Refolding dynamics of stretched biopolymers upon force quench

Changbong Hyeon[a,1], Greg Morrison[b,c], David L. Pincus[c], and D. Thirumalai[c,d,1]

[a]Department of Chemistry, Chung-Ang University, Seoul 156-756, Republic of Korea; [b]School of Engineering and Applied Science, Harvard University, Cambridge, MA 02138; and [c]Biophysics Program, Institute for Physical Science and Technology; and [d]Department of Chemistry and Biochemistry, University of Maryland, College Park, MD 20742

Single-molecule force spectroscopy methods can be used to generate folding trajectories of biopolymers from arbitrary regions of the folding landscape. We illustrate the complexity of the folding kinetics and generic aspects of the collapse of RNA and proteins upon force quench by using simulations of an RNA hairpin and theory based on the de Gennes model for homopolymer collapse. The folding time, τ_F, depends asymmetrically on $\delta f_S = f_S - f_m$ and $\delta f_Q = f_m - f_Q$ where f_S (f_Q) is the stretch (quench) force and f_m is the transition midforce of the RNA hairpin. In accord with experiments, the relaxation kinetics of the molecular extension, $R(t)$, occurs in three stages: A rapid initial decrease in the extension is followed by a plateau and finally, an abrupt reduction in $R(t)$ occurs as the native state is approached. The duration of the plateau increases as $\lambda = \tau_Q/\tau_F$ decreases (where τ_Q is the time in which the force is reduced from f_S to f_Q). Variations in the mechanisms of force-quench relaxation as λ is altered are reflected in the experimentally measurable time-dependent entropy, which is computed directly from the folding trajectories. An analytical solution of the de Gennes model under tension reproduces the multistage stage kinetics in $R(t)$. The prediction that the initial stages of collapse should also be a generic feature of polymers is validated by simulation of the kinetics of toroid (globule) formation in semiflexible (flexible) homopolymers in poor solvents upon quenching the force from a fully stretched state. Our findings give a unified explanation for multiple disparate experimental observations of protein folding.

The folding of RNA molecules (1) and proteins (2–5) should be thought of as dynamic changes in the distribution of conformations that result in collapse, the formation of intermediates, and barrier crossings to reach the folded structures. Such a statistical, mechanical perspective of the folding process is finding support in single-molecule measurements that manipulate the structural ensembles by using an external mechanical force (f) (6–8). The development of the force-clamp methods (6, 8), which allow the application of a constant force to specific locations on a biomolecule, have made it possible to explore in the most straightforward manner the folding of proteins and RNA initiated from regions of the folding landscape that are inaccessible in conventional ensemble experiments. The use of an initial stretching force, f_S, which fully unfolds the biomolecule, followed by a subsequent quench to a sufficiently low force, f_Q, which populates the native basin of attraction (NBA), holds the promise of unearthing all aspects of the folding reaction, including the dynamics of the collapse process and the extent to which the folding pathways are heterogeneous.

Although there is great heterogeneity in the folding trajectories, Fernandez and Li (6) noted that upon the quench $f_S \rightarrow f_Q$, ubiquitin (Ub) folded in three stages, as reflected in the time-dependent changes in the extension, $R(t)$. Typically, after a very rapid decrease in $R(t)$, there is a long plateau in $R(t)$ that is suggestive of the formation of metastable states, which we term force-induced metastable intermediates (FIMI's). Finally, $R(t)$ decreases sharply in the last stage as the extension corresponding to the native state is reached. Although it is not emphasized, a similar behavior has been observed in the folding of TAR RNA

(see figure 3*A* of ref. 8) upon abrupt decrease in f, even though the forces used in atomic force microscopy (AFM) experiments on Ub and laser optical tweezers (LOT) experiments on the RNA differ greatly. In contrast, it has been argued that Ub folds in a discrete manner based on refolding initiated by decreasing force slowly by using AFM (7) and not in the way described in ref. 6. The seemingly contradictory conclusions on the refolding of an initially extended Ub require a full theoretical description of the molecular events that ensue when force is decreased from f_S.

In order to quantitatively describe the results of single-molecule measurements, the interplay between a number of factors has to be taken into account. The variables that can be controlled in experiments are f_S, f_Q, and v_L, the speed with which the position of the transducer is contracted (over a length ΔL) or equivalently, r_c, the rate of force contraction. Because the routes navigated by the biomolecule will depend on the precise protocol used (8, 9), these factors must be taken into account when comparing different experiments, even if the biomolecule under consideration is the same. For a system whose global folding obeys two-state kinetics, the two relevant time scales are the force-relaxation time, τ_Q (with $\tau_Q \approx \Delta L/v_L$ or equivalently, $\tau_Q \approx (f_S - f_Q)/r_c$) and the overall folding time, $\tau_F(f_S, f_Q)$. The dynamics of the folding trajectories, $R(t)$, should depend on $\lambda = \tau_Q/\tau_F$. If $\lambda \ll 1$, then folding takes place far from equilibrium, whereas in the opposite limit refolding occurs under near-equilibrium conditions. Thus, by altering the force-quench protocol, it is possible to control the collapse process and the folding routes for a given protein or RNA, even if f_Q and f_S are fixed.

