

Complex Unfolding Kinetics of Single-Domain Proteins in the Presence of Force

Michael Schlierf,^{†△} Zu Thur Yew,^{‡△} Matthias Rief,[¶] and Emanuele Paci^{‡§*}

[†]Loomis Laboratory, University of Illinois at Urbana-Champaign, Champaign, Illinois; [‡]Institute of Molecular and Cellular Biology, [§]School of Physics and Astronomy, University of Leeds, Leeds, United Kingdom; and [¶]Physik Department, Technical University of Munich, Munich, Germany

ABSTRACT Single-molecule force spectroscopy is providing unique, and sometimes unexpected, insights into the free-energy landscapes of proteins. Despite the complexity of the free-energy landscapes revealed by mechanical probes, forced unfolding experiments are often analyzed using one-dimensional models that predict a logarithmic dependence of the unfolding force on the pulling velocity. We previously found that the unfolding force of the protein filamin at low pulling speed did not decrease logarithmically with the pulling speed. Here we present results from a large number of unfolding simulations of a coarse-grain model of the protein filamin under a broad range of constant forces. These show that a two-path model is physically plausible and produces a deviation from the behavior predicted by one-dimensional models analogous to that observed experimentally. We also show that the analysis of the distributions of unfolding forces ($p[F]$) contains crucial and exploitable information, and that a proper description of the unfolding of single-domain proteins needs to account for the intrinsic multidimensionality of the underlying free-energy landscape, especially when the applied perturbation is small.

INTRODUCTION

Over the past decade, single-molecule atomic force microscopy studies have revealed the complexity of the folding and unfolding of single-domain proteins. Examples include the observation of parallel unfolding pathways (1,2), unfolding intermediates (3), and compact collapsed structures along the folding pathway (4). Recently, the mechanical unfolding kinetics of the native to intermediate ($N \rightarrow I$) transition of the protein filamin was characterized in a broad range of pulling velocities (1–4000 nm/s), and an unexpected plateau in the modal (i.e., most probable) unfolding force, F^* , was observed at low pulling velocities/loading rates (Fig. 1) (5). The convex curvature in the plot of F^* as a function of $\ln[R_{F^*}]$ represents a sharp departure from the behavior predominantly observed for simple two-state proteins where F^* is generally assumed to vary linearly with the logarithm of the loading rate, $\ln[R_{F^*}]$ (6). The linear relationship between F^* and $\ln[R_{F^*}]$ relies on the validity of the irreversible two-state Bell model (7), which states that the unfolding rate decreases exponentially with the applied force,

$$k_u(F) = k_u(0)\exp(-Fx_u/k_B T),$$

where $k_u(0)$ is the unfolding rate in the absence of force and x_u is the force dependence of the unfolding barrier. However, if x_u changes with the force (8–11) or the refolding reaction is nonnegligible (12), the dependence of F^* on $\ln[R_{F^*}]$ will no longer be linear. A close inspection of the force traces of filamin showed no evidence of refolding or additional unfolding intermediates at low pulling velocities (5). In addition, using a two-state model that accounts for the force-dependence of x_u (10) did not result in better agreement with the experimental data (5). Although anomalous curvatures are not uncommon in the mechanical unbinding of receptor-ligand complexes (13–16), to our knowledge they have never been directly observed for the unfolding of single-domain proteins. One reason could be the highly cooperative unfolding kinetics of small, single-domain proteins. Previous simulation (17–19) and experimental (20) studies, however, have suggested that such anomalous kinetic signatures can, in principle, be observed for single-domain proteins.

The convex curvature is typically explained by the presence of two sequential barriers along the unfolding pathway of the protein (Fig. 2). At low pulling velocities, unfolding or unbinding occurs over the outer barrier, located far from the native state (Fig. 2 A), while at high pulling velocities, the outer barrier is sufficiently lowered by the force that the inner barrier becomes rate-limiting instead (Fig. 2 B). This sequential barriers (SB) model was first proposed by Merkel et al. (13), who used it to rationalize the mechanical unbinding kinetics of the streptavidin-biotin complex. If the free-energy landscape of the protein can be represented by a single reaction coordinate, the end-to-end extension, the SB model predicts that the plot of F^* as a function of $\ln[R_{F^*}]$ will consist of a series of quasilinear regimes that get increasingly steeper

Submitted April 19, 2010, and accepted for publication June 17, 2010.

