In a recent experiment (1), the folding of a single polyprotein consisting of nine N-C linked ubiquitin repeats was monitored under a constant force with an atomic force microscope (AFM). Unlike the discrete, staircase-like increase in the length of the polyubiquitin chain observed in unfolding experiments at higher forces, folding was characterized by a slow contraction in the end-to-end length of the protein \( R_{\text{NC}} \), terminated by a sharp drop to the length at which all the ubiquitin modules are folded. This result has been interpreted as either a breakdown of two-state kinetics under force or evidence of initial aggregation events prior to folding (3), because monomeric ubiquitin is known to aggregate at concentrations above 2 μM (4). By simulating the dynamics of folding under force using a simplified representation of the protein, we show that the absence of steplike changes in the overall length upon folding of individual modules can be explained by the elastic properties of the remaining unfolded modules.

In our simulations, the protein is represented by a chain of alpha carbon beads with a Gö-like potential, in which attractive interactions occur only between residues that form contacts in the folded structure (5). However, by including favorable interactions for both intra- and intermodule native contacts, folding competes with aggregation. Following the experimental protocol, we started with an unfolded four-module polyubiquitin model under high force and then quenched the force to a low value that favors folding. The simulations capture the experimentally observed behavior (1): A rapid initial collapse is followed by a period of slow contraction, characterized by large fluctuations in the protein extension, and a final jump to the fully folded length (Fig. 1A).

What is causing this behavior? In contrast to the experiment in (1), we can monitor the folding of each subunit individually in our simulation. The fractions \( Q_i \) of native contacts for each subunit \( i \) exhibit sharp cooperative transitions from typical unfolded values of \( Q < 0.5 \) to a folded \( Q \) close to 1 (Fig. 1, B to E). At the same time, the end-to-end distance (between the first and last residue) of each subunit drops to near-native values upon unfolding, with a concomitant reduction in fluctuations. Remarkably, though, the folding events (red arrows in Fig. 1A) are not at all obvious from monitoring the overall length of the protein. The reason is that the remaining unfolded modules act as soft entropic springs whose large fluctuations mask the decrease in contour length when a single module folds. This effect is not seen in high-force unfolding experiments because the unfolded modules are fully stretched, so the contour length increases in discrete steps upon each unfolding event (2). Protein elasticity also explains the slower response for the final jump to the folded state in the AFM experiments (1, 3) compared with recent single-molecule fluorescence studies (6).

There is little evidence of aggregation in our simulations, despite the inclusion of native-like interactions between different modules. The fraction of intermodule native-like contacts between any two modules is always <0.05. Modifying the model to include additional non-native contacts in the energy function resulted in a more frustrated system and slower folding, but did not qualitatively alter the results presented above. In connection with this, we note that ubiquitin is in fact biosynthesized as a polyubiquitin fusion protein similar to the one studied in the AFM experiment (7), although the cell is clearly a more complex folding environment.

Although these simulations do not preclude other explanations (which may well contribute in the experiment), they demonstrate that a nontrivial folding model with cooperative folding transitions is sufficient to describe the essential features of the measurements. This interpretation is also consistent with other experimental observations. The increase in folding time with contour length (1) can be related to the expected increase in folding time with the number of unfolded repeats for independently folding modules (8). In addition, the approximate extrapolation of experimental folding time to zero force, 0.01 s (1), is within an order of magnitude of that found using chemical denaturant under similar conditions (4).

Individual protein-folding events may be fully resolvable with the experimental setup of Fernandez et al. (1, 2) by operating at higher forces whereby fluctuations in the end-to-end length are reduced, although folding is slowed exponentially by force. Alternatively, stiffer linkers could be useful—for example, polyproteins consisting of multiple titin modules, unfolding only at high forces, and a single ubiquitin module.

Robert B. Best and Gerhard Hummer*
Laboratory of Chemical Physics
National Institute of Diabetes and Digestive and Kidney Diseases
National Institutes of Health,
Building 5, Bethesda MD 20892–0520, USA.
*To whom correspondence should be addressed.
E-mail: gerhard.hummer@nih.gov
References and Notes
8. If the folding-time distribution of a single protein module is exponential with mean folding time $t$, the mean folding time for an $n$-module protein is given by $\tau_n = t(1 + \psi(n + 1)) = t\gamma + \ln(n + 1) - \frac{(2n + 2)^{-1}}{2}$, where $\gamma = 0.5772$ and $\psi(x)$ are Euler's constant and psi function, respectively.

1 November 2004; accepted 24 February 2005
10.1126/science.1106969