechanical stability of both the folded and unfolded states of ubiquitin. It is therefore appealing to investigate the response of the recently identified ensemble of collapsed states (25) upon the application of a stretching force under different solvent environments. To study the force-dependent unraveling rate of this ensemble of conformations in standard PBS solution, we devised a force quench protocol (Fig. 4A) described in detail in the *SI Text* section that specifically probes the mechanical resistance of these structures under force. Fig. 4E shows a logarithmic plot of the rate of unraveling of the collapsed conformations,  $k_1$ , as a function of the pulling force (gray triangles) in standard PBS aqueous solution (25). We fitted the force-dependency of  $k_1$  with a simple Arrhenius term, obtaining  $k_{0,1} = 0.7 \text{ s}^{-1}$  and  $\langle \Delta x_1 \rangle = 2.0 \text{ \AA}$ . Assuming the same prefactor value used to describe the unfolding energy barrier of the native conformations (34) of  $A = 10^6 \text{ s}^{-1}$ , we calculate an average activation energy of  $\langle \Delta G_1 \rangle$  approximately 8.4 kcal/mol (25). It is remarkable that the distance to the transition state for the collapsed structures is similar to that of the native state  $\Delta x_1$  approximately 2 Å. We now extend these results by studying the effect of solvent substitution on the force dependency of the unraveling rate of the collapsed conformations (Fig. 4E). Fig. 4C and D show averaged traces and their corresponding double exponential fits obtained at different forces in solutions containing 40% vol/vol ethanol and 30% vol/vol glycerol. Fig. 4E shows the force dependency of the unraveling rate of the collapsed conformations in 40% vol/vol ethanol (green circles) and 30% vol/vol glycerol (red squares). These results demonstrate that both ethanol and glycerol have a minor effect on the force-dependency of the collapsed states. Bell fits to the data (solid lines) give activation energies and distances to the transition state similar to those found in PBS solution ( $\Delta x_{\text{ethanol}} = 1.4 \text{ \AA}$ ,  $\Delta x_{\text{gly}} = 1.4 \text{ \AA}$ ). Extrapolated to zero force, the unraveling rates converge to the same value of  $\approx 0.5\text{--}0.7 \text{ s}^{-1}$ , respectively (Fig. 4D), which implies a vanishingly small change in the height of the unraveling barrier ( $\Delta \Delta G_1 = 0.2 \text{ kcal/mol}$ ).

### Discussion

Protective and denaturing osmolytes such as ethanol and glycerol are known to destabilize or stabilize the native state of proteins against a wide variety of environmental stresses such as high

pressure (37) or temperature (38). Here, we have discovered that ethanol and glycerol affect the activation energy barrier to mechanical unfolding of the protein ubiquitin, expanding the repertoire of known function of these osmolytes. Specifically, our results show that although ethanol slightly destabilizes the native state of ubiquitin ( $\Delta \Delta G = 0.05 \text{ kcal/mol}$ ), glycerol significantly enhances it by  $\Delta \Delta G = 3.8 \text{ kcal/mol}$ . Although the effects of both ethanol and glycerol are in qualitative agreement with the results obtained from transfer free energy calculations (39), the molecular origin of these changes to the free energy remain elusive. Indeed, despite extensive studies it has proved difficult to provide a universal molecular mechanism of solvent effects on protein destabilization/stabilization. Although a ‘direct mechanism’ suggests that the cosolvent directly and favorably interacts with the protein backbone (40), an ‘indirect mechanism’ suggests that the cosolvent disrupts the water structure, thereby modifying its solvation properties (41). Our experiments at the single molecule level provide a different perspective to this important area by directly capturing the role of solvent molecules in the structure of the mechanical unfolding transition state of a protein. Crucially, this level of experimental control provides a unique test of either a direct or indirect solvent stabilization mechanism in the context of mechanical stability (35). In our experiments probing mechanical unfolding of ubiquitin in 40% vol/vol ethanol show that  $\Delta x$  remains unchanged in this aqueous ethanol solution as compared with water ( $\Delta x_{\text{ethanol}} = 1.7 \text{ \AA}$ ,  $\Delta x_{\text{water}} = 1.6 \text{ \AA}$ ), suggesting that ethanol molecules are not actively involved in the unfolding transition state structure of this protein. These results therefore suggest that ethanol slightly destabilizes ubiquitin through an indirect mechanism, disrupting the water structure. Indeed, previous structural studies of aqueous alcohol solutions have shown that the extended hydrogen bond network of water is significantly altered in the presence of an alcohol (42). In contrast to ethanol, the distance to the transition state is measured to increase from  $\Delta x_{\text{water}} = 1.6 \text{ \AA}$  in water to  $\Delta x_{\text{gly}} = 2.3 \text{ \AA}$  upon addition of 30% vol/vol glycerol, suggesting that the larger molecule glycerol plays an active role in the transition state structure, which is consistent with a direct mechanism of protein stabilization. The involvement of glycerol molecules in the mechanical transition state of unfolding observed for ubiquitin, resulting in an increase in  $\Delta x$ , is consistent with the results obtained for the protein I27 (35). Interestingly, while the increase in  $\Delta x$  for the protein I27 in glycerol was approximately 1.6 Å the increase in  $\Delta x$  for ubiquitin is only 0.7 Å. This smaller increase in  $\Delta x$  suggests that