Here, we explore quantitatively the scenarios that emerge for the folding of a RNA hairpin as a function of f_Q and f_S. The value of λ is controlled by altering $\tau_F(f_Q, f_S)$ in this work, rather than by adjusting the force contraction timescale τ_Q, as in refs. 6 and 7. We place particular emphasis on the general features of the physics underlying the early stages of length contraction. The folding trajectories for a RNA hairpin exhibit a long plateau of varying length in $R(t)$, associated with the development of a FIMI. More generally, we show that the folding trajectories can be controlled by the choice of an appropriate force-quench protocol, which is succinctly expressed in terms of λ. The origin of the long-lived FIMIs when λ is small is explained by adapting de Gennes' picture of the kinetics of a coil to globule transition in flexible homopolymers. The combination of simulations and theory shows that the formation of long-lived FIMIs should be a generic feature of any biomolecule that adopts a compact structure, which implies that the polymeric nature of RNA and proteins, along with interactions that induce a
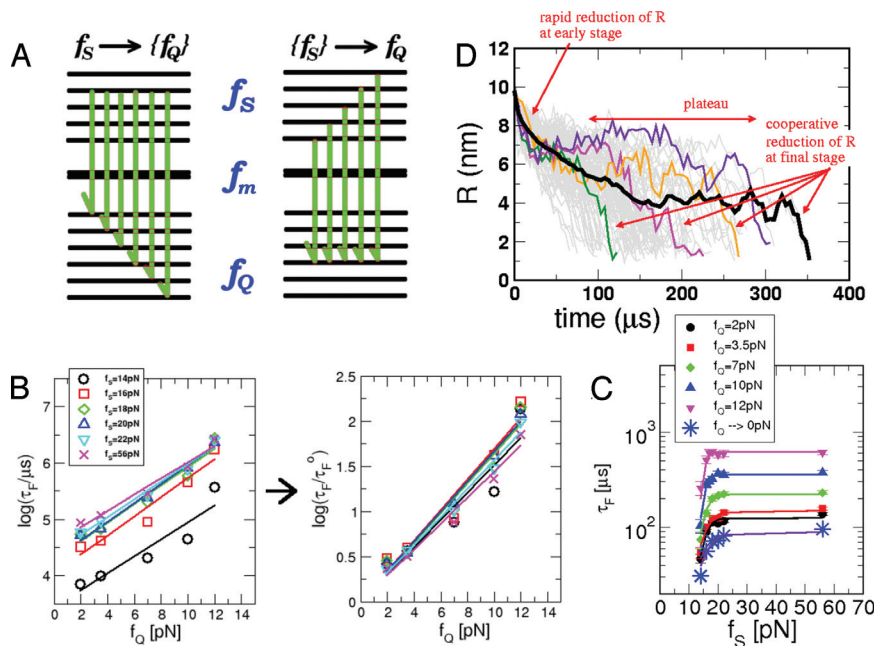
Fig. 1. Force-quench kinetics of P5GA hairpins. (*A*) Refolding from a fixed f_S to varying f_Q (left) or from varying f_S to a fixed f_Q (right). (*B*) τ_F versus f_Q starting from different f_S. A fit to the Bell equation gives $\Delta x_{U \to N}^{\ddagger} \approx (0.6 - 0.7)$ nm. The extrapolated value τ_F^o to zero quench force is plotted in *C* (∗ symbols). (*C*) τ_F versus f_S for each value of f_Q. The variations of τ_F with f_S are fit to Eq. **1**. From the fit, we find $(\Delta x_{UN}/\text{nm}, \tau(f_Q)/\mu s) = (1.5, 89.7), (2.5, 126), (1.8, 150), (2.6, 224), (3.8, 357), (4.4, 622)$ for $f_Q = 0, 2, 3.5, 7, 10, 12$ pN, respectively. (*D*) Time traces of the molecular extension, $R(t)$, upon force quench. Multiple time traces are plotted in gray, and the averaged time trace is shown by the thick black line. A few representative time traces are in color. The ensemble of time traces are shown for $f_S = 70$ pN and $f_Q = 3.5$ pN.

globular state for sufficiently low f_Q, will generically lead to a long plateau in $R(t)$ as long as λ is small. These conclusions are supported by explicit simulations of force-quench studies of a flexible homopolymer in a poor solvent and a semiflexible chain model for DNA, in which the compact state adopts a toroidal structure. Our theory, which emphasizes λ as the relevant variable in determining the folding routes, naturally explains the apparent differences in the interpretation of force-quench folding of Ub by different groups (6, 7).

Results

Asymmetry in the Force-Quench Kinetics. Force-quench refolding of a RNA hairpin (or a protein) can be initiated by using two distinct modes. In the first, the RNA is equilibrated at an initial $f_S > f_m$, where f_m is the transition force at which the probabilities of being folded and unfolded are equal, and subsequently the force is quenched to various f_Q values ($f_S \to \{f_Q\}$), with $f_Q < f_m$ (Fig. 1). We use simulations of the P5GA hairpin (10), with $f_m = 14.7$ *pN* at $T = 300$ K (see *Methods*) to extract the dependence of τ_F on f_Q from the dynamics of folding trajectories (Fig. 1*D* and Fig. 2). Similarly, refolding can also be initiated by varying $f_S (> f_m)$ and reducing the force to a single $f_Q (< f_m)$ ($\{f_S\} \to f_Q$) (see Fig. 1*A*).

The mean refolding time ($\tau_F = \frac{1}{N} \sum_{i=1}^{N} \tau_F(i)$) for different values of f_Q is fit to a Bell equation (6, 11, 12) (Fig. 1*B*), $\tau_F(f_Q, f_S) = \tau_F^o(f_S) e^{f_Q \Delta x_{U \to N}^{\ddagger}/k_B T}$ where $\tau_F^o(f_S)$ is the folding time at $f_Q = 0$, and $\Delta x_{U \to F}^{\ddagger}$ is the distance between the transition state and the free energy minimum in the unfolded basin of attraction (UBA), which assumes that R is a good reaction coordinate. The slight upward curvature in $\log(\tau_F)$ as a function of f_Q is due to the dependence of $\Delta x_{U \to N}^{\ddagger}$ on f_Q (13–15). When τ_F is rescaled by $\tau_F^o(f_S)$, the refolding times at different f_S values nearly collapse onto a single curve (see the right-hand frame of Fig. 1*B*). It follows that the time scale for crossing an effective free energy barrier (in the late stages of folding) is solely determined by f_Q and is independent of f_S. The microscopic model in refs. 15 and 16 provides an excellent fit to the simulation data (see also Fig. S2 and its legend in the *SI Appendix*). Here, we use the Bell equation to provide a physically simpler picture of the effect of force on the free energy profile.

Unlike the behavior of τ_F with varying f_Q, the dependence of τ_F on f_S cannot be explained by using the Bell model. We find that τ_F decreases sharply with f_S for $f_S < 20$ pN but saturates to a limiting value (which is also dependent upon f_Q) if $f_S > 20$ pN (Fig. 1*C*).