**Fig. 4.** The collapsed conformations are solvent-insensitive. (A) We use the 3-pulse protocol ( $F_1$  to  $F_3$ ) to measure the force-dependent extension of the collapsed ensemble. After a standard force-quench sequence ( $F_1$  and  $F_2$ ), the force-dependent rate of extension of the collapsed conformations  $k_1$  is measured by varying  $F_3$  within the range 30–70 pN in (B) PBS aqueous solution, data from (25), (C) ethanol solution (40% vol/vol), and (D) glycerol solution (30% vol/vol). We measured the unfolding rate from the weighted average of biexponential fits to the unraveling time course of 25–50 recordings at each force. (E) We fit an Arrhenius term to  $k_1(F)$  to estimate the average size of the activation energy barrier,  $\langle \Delta G_1 \rangle$ , and distance to the transition state,  $\langle \Delta x_1 \rangle$  of the collapsed conformations in PBS solution and in solutions containing 40% ethanol and 30% glycerol. The data shows that the Arrhenius fits for all 3 solvent environments converge to the same rate at 0 force, indicating that  $\langle \Delta G_1 \rangle$  is solvent-independent.

the mechanical transition state structure in ubiquitin, thought to be placed between the  $\beta 1$ – $\beta 5$  sheets (43, 44), is less sensitive to the size of the solvent molecules.

In addition to the native state of ubiquitin, the extended state is also found to be highly sensitive to the solvent environment. While protein collapse is significantly hampered in an aqueous ethanol environment, it is readily enhanced in aqueous glycerol. Protein stability is known to be strongly determined by its solvation properties with the surrounding solvent environment. The observation of an enhancement in the hydrophobic collapse of ubiquitin in glycerol is not inconsistent with a proposed backbone based theory of protein folding, in which unfavourable interactions between the protein backbone and glycerol's polar surface area drive the collapse of the unfolded protein (9, 45). However, the heterogeneity observed in the collapsing trajectories can only be explained in terms of the different initial solvation properties of side chains in the unfolded protein. These initial conformations result from the different combination of dihedral angles explored by the extended polyprotein (46). The heterogeneity in the collapsing trajectories, mainly driven by hydrophobic interactions, is likely to

result in the conformational disorder observed in the ensemble of collapsed states (Fig. 1, *Bottom Inset*).

Strikingly, the behavior of both the native and extended proteins contrasts with the measured solvent insensitivity of the collapsed conformations. It is interesting to consider what structural properties of the collapsed conformations could result in this solvent insensitivity. One plausible explanation is that solvent molecules are excluded from the internal structure of the collapsed conformations. Thus solvent molecules would not play an integral role in their transition state structure. Indeed, this scenario would account for the similar  $\Delta x$  values measured for the different solvents (Fig. 4 and Table S1), suggesting that solvent bridging is not an important mechanism for these structures. Instead, each collapsed conformation may be held together by interactions that are mostly hydrophobic in origin. The mechanical resistance exhibited by the collapsed conformations would thus reflect the strength of this hydrophobic packing. Interestingly, such a hydrophobic packing scenario has been observed for the protein ankyrin, which consists of smaller ankyrin repeats which stack into a helical supramolecular structure via hydrophobic forces (47). Similarly to our single

molecule experiments, such ankyrin structures are largely insensitive to osmolytes (48), yet exhibit mechanical resistance which is similar in scale to that measured for the collapsed conformations (49). Alternatively, the solvent insensitivity of the collapsed conformations may reflect structures which are preferentially solvated by water molecules from the ethanol/glycerol aqueous mixtures. Thus the transition state structures of these structures will only involve water molecules resulting in an unchanging  $\Delta x$  and  $\Delta G$ . In this scenario, the collapsed conformations would require final expulsion of water molecules to reach the native fold. Indeed, this is not inconsistent with previous computational studies which suggest that the final step of the folding process from a near-native collapsed structure to the final folded conformation involves a desolvation process that squeezes out water molecules in the vicinity of the partially hydrated core (1). It is feasible that in our experiments, such a folding step is represented at the end of the collapsing trajectories (Figs. 1 and 3), which are largely affected by the solvent environment. Indeed, once the protein reaches the folded length, it remains a puzzle how the native contacts are formed to regain its mechanical stability. We hypothesize that this transition encompasses the correct rearrangement of side-chains and contact pairs in each individual collapsed structure to reach the final native structure, and that this last step is not crucially determined by the solvent environment.

Here, we demonstrate the success of using solvent substitution combined with single molecule force-clamp spectroscopy as an

independent probe of the evolving structures of a polypeptide as it folds. These experiments open the way for a detailed examination of the transition state structures that populate the folding trajectory of a protein.

## Materials and Methods

**Protein Engineering.** WT-Ubiquitin polyprotein was subcloned using the *Bam*HI and *Bgl*II restriction sites (50). The nine-domain ubiquitin was cloned into the pQE80L (Qiagen) expression vector, and transformed into the BLR(DE3) *Escherichia coli* expression strain and purified by histidine metal-affinity chromatography with Talon resin (Clontech) and by gel-filtration using a Superdex 200 HR column (GE Biosciences).

**Force Spectroscopy.** Force-clamp atomic force microscopy experiments were conducted at room temperature using a home-made set-up under force-clamp conditions described elsewhere (10). Experiments were carried out in a sodium phosphate buffer solution, specifically, 50 mM sodium phosphate ( $\text{Na}_2\text{HPO}_4$  and  $\text{NaH}_2\text{PO}_4$ ), and 150 mM NaCl, pH = 7.4. Samples of glycerol ( $\geq 99\%$ ) and ethanol ( $\geq 99\%$ ) were obtained from Sigma-Aldrich and used without additional purification. Solvent mixtures were prepared to obtain the desired volume fraction, vol/vol, ratio of the co-solvent and viscosity.

**ACKNOWLEDGMENTS.** We thank Jasna Brujić and members of the J.M.F. laboratory for discussions. S.G.-M. thanks the Generalitat de Catalunya for a postdoctoral fellowship through the NANO and Beatriu de Pinós programs, and also the Fundación Caja Madrid for financial support. This work was supported by National Institutes of Health Grants HL66030 and HL61228 (to J.M.F.).