The saturation of τ_F for large f_S suggests that the initial conformations of P5GA belong to the same structural ensemble [see the variations of the end-to-end distribution, $P_S(R)$, at each f_S in Fig. S1*A*]. At low f_S ($f_m \lesssim f_S < 20$ pN), the molecules are partitioned into different ensembles. The initial population of molecules in the UBA, determined by f_S, affects the average folding time. Because P5GA is two-state folder under force, the fraction of molecules in the UBA at f_S is $\varphi_{UBA}(f_S) = 1/[1 + e^{(\Delta F_{UN}^o - f_S \Delta x_{UN})/k_B T}]$, where $\Delta F_{UN}^o = F_U^o - F_N^o$ and $\Delta x_{UN} = x_U - x_N$ are the difference in free energy and molecular extension between unfolded and folded conformations, respectively. By combining the f_Q dependence of τ_F from Bell's model with $\varphi_{UBA}(f_S)$, we obtain a unified expression that describes force-quench folding time in both modes (see *SI Appendix* for details),

$$\tau_F(\delta f_S, \delta f_Q) = \frac{\tau(0)}{1 + e^{-\delta f_S \Delta x_{UN}/k_B T}} e^{-\delta f_Q \Delta x_{U \to N}^{\ddagger}/k_B T}, \qquad \textbf{[1]}$$
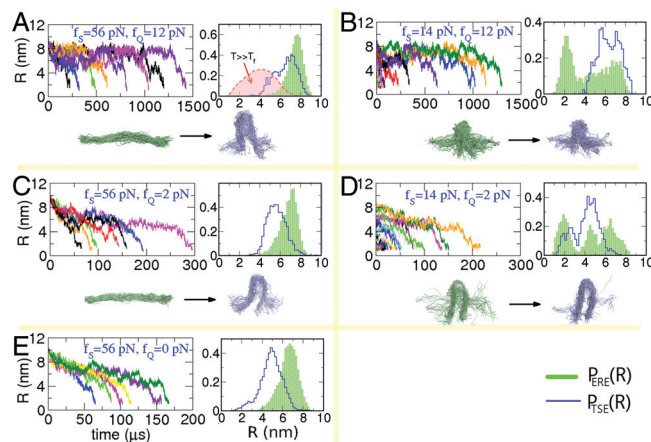
Fig. 2. Analysis of the folding trajectories associated with force quench for various values of (f_S, f_Q). (*A*) $f_S = 56$ pN → $f_Q = 12$ pN. (*B*) $f_S = 14$ pN → $f_Q = 12$ pN. (*C*) $f_S = 56$ pN → $f_Q = 2$ pN. (*D*) $f_S = 14$ pN → $f_Q = 2$ pN. (*E*) $f_S = 56$ pN → $f_Q = 0$ pN. $P_{ERE}(R)$ and $P_{TSE}(R)$ are shown in the right-hand frames of *A*–*E*, and the corresponding structures are shown for the ERE (left) and TSE (right). Thermally denatured ensemble $P_{TDE}(R)$ for $T \gg T_f$ is shown in *A*.

Table 1. Rate of force quench in various experiments

Experiments	v_L (Slew rate)	τ_Q (Quench time)	$\tau_F(f_Q)$ (Folding time)	[a] $\lambda = \tau_Q/\tau_F(f_Q)$
Polyubiquitin (6)	$\sim 10^4$ nm/s	3 msec	0.2 s ($f_Q = 15$ pN)	0.015
TAR RNA (8)	> 200 nm/s	0.1 sec	~ 1 sec	0.1
RNAase H ($I \rightarrow N$) (33)	$(10 - 10^3)$ nm/s	$(0.05 - 5)$ sec	5 sec ($I \rightarrow U$)	$(0.01 - 1)$
RNAase H ($N \rightarrow U$) (33)	$(10 - 10^3)$ nm/s	$(0.05 - 5)$ sec	3,000 sec ($N \rightarrow U$)	$(1/60,000 - 1/600)$
Ubiquitin dimer (7)	5 nm/s	~ 0.16 sec	$\gtrsim 5$ msec ($f_Q = 0$ pN)	$\lesssim 32$
P5GA hairpin (simulation)	$\sim 10^{11}$ nm/s	47 ps	$(10 - 1,000)$ μs	$(10^{-9} - 10^{-6})$
WLC (simulation)	$\sim 10^{13}$ nm/s	0.5 ps	$\gtrsim 100$ μs	$\lesssim 5 \times 10^{-9}$

[a]For $\lambda \ll 1$, we expect, a nonequilibrium effect, which leads to a population of long-lived FIMIs. In the opposite limit, folding occurs close to equilibrium conditions. The values of λ for Ub explain the differences in the refolding found in refs. 6 and 7. The value of λ in all our simulations are extremely small and cannot be realized in AFM experiments. However, we expect qualitative changes in $R(t)$ only when $\lambda > 1$, and hence our simulation results are qualitatively similar to the small $\lambda(\ll 1)$ used in ref. 6. WLC, worm-like chain.

where $\delta f_S = f_S - f_m$, $\delta f_Q = f_m - f_Q$, $\tau(0) = \tau_o e^{\Delta F_{U \rightarrow N}^{\ddagger}/k_B T}$ with τ_o being the time scale in the absence of an activation barrier, and we have used $\Delta F_{UN}^o - f_m \Delta x_{UN} = 0$. The fits using the two parameters $\tau_F(f_Q, f_S)(\equiv \tau(0) e^{-\delta f_Q \Delta x_{U \rightarrow N}^{\ddagger}/k_B T})$ and Δx_{UN}, provides an excellent description of the simulation data (Fig. 1C). The analytic result in Eq. **1** also shows the inherent asymmetry in the two modes for initiating folding by force quench.

Nonequilibrium Force-Quench Folding Is Reflected in $R(t)$. The dramatic variations in the refolding time $\tau_F(f_Q, f_S)$ (Fig. 1B and C) are reflected in the folding trajectories, $R(t)$, as f_S and f_Q are varied. In Fig. 1D (see also Fig. 2), which shows the ensemble of force-quench trajectories with $f_S = 70$ pN and $f_Q = 3.5$ pN, we observe a rapid decrease in $R(t)$ in the early stages of folding (stage 1), followed by a slowly decaying plateau of varying length (stage 2). The second stage is followed by an abrupt reduction to $R \lesssim 2$ nm, corresponding to the hairpin formation (6, 13, 17, 18) (stage 3). The three stages of folding are well separated temporally, and the plateau length increases with increasing f_Q. Although multistage folding has been discussed in the context of T-jump protein folding (5, 19, 20), the dramatic manifestation of time scale separation at an "early" stage of protein folding is unique to force-quench experiments. The nature of FIMI is different from that of a thermally collapsed intermediate state in that the FIMI state contains no intramolecular contacts (see Fig. 2A and discussion of P_{TDE} below). The rapid increase in the folding times (Fig. 1B and 1D) is directly linked to the long-lived FIMIs. The plateau persists until the system reaches the transition state, $R \approx R_{TS} = 4$ nm, where R_{TS} is the position of transition barrier at $f = f_m$ (Fig. S3 of the *SI Appendix*).