- Cheung MS, Garcia AE, Onuchic JN (2002) Protein folding mediated by solvation: Water expulsion and formation of the hydrophobic core occur after the structural collapse. *Proc Natl Acad Sci USA* 99:685–690.
- Tarek M, Tobias DJ (2002) Role of protein-water hydrogen bond dynamics in the protein dynamical transition. *Phys Rev Lett* 88:138101.
- Qu YX, Bolen CL, Bolen DW (1998) Osmolyte-driven contraction of a random coil protein. *Proc Natl Acad Sci USA* 95:9268–9273.
- Pal SK, Peon J, Zewail AH (2002) Biological water at the protein surface: Dynamical solvation probed directly with femtosecond resolution. *Proc Natl Acad Sci USA* 99:1763–1768.
- Sorenson JM, Hura G, Soper AK, Pertsemlidis A, Head-Gordon T (1999) Determining the role of hydration forces in protein folding. *J Phys Chem B* 103:5413–5426.
- Rariy RV, Klibanov AM (1997) Correct protein folding in glycerol. *Proc Natl Acad Sci USA* 94:13520–13523.
- Gekko K, Timasheff SN (1981) Mechanism of protein stabilization by glycerol: Preferential hydration in glycerol-water mixtures. *Biochemistry* 20:4667–4676.
- O'Brien EP, Ziv G, Haran G, Brooks BR, Thirumalai D (2008) Effects of denaturants and osmolytes on proteins are accurately predicted by the molecular transfer model. *Proc Natl Acad Sci USA* 105:13403–13408.
- Street TO, Bolen DW, Rose GD (2006) A molecular mechanism for osmolyte-induced protein stability. *P Natl Acad Sci USA* 103:13997–14002.
- Schlierf M, Li H, Fernandez JM (2004) The unfolding kinetics of ubiquitin captured with single-molecule force-clamp techniques. *Proc Natl Acad Sci USA* 101:7299–7304.
- Wiita AP, Ainaravaru SRK, Huang HH, Fernandez JM (2006) Force-dependent chemical kinetics of disulfide bond reduction observed with single-molecule techniques. *Proc Natl Acad Sci USA* 103:7222–7227.
- Li H, Carrion-Vazquez M, Oberhauser AF, Marszalek PE, Fernandez JM (2000) Point mutations alter the mechanical stability of immunoglobulin modules. *Nat Struct Biol* 7:1117–1120.
- Brockwell DJ, et al. (2005) Mechanically unfolding the small, topologically simple protein L. *Biophys J* 89:506–519.
- Perez-Jimenez R, Garcia-Manyes S, Ainaravaru SR, Fernandez JM (2006) Mechanical unfolding pathways of the enhanced yellow fluorescent protein revealed by single molecule force spectroscopy. *J Biol Chem* 281:40010–40014.
- Lu H, Schulten K (2000) The key event in force-induced unfolding of titin's immunoglobulin domains. *Biophys J* 79:51–65.
- Dougan L, Feng G, Lu H, Fernandez JM (2008) Solvent molecules bridge the mechanical unfolding transition state of a protein. *Proc Natl Acad Sci USA* 105:3185–3190.
- Dougan L, Koti AS, Genchev G, Lu H, Fernandez JM (2008) A single-molecule perspective on the role of solvent hydrogen bonds in protein folding and chemical reactions. *Chemphyschem* 22:2836–2847.
- Walther KA, Grater F, Dougan L, Badilla CL, Berne BJ, Fernandez JM (2007) Signatures of hydrophobic collapse in extended proteins captured with force spectroscopy. *Proc Natl Acad Sci USA* 104:7916–7921.
- Lipman EA, Schuler B, Bakajin O, Eaton WA (2003) Single-molecule measurement of protein folding kinetics. *Science* 301:1233–1235.
- Hoffmann A, et al. (2007) Mapping protein collapse with single-molecule fluorescence and kinetic synchrotron radiation circular dichroism spectroscopy. *Proc Natl Acad Sci USA* 104:105–110.
- Cecconi C, Shank EA, Bustamante C, Marqusee S (2005) Direct observation of the three-state folding of a single protein molecule. *Science* 309:2057–2060.
- Fernandez JM, Li H (2004) Force-clamp spectroscopy monitors the folding trajectory of a single protein. *Science* 303:1674–1678.
- Schlierf M, Berkemeier F, Rief M (2007) Direct observation of active protein folding using lock-in force spectroscopy. *Biophys J* 93:3989–3998.
- Garcia-Manyes S, Brujić J, Badilla CL, Fernandez JM (2007) Force-clamp spectroscopy of single-protein monomers reveals the individual unfolding and folding pathways of I27 and ubiquitin. *Biophys J* 93:2436–2446.