The process by which RNA navigates the folding landscape in stage 2 of the collapse process, during which $R(t)$ is nearly a constant (Fig. 1D), is assessed by monitoring the conformational changes in the hairpin. In order to illustrate these structural changes, we computed the distributions, $P_{ERE}(R)$ (ERE stands for entropically relaxed ensemble) "immediately after" the force-quench (between stages 1 and 2), and $P_{TSE}(R)$ (TSE denotes the transition-state ensemble) that gives the distribution function "immediately before" the final decrease of $R(t)$ (between stages 2 and 3; see the *SI Appendix* for further discussion). The differences between $P_{ERE}(R)$ and $P_{TSE}(R)$ are illustrated in Fig. 2. As long as $f_S \gg f_m$, $P_{ERE}(R)$ consists of a set of homogeneously stretched structures, whereas the structural ensemble of hairpin loops with disordered ends become dominant for $f_S \lesssim f_m$ (see Fig. 2 A–D). For $f_S = 56$ pN, $P_{TSE}(R)$ is peaked at smaller values of $R \approx (6 - 7)$ nm, and becomes broader with decreasing f_Q (see Fig. 2 A, C, and E). The ensemble of structures that are sampled in this second stage are the FIMIs, which are analogues of minimum-energy compact structures, postulated by Camacho and Thirumalai (21) for folding initiated by temperature quench. A comparison of $P(R)$ in Fig. 2 A, C, and E with $P_{TDE}(R)$ (TDE denotes the thermally denatured ensemble; see the pink distribution in Fig. 2A) shows

that the conformations in the FIMI are unlikely to be sampled during a typical refolding initiated by temperature quench (17).

At $f_S = 14$ pN $\lesssim f_m$ (Fig. 2 B and D), $P_{ERE}(R)$ differs substantially from the distributions for $f_S = 56$ pN $\gg f_m$ (Fig. 2 A, C, and E). The ERE has an approximately bimodal distribution, similar to the initial distribution $P_S(R)$ at $f = f_m$ (10, 22) (compare $P_{ERE}(R)$ in Fig. 2 B and D with Fig. S1 of the *SI Appendix* at $f_S = 14$ pN) and is far more heterogeneous than the ensemble at $f_S = 56$ pN. With $f_S \approx f_m$, the folding is initiated predominantly from molecules with a partially formed hairpin loop. In this case, the molecule refolds directly to the native state without a well-defined plateau in $R(t)$. A plateau in $R(t)$ will be manifested most clearly only when the force is quenched from a large $f_S(\gg f_m)$ and $f_Q(\lesssim f_m)$. The simulations also suggest that the lifetimes of the FIMIs depend on the protocol used to quench the force, which is determined by λ (see Table 1).

The evolution from the ERE to the TSE, which occurs over a time scale determined by the length of the plateau, is associated with a search process leading to a nucleation step in the hairpin formation. We have shown previously, in a force-quench simulation with $f_Q = 0$, that the dynamics of the plateau is associated with a search for the native dihedral angles (trans → gauche transition) in the tetra-loop region (13, 17). However, the plateau formation in $R(t)$ is a generic feature of force-quench dynamics and has also been observed in recent experiments on polyubiquitin, I27, and the PEVK protein from titin (6, 23). Interestingly, the formation of long-lived FIMIs also occurs in force-quench dynamics of semiflexible chains and flexible homopolymers (see *Discussion*).

Link Between Entropy Production and the Collapse Dynamics. The transition from the stretched state to the folded state that occurs through FIMIs can be described by using time-dependent changes in the entropy. From the ensemble of force-quench time traces, $R(t)$, we estimate the entropy of the molecule as a function of time by using

$$S(t)/k_B = -\int_0^\infty dR\, P(R,t) \log P(R,t), \qquad [2]$$

where $P(R,t)$ is the end-to-end distribution function obtained directly from the ensemble of trajectories. At long times, $P(R, t \rightarrow \infty)$ will reduce to the distribution for a given system. The time-dependent entropy is computed directly from the data generated from experiments or simulations, giving this approach a practical use, and thus captures the nonequilibrium behavior in all force regimes and over all time scales. Note that Eq. **2** makes no assumptions about a particular model for $P(R,t)$ but is strictly valid only if R can adequately describe the conformation of the molecule of interest. Fig. 3 shows $S(t)$ computed for four different pairs of (f_S, f_Q). Interestingly, changes in $S(t)$ are nonmonotonic when the system begins from a high stretch force ($f_S = 56$ pN, Fig. 3A), whereas the $S(t)$ decreases monotonically upon force-quench from
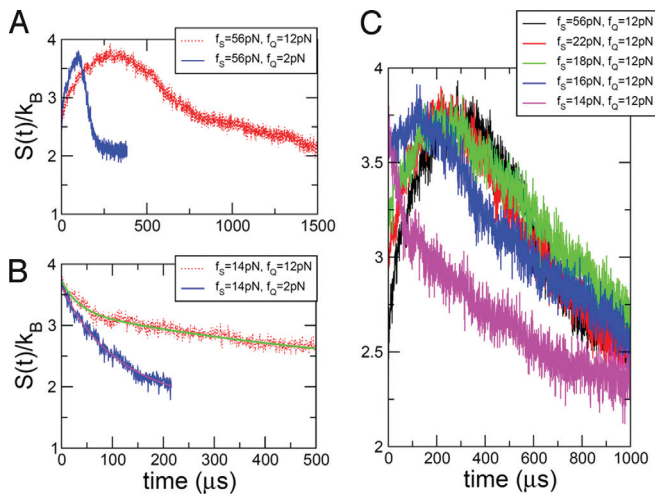
Fig. 3. The time evolution of the entropy, measured by using Eq. 2. (*A*) $S(t)/k_B$ from $f_S = 56$ pN to $f_Q = 12$ pN (red) and 2 pN (blue). (*B*) $S(t)/k_B$ from $f_S = 14$ pN to $f_Q = 12$ pN (red) and 2 pN (blue). (*C*) The time evolution of $S(t)/k_B$ for varying f_S with fixed $f_Q = 12$ pN. In this frame, we take the running average of $S(t)/k_B$ every 4.7 μs.

$f_S = 14$ pN (Fig. 3*B*). The quantitative analysis using $S(t)$ clarifies the dynamics in the ensemble of time traces (Fig. 2). The evolution of the end-to-end distribution, $P_S(R) \rightarrow P_{ERE}(R) \rightarrow P_{TSE}(R) \rightarrow P_{NBA}(R)$, is due to changes in the dynamic evolution of the entropy (Eq. 2) which progresses from a low (stretched) $\rightarrow$ medium (ERE) $\rightarrow$ high (FIMIs, TSE) $\rightarrow$ low (NBA) value as the hairpin folds, especially when the initial ensemble of states is fully stretched ($R \gg 4$ nm in Fig. S2 of the *SI Appendix*). At $f_S = 14$ pN, $P_S(R)$ is bimodal and $S(t)$ with a maximum value at $t = 0$, and decreases monotonically as $S(t)/k_B \sim \Phi e^{-t/\tau_f} + (1-\Phi)e^{-t/\tau_s}$ where Φ, τ_f and τ_s are f_Q-dependent for the same value of f_S. The time constants τ_f and τ_s correspond to collapse of fast- and slow-track molecules in the multidimensional free energy landscape (with $\tau_f < \tau_s$), and Φ is the fraction of fast collapsing molecules (1). The fast and slow tracks are associated with the dynamics of collapse, starting from within the NBA ($R < 4$ nm) and UBA ($R > 4$ nm), respectively. We find $\Phi = 0.29$, $\tau_f = 33.8$ μs, $\tau_s = 715$ μs for $f_Q = 12$ pN, and $\Phi = 0.47$, $\tau_f = 81.3$ μs, $\tau_s = 181$ μs for $f_Q = 2$ pN. A smaller value of f_Q increases the fraction of conformations in the folding route starting from NBA to the native state. The force-quench dynamics from the force-denatured ensemble with $f_S = 14$ pN is an example of the kinetic partitioning mechanism (1) and shows that the mechanisms of hairpin formation can be surprisingly complex (8).

The transition from the monotonic to nonmonotonic decay of $S(t)$ occurs at $f_S \approx 16$ pN (Fig. 3*C*), and the time scales for $S(t)$ to attain its maximum value are similar for $f_S = 18{-}56$ pN. However, the entropies of the initial structural ensemble, $S(0)$, are different and decrease as f_S increases. This finding provides insights into why τ_F saturates for $f_S \geq 18$ pN in Fig. 1*C*, because the P5GA hairpin is predominantly in the UBA for $f_S > f_m + \delta f_{1/2} = 14.7 + 1.2$ pN ≈ 16 pN, where $\delta f_{1/2}$ is the half-width of force fluctuations around f_m (see *SI Appendix* for details).

The Expanding Sausage Model Under Tension Captures the Physics of Force-Quench Refolding of Biopolymers. Although the nonequilibrium behavior of the hairpin, reflected in $R(t)$ and the time scales for hairpin formation, is shown for P5GA (Figs. 1 and 2), the dynamics of contraction of the molecule upon rapid quench should be a generic phenomenon, provided λ is small and a driving force leads to the development of a compact structure for sufficiently low f_Q. In order to illustrate the generality of the FIMI that gives rises to the second stage of contraction (Figs. 1 and 2 and ref. 6), we provide a theoretical description by adapting

de Gennes' model (24) for collapse kinetics of a flexible coil. By using an "expanding sausage model," de Gennes envisioned that the collapse (upon temperature quench to $T = T_\Theta - \Delta T < T_\Theta$, the Flory Θ-temperature) progresses from a randomly distributed string of thermal blobs of size $\xi(= ag^{1/2})$ (where $g[= (T_\Theta/\Delta T)^2]$ is the number of monomer constituting each blob), to a sausage-like collapsed structure. The driving force in this picture (see Fig. S4*A*) is the minimization of the surface tension in a poor solvent ($T < T_\Theta$).

To describe the second stage in the force-quench dynamics (or more generally for any biopolymer that undergoes a stretch $\rightarrow$ collapse transition) using the generalization of the expanding sausage model, we write the free energy under tension as

$$F(R) = 2\frac{k_B T}{\xi^2}\sqrt{\pi\Omega}R^{1/2} - f_Q R$$
$$+ k_B T \frac{L}{l_p} \int_0^{R/L} dx \left[\frac{1}{4(1-x)^2} - \frac{1}{4} + x \right], \quad [3]$$

where the first term is the driving force for collapse ($F_C = \gamma A$ where A is the surface area of a thermal blob). We rewrite $F_C \sim 2k_B T(\pi\Omega R)^{1/2}/\xi^2$, where the surface tension for a thermal blob is $\gamma \approx k_B T/\xi^2$, and Ω is the volume of the sausage, assumed to be constant throughout the collapse process (24). The second term in Eq. 3, $F_f = -f_Q R$, accounts for the mechanical work. The third term in Eq. 3, which enforces chain inextensibility under the applied tension, is required to incorporate the sharp increase of mechanical force at $R \approx L$, observed in the P5GA hairpin (see Fig. S1*B*). A simplified version of the sausage model, where inextensibility is not taken into account, is presented in the *SI Appendix*. The details of the free energy (Eq. 3) are not expected to change the qualitative picture of the collapse kinetics that emerges when using the expanding sausage model to describe force-quench dynamics as long as each term is taken into account. The effect of the stretch force f_S on the folding kinetics is introduced through the initial equilibrium extension R_0, which for $f_S \gg k_B T/l_p (\equiv f_p)$ is $R_0(f_S) \approx L(1 - \sqrt{f_p/4f_S}) \lesssim L$.

We equate the rate of free energy reduction of the system, $\dot{F}(R)$, with the entropy production due to dissipation, $T\dot{S}(R) \approx \eta \times R \times \dot{R}^2$ (24), where we have used the Stokes–Einstein relation and a damping force $f \approx -\zeta v \approx -\eta R v$. The friction coefficient $\zeta \approx \eta R$, where R is the size of collapsed polymer in the longitudinal direction and $v(\approx \dot{R})$, is the velocity. By equating $\dot{F}(R) \approx -T\dot{S}(R)$, we get

$$t/\tau_c = \int_{R/L}^{R_0/L} \frac{x\,dx}{x^{-1/2} - \bar{f}_Q + \bar{f}_p \left(\frac{1}{4(1-x)^2} - \frac{1}{4} + x \right)}, \quad [4]$$