- Garcia-Manyes S, Dougan L, Badilla CL, Brujić J, Fernandez JM (2009) Direct observation of an ensemble of stable collapsed states in the mechanical folding of ubiquitin. *Proc Natl Acad Sci USA* doi/10.1073/pnas.0901213106.
- Honeycutt JD, Thirumalai D (1990) Metastability of the folded states of globular proteins. *Proc Natl Acad Sci USA* 87:3526–3529.
- Camacho CJ, Thirumalai D (1993) Minimum energy compact structures of random sequences of heteropolymers. *Phys Rev Lett* 71:2505–2508.
- Camacho CJ, Thirumalai D (1993) Kinetics and thermodynamics of folding in model proteins. *P Natl Acad Sci USA* 90:6369–6372.
- Phillips CL, Thrower J, Pickart CM, Hill CP (2001) Structure of a new crystal form of tetraubiquitin. *Acta Crystallogr D Biol Crystallogr* 57:341–344.
- Cook WJ, Jeffrey LC, Kasperek E, Pickart CM (1994) Structure of tetraubiquitin shows how multiubiquitin chains can be formed. *J Mol Biol* 236:601–609.
- Cook WJ, Jeffrey LC, Carson M, Chen Z, Pickart CM (1992) Structure of a diubiquitin conjugate and a model for interaction with ubiquitin conjugating enzyme (E2). *J Biol Chem* 267:16467–16471.
- Brujić J, Hermans RI, Walther KA, Fernandez JM (2006) Single-molecule force spectroscopy reveals signatures of glassy dynamics in the energy landscape of ubiquitin. *Nat Phys* 2:282–286.
- Bell GI (1978) Models for the specific adhesion of cells to cells. *Science* 200:618–627.
- Li MS, Klimov DK, Thirumalai D (2004) Thermal denaturation and folding rates of single domain proteins: Size matters. *Polymer* 45:573–579.
- Dougan L, Fang G, Lu H, Fernandez JM (2008) Solvent molecules bridge the unfolding transition state of a protein. *Proc Natl Acad Sci USA* 105.
- Li MS, Hu CK, Klimov DK, Thirumalai D (2006) Multiple stepwise refolding of immunoglobulin domain I27 upon force quench depends on initial conditions. *Proc Natl Acad Sci USA* 103:93–98.
- Ruan K, Xu C, Li T, Li J, Lange R, Balny C (2003) The thermodynamic analysis of protein stabilization by sucrose and glycerol against pressure-induced unfolding. *Eur J Biochem* 270:1654–1661.
- Gekko K, Timasheff SN (1981) Thermodynamic and kinetic examination of protein stabilization by glycerol. *Biochemistry* 20:4677–4686.
- Tanford C (1962) Contribution of hydrophobic interactions to stability of globular conformation of proteins. *J Am Chem Soc* 84:4240.
- Hua L, Zhou RH, Thirumalai D, Berne BJ (2008) Urea denaturation by stronger dispersion interactions with proteins than water implies a 2-stage unfolding. *Proc Natl Acad Sci USA* 105:16928–16933.
- Bennion BJ, Daggett V (2003) The molecular basis for the chemical denaturation of proteins by urea. *Proc Natl Acad Sci USA* 100:5142–5147.
- Dougan L, et al. (2004) Methanol-water solutions: A bi-percolating liquid mixture. *J Chem Phys* 121:6456–6462.
- Irbäck A, Mitternacht S, Mohanty S (2005) Dissecting the mechanical unfolding of ubiquitin. *Proc Natl Acad Sci USA* 102:13427–13432.
- Kleiner A, Shakhnovich E (2007) The mechanical unfolding of ubiquitin through all-atom Monte Carlo simulation with a Go-type potential. *Biophys J* 92:2054–2061.
- Rose GD, Fleming PJ, Banavar JR, Maritan A (2006) A backbone-based theory of protein folding. *Proc Natl Acad Sci USA* 103:16623–16633.
- Auton M, Bolen DW (2005) Predicting the energetics of osmolyte-induced protein folding/unfolding. *Proc Natl Acad Sci USA* 102:15065–15068.
- Lee G, Abdi K, Jiang Y, Michael P, Bennett V, Marszalek PE (2006) Nanospring behaviour of ankyrin repeats. *Nature* 440:246–249.
- Wetzel SK, Settanni G, Kenig M, Binz HK, Pluckthun A (2008) Folding and unfolding mechanism of highly stable full-consensus ankyrin repeat proteins. *J Mol Biol* 376:241–257.
- Li L, Wetzel S, Pluckthun A, Fernandez JM (2006) Stepwise unfolding of ankyrin repeats in a single protein revealed by atomic force microscopy. *Biophys J* 90:L30–32.
- Carrion-Vazquez M, et al. (1999) Mechanical and chemical unfolding of a single protein: a comparison. *Proc Natl Acad Sci USA* 96:3694–3699.