where $R_0/L \approx 1 - \sqrt{f_p/4f_S}$ for $f_S \gg f_p$ and the forces are scaled as $\bar{f}_\alpha = f_\alpha \tau_c/\eta L^2$ for $\alpha = S, Q$, and p, and with $\tau_c \equiv \frac{\eta \xi^2 L^{5/2}}{k_B T \sqrt{\pi\Omega}}$. The ratio $R(t)/R_0 \equiv \Sigma(t)$, which is interpreted as the survival probability of the unfolded state upon force quench f_Q, can be determined by numerically integrating Eq. 4 (see the left columns of Figs. 4 *B*–*D*). The average folding time (or mean first passage time) is $\tau_F(\{f_\alpha\}) = \langle \tau \rangle_{\{f_\alpha\}} = \int_0^{\tau^*} dt\, \Sigma(t; \{f_\alpha\})$, where τ^* satisfies $R(\tau^*) = 0$. The results of τ_F for various combinations of $(\bar{f}_Q, \bar{f}_S, \bar{f}_p)$ are given in the right columns of Fig. 4 *B*–*D*. In agreement with our simulation results (see also refs. 13, 17, and 18), as well as force-quench experiments on poly-Ub (6), the sausage model for the worm-like chain (WLC) shows a sharp decrease in $R(t)$ at an early stage of the force-quench dynamics, after which a plateau in R of variable length develops, before an abrupt collapse to $R = 0$ (see Fig. 2). A long time plateau is observed for a stiffer biopolymer (small f_p) and for larger f_Q (Fig. 4 *B* and *D*). In agreement with the more-detailed P5GA simulations, τ_F increases exponentially at small quench forces ($\bar{f}_Q \lesssim 1$) with a slight upward curvature for larger f_Q (compare Fig. 4*B*).
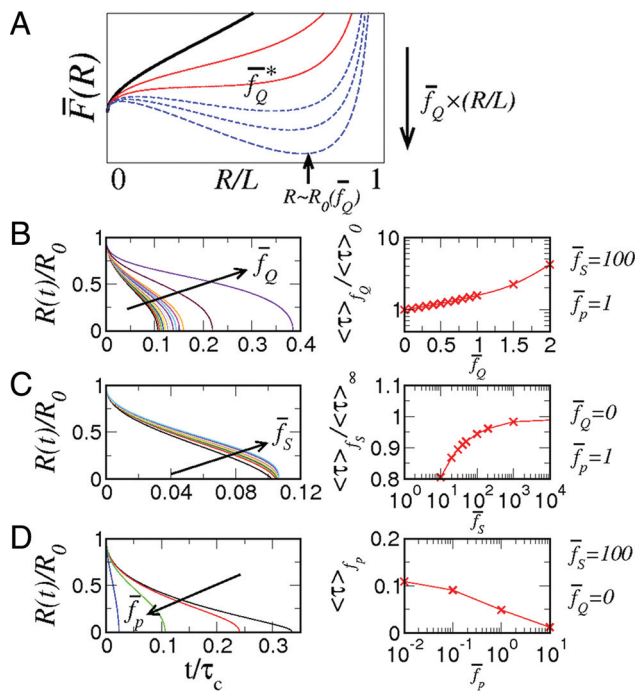
Fig. 4. The expanding sausage model for force-quench dynamics. (*A*) The free energy profile as a function of R for increasing tension. The free energy profile $F(R)$ is plotted in dimensionless form by using $\bar{F}(R/L) = \frac{F\tau_C}{\eta L^3} = 2(\frac{R}{L})^{1/2} + \bar{f}_P \frac{R^2}{4L^2}(\frac{3-2R/L}{1-R/L}) - \bar{f}_Q(\frac{R}{L})$. Two stable, free energy minima occur at $R = 0$ and $R = R_0(f_Q) \lesssim L$ when $\bar{f}_Q > \bar{f}_Q^*$. The profiles show that R^{TS} decreases and R_0 increases with increasing f. In *B–D*, the dynamics of $R(t)$ (left) and the average refolding time $\langle\tau\rangle = \tau_F$ (right) are shown for varying $\bar{f}_Q$, $\bar{f}_S$, and chain flexibility ($\bar{f}_P = k_BT/l_p$), respectively.

The dependence of τ_F on the stretch force f_S is typically small and saturates when $\bar{f}_S \gtrsim 10^3$ (compare Figs. 1*C* and 4*C*).

A word of caution is appropriate. Eq. **4**, which is derived in the absence of fluctuating force (noise term in the Langevin equation), is valid only for $\bar{f}_Q < \bar{f}_Q^*$, where $\bar{f}_Q^*$ is the force beyond which a finite free energy barrier for refolding develops (the dotted energy profiles in Fig. 4*A*). Above $\bar{f}_Q = \bar{f}_Q^*$, there is a barrier to refolding, $\delta\bar{F}^{\ddagger}$, and a valid solution to Eq. **4** cannot be found for all $t > 0$. De Gennes' approach (24) is not applicable for rate processes (25) that involve crossing a free energy barrier. Physically, the residence time in the FIMI state becomes infinite in the absence of noise. The analog of $\bar{f}_Q^*$ was referred to as "spinodal" or as a "coercive field" in ref. 26 in which the nonequilibrium relaxation of spin systems was studied. Therefore, to account for the force-quench dynamics when $\bar{f}_Q > \bar{f}_Q^*$, we use Kramers' theory for barrier crossing, assuming that R (under tension) is a good reaction coordinate, an approach that is consistent with other estimations of the transition time (27, 28). In so doing, it is assumed that under tension ($f_Q \neq 0$) that results in the suppression of internal fluctuations, R is a good reaction coordinate. The refolding time τ_F scales with the effective height of the free energy barrier as $\tau_F \sim \exp(\delta F^{\ddagger})$ where $\delta\bar{F}^{\ddagger} = \bar{F}(R^{TS}/L) - \bar{F}(R_0/L)$. For $\bar{f}_Q \gg \bar{f}_Q^*$, one can show (see *SI Appendix* for details)

$$\tau_F \sim \exp\left[\left(\frac{1-2\bar{f}_P^2}{\bar{f}_Q} - 2 + \bar{f}_Q\right) + \left(2 - \bar{f}_Q + \frac{3}{2}\bar{f}_P\right)\sqrt{\frac{\bar{f}_P}{\bar{f}_Q}}\right].$$
[5]

For small $\bar{f}_Q$ or increasing flexibility ($\bar{f}_P \gg 1$), τ_F in Eq. **5** deviates significantly from Bell's expression [$\tau_F \sim \exp(\bar{f}_Q)$] (13–15). We stress that Eq. **5** is only approximate and is shown to emphasize that activated transitions (in contrast to Eq. **4**) dominate the approach to the ordered state.

Although additional factors, such as topological constraints due to self-avoidance (29) or other interactions, may alter the dynamic characteristics (for instance, the dependence of τ_F on f_S or f_Q), the generic feature of the emergence of a FIMI state upon force quench, also evidenced in the force-quench collapse dynamics of polymer chain under poor solvent conditions (see *Discussion*), is satisfactorily captured by the simple model.

Discussion

Implications for Experiments. Our finding that the very nature of folding can be altered by the depth of quench as well as the rate of quench readily explains the different conclusions reached in the refolding of Ub (7). As shown in Figs. 1 and 2, the lifetime of the FIMIs, leading to the plateau in $R(t)$, arises only if the ratio $\lambda = \tau_Q/\tau_F(f_Q, f_S)$ is small. Under the experimental conditions in ref. 6, $\lambda \approx 0.15$, whereas in the more recent AFM experiments, $\lambda \approx 32$ (7) (see Table 1 for details). Consequently, the experiments by Schlierf and Rief (7) are closer to equilibrium conditions than the force-quench refolding reported in ref. 6. We predict that the apparent differences between the conclusions reached in refs. 6 and 7 are solely due to differing values of λ. The generality of these findings can be further appreciated from the study of Li et al. on TAR RNA ($\lambda = 0.1$, see Table 1), which showed evidence for the development of a plateau (see figure 3*A* in ref. 8), although its importance was not discussed. Our findings also imply that only by varying λ over a wide range is it possible to discern the multiple stages in folding and firmly establish a link between collapse and folding transitions.

FIMIs Are a Generic Feature of Force-Quench Dynamics. The sharp transitions in $R(t)$, from its value in the second stage to the native compact structure in the third stage, is reminiscent of a weak first-order phase transition. Analyses of $P_{TSE}(R)$ for P5GA (Fig. 2*A*) show that the formation of the hairpin loop from a FIMI is the key nucleation event in bringing the stretched RNA into its native-like hairpin structure. Because the simulations on RNA hairpins and the expanding sausage model suggest the long plateau only depends on the characteristic time scales associated with the folding polymer and the experimentally controlled variables in force quench, a similar behavior should be observed even in polymers that do not have a unique native structure but can undergo a transition to a globular state. To this end, we consider the dynamics of $R(t)$ in homopolymers (flexible and worm-like chains) in a poor solvent upon force quench. The simulation details are in the *SI Appendix*.

Collapse of semiflexible chain. In the absence of force, the thermodynamically stable state of a semiflexible chain in a poor solvent is a toroid (see Fig. 5*A*). We consider the collapse of a stiff worm-like chain (bending rigidity $k_b = 80k_BT$; see Eq. **8** of the *SI Appendix*) that is initially equilibrated at $f = f_S = 83$ pN, and the force is suddenly quenched to $f_Q = 4$ pN. From the time traces of $R(t)$ in Fig. 5*A*, we surmise that the transition to the toroidal structure occurs in three stages, just as for the P5GA hairpin and Ub (6). Within our simulation times ($\lesssim$ 170 μs), some molecules adopt toroidal (magenta) or racquet structures (green), whereas others (brown) fluctuate around large values of R without adopting any ordered structure. These findings for the formation of toroids for the WLC are in quantitative accord with experiments on the dynamics of force-quench collapse of λ−DNA, which showed that toroid formation occurs in a stepwise, quantized manner (30) and is in agreement with the forced unfolding of the lattice model of he proteins (31).

Fig. 5. Polymer collapse in a poor solvent upon force quench. (*A*) The collapse of a WLC in a poor solvent. (*B*) Toroidal, single racquet, or multiple racquet structures are formed in various trajectories. (*C*) The entropy production using Eq. **2**, displaying the low → high → low pathway from high f_S to low f_Q. (*D*) Linear reduction in force over a period, τ_Q, during which the free energy gradually changes from a profile with $f > f_m$ to one with $f < f_m$. Depending on the time scale for hopping between the native and unfolded basins of attraction, τ_{hop}, in comparison with τ_Q, the pattern of relaxation dynamics upon force quench can be greatly affected.

The entropy production, $S(t)$, of a collapsing WLC is shown in Fig. 5*B* only for those molecules that form an ordered structure (the trajectories grouped in the red curly bracket in Fig. 5*A*; see also Fig. S5 in the *SI Appendix* for complete set of trajectories). The low entropy of the initial, highly stretched state increases until it attains a maximum value, corresponding to an ensemble dominated by the FIMIs, after which the entropy slowly decays to a value in the more ordered, low-entropy state. This finding is in agreement with the three-stage collapse dynamics of the simpler RNA hairpin with a well-defined folded structure.

Collapse of flexible polymers in a poor solvent. To further demonstrate that bending rigidity in the semiflexible chain plays no special role in leading to FIMIs, we have considered the stretch → globule transition of a flexible hompolymer in a poor solvent under force-quench conditions (see *SI Appendix* text for details). The final globular structure, in the absence of force, forms continuously from the ends of the chain (a pearl-necklace ensemble),

with no evidence of a plateau in the collapse (see Fig. S6*A* of the *SI Appendix*). However, the continuous collapse transition becomes cooperative and weakly first order by choosing a large quench force $f_Q = 75$ pN (Fig. S6*B*), due to the formation of FIMIs.

These two examples for homopolymers and the results for RNA hairpin show that the value of λ, and not the specific architecture of the folded structure, dictate the collapse in the intermediate stages. The expected changes in $R(t)$ of a biomolecule (or a homopolymer in a poor solvent) upon a rapid $f_S \rightarrow f_Q$ quench (small λ) must occur in multiple stages (Fig. S5), with stages 1 and 2 being determined by entropy growth (stage 1). The search for the folding nuclei (32) that can further drive structure formation (stage 3) and sequence-dependent effects arises in the final stages of folding (Fig. 5*C*).

If the force is quenched rapidly, the biomolecule will be far from equilibrium when the first nucleation event occurs, which will drastically change the dynamics of the folding process. Folding occurs in a near-equilibrium manner either if the force is relaxed slowly (Fig. 5*D*), so that the internal modes of the biomolecule equilibrate, or if the quench force is sufficiently large that there is a significant barrier to folding. For a slow decrease in the force, the relevant time scale is $\tau_Q = \Delta L / v_L$, whereas the relevant time scale for the barrier crossing event will be $\tau_{hop}(f_Q)$ (Fig. 5*D*). Thus, near folding occurs at near equilibrium for $\lambda = \tau_Q / \tau_{hop} \gg 1$, whereas nonequilibrium folding occurs in the opposite limit. Variations in the experimental protocol have to be taken into account in elucidating the pathways explored during the folding process.

Methods

The P5GA hairpin, as represented by using the self-organized polymer model (see ref. 22 for the energy function and simulation method), is a two-state folder with a transition midforce of $f_m = 14.7$ pN (10). We probe the refolding kinetics for multiple pairs of initial and quench forces, (f_S,f_Q), whose magnitudes range from $14 \leq f_S \leq 70$ pN and $2 \leq f_Q \leq 12$ pN (see Fig. 1*A*). The force-quench refolding kinetics for 100 molecules were simulated beginning with an initial tension (f_S) and quenched to a final tension (f_Q). The refolding time of the ith molecule, $\tau_F(i)$, was defined as the first time the hairpin gained more than 95% of native contacts after the tension was reduced to $f = f_Q$.

1. Thirumalai D, Hyeon C (2005) RNA and protein folding: Common themes and variations. *Biochemistry* 44:4957–4970.
2. Onuchic JN, Wolynes PG (2004) Theory of protein folding. *Curr Opin Struct Biol* 14:70–75.
3. Dill KA, Ozkan SB, Shell MS, Weikl TR (2008) The protein folding problem. *Annu Rev Biophys* 37:289–316.
4. Shakhnovich E (2006) Protein folding thermodynamics and dynamics: Where physics, chemistry, and biology meet. *Chem Rev* 106:1559–1588.
5. Thirumalai D (1995) From minimal models to real proteins: Time scales for protein folding kinetics. *J Physique I* 5:1457–1467.
6. Fernandez JM, Li H (2004) Force-clamp spectroscopy monitors the folding trajectory of a single protein. *Science* 303:1674–1678.
7. Schlierf M, Rief M (2009) Surprising simplicity in the single-molecule folding mechanics of proteins. *Angew Chem Int Ed* 48:820–822.
8. Li PTX, Collin D, Smith SB, Bustamante C, Tinoco I Jr (2006) Probing the mechanical folding kinetics of TAR RNA by hopping, force-jump, and force-ramp methods. *Biophys J* 90:250–260.
9. Li PTX, Bustamante C, Tinoco I Jr (2007) Real-time control of the energy landscape by force directs the folding of RNA molecules. *Proc Natl Acad Sci USA* 104:7039–7044.
10. Hyeon C, Morrison G, Thirumalai D (2008) Force dependent hopping rates of RNA hairpins can be estimated from accurate measurement of the folding landscapes. *Proc Natl Acad Sci USA* 105:9604–9606.
11. Bell GI (1978) Models for the specific adhesion of cells to cells. *Science* 200:618–627.
12. Li MS, Hu CK, Klimov DK, Thirumalai D (2006) Multiple stepwise refolding of immunoglobulin I27 upon force quench depends on initial conditions. *Proc Natl Acad Sci USA* 103:93–98.
13. Hyeon C, Thirumalai D (2006) Forced-unfolding and force-quench refolding of RNA hairpins. *Biophys J* 90:3410–3427.
14. Hyeon C, Thirumalai D (2007) Measuring the energy landscape roughness and the transition state location of biomolecules using single molecule mechanical unfolding experiments. *J Phys Condens Matter* 19:113101.
15. Dudko OK, Hummer G, Szabo A (2006) Intrinsic rates and activation free energies from single-molecule pulling experiments. *Phys Rev Lett* 96:108101.
16. Dudko OK, Hummer G, Szabo A (2008) Theory, analysis, and interpretation of single-molecule force spectroscopy experiments. *Proc Natl Acad Sci USA* 105:15755–15760.
17. Hyeon C, Thirumalai D (2008) Multiple probes are required to explore and control the rugged energy landscape of RNA hairpins. *J Am Chem Soc* 130:1538–1539.
18. Hyeon C, Thirumalai D (2005) Mechanical unfolding of RNA hairpins. *Proc Natl Acad Sci USA* 102:6789–6794.
19. Sali A, Shakhnovich E, Karplus M (1994) How does a protein fold? *Nature* 369:248–251.
20. Camacho C, Thirumalai D (1993) Kinetics and thermodynamics of folding in model proteins. *Proc Natl Acad Sci USA* 90:6369–6372.
21. Camacho C, Thirumalai D (1993) Minimum energy compact structures of random sequences of heteropolymers. *PhyS Rev Lett* 71:2505–2508.
22. Hyeon C, Thirumalai D (2007) Mechanical unfolding of RNA: From hairpins to structures with internal multiloops. *Biophys J* 92:731–743.
23. Walther KA, et al. (2007) Signatures of hydrophobic collapse in extended proteins captured with force spectroscopy. *Proc Natl Acad Sci USA* 104:7916–7921.
24. de Gennes PG (1985) Kinetics of collapse of a flexible coil. *J Physique Lett* 46:L639–L642.
25. Chandrasekhar S (1943) Stochastic problems in physics and astronomy. *Rev Mod Phys* 15:1–89.
26. Binder K (1973) Time-dependent Ginzburg–Landau theory of nonequilibrium relaxation. *Phys Rev B* 8:3423–3438.
27. Yoshinaga N (2008) Folding and unfolding kinetics of a single semiflexible polymer. *Phys Rev E* 77:061805.
28. Halperin A, Goldbart PM (2000) Early stage of homopolymer collapse. *Phys Rev E* 61:565–573.
29. Grosberg AY, Nechaev SK, Shakhnovich EI (1988) The role of topological constraints in the kinetics of collapse of macromolecules. *J Phys* 49:2095–2100.
30. Fu WB, et al. (2006) Compaction dynamics of single DNA molecule under tension. *J Am Chem Soc* 128:15040–15041.
31. Klimov DK, Thirumalai D (1999) Stretching single-domain proteins: Phase diagram and kinetics of force-induced unfolding. *Proc Natl Acad Sci USA* 96:6166–6170.
32. Guo Z, Thirumalai D (1997) The nucleation-collapse mechanism in protein folding: Evidence for the non-uniqueness of the folding nucleus. *Folding & Design* 2:277–341.
33. Cecconi C, Shank EA, Bustamante C, Marqusee S (2005) Direct observation of three-state folding of a single protein molecule. *Science* 309:2057–2060.

PHYSICS

BIOPHYSICS AND COMPUTATIONAL BIOLOGY