[△]Michael Schlierf and Zu Thur Yew contributed equally to this work.

*Correspondence: e.paci@leeds.ac.uk

Michael Schlierf's present address is B CUBE, Molecular Bioengineering, TU Dresden, Dresden, Germany.

This is an Open Access article distributed under the terms of the Creative Commons-Attribution Noncommercial License (<http://creativecommons.org/licenses/by-nc/2.0/>), which permits unrestricted noncommercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Editor: Gerhard Hummer.

© 2010 by the Biophysical Society
0006-3495/10/09/1620/8 \$2.00

doi: 10.1016/j.bpj.2010.06.039

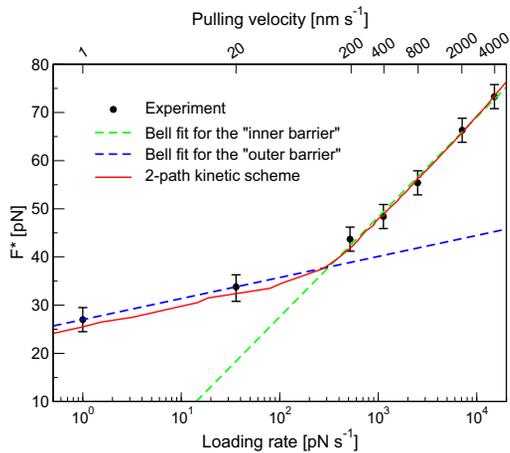


FIGURE 1 The modal unfolding force, F^* , as a function of the loading rate, R_{F^*} , calculated at each F^* . The pulling velocities at which the F^* were determined are indicated (top of the graph). The loading rates, R_{F^*} , were calculated by modeling the polymeric spacer as a wormlike chain (see Methods for details). (Error bars) Bin-width used to construct the unfolding force histograms. (Dashed lines) Piecewise fits of the data at 200–4000 nm/s (dark dashed line/red online) and at 1–20 nm/s (light dashed line/green online) to $(k_B T/x_{in/out}) \ln(R_{F^*}) + C_{in/out}$, respectively. The force-dependencies are $x_{in}^{F^*} = 4.6 \text{ \AA}$ and $x_{out}^{F^*} = 21.6 \text{ \AA}$ for the high- and low-velocity regimes, respectively. (Continuous line/red online) Fit to the kinetic scheme shown at the bottom of Fig. 5.

with the loading rate (Fig. 2, A and B). It can be shown that within each quasilinear regime (12,21)

$$F^* \approx (k_B T/x_{in/out}) \ln[R_{F^*}] + C_{in/out},$$

where $x_{in/out}$ can be thought of as the distance between the native minimum and the inner or outer barrier (Fig. 2), and $C_{in/out}$ contains information about the heights of inner or outer barriers in the absence of force. Therefore, the steepness of each the quasilinear regime is inversely related to $x_{in/out}$.

In contrast to the conventional explanation, the convex curvature observed for filamin was explained using a kinetic scheme with two unfolding pathways (see Fig. 5, denoted henceforth as the two-path kinetic scheme) (5). From Fig. 1, it can be seen that the SB model describes the R_{F^*} -dependence of F^* of filamin quite nicely (dashed lines), which begs the question: why use a more complex model when a more parsimonious one is able to adequately describe the data?

Due to the difficulty in performing single-molecule mechanical unfolding experiments, the unfolding kinetics of a protein is often studied in a narrow range of pulling velocities, and/or with limited sampling within the range of the pulling velocities investigated. In addition, the measured unfolding forces are often noisy (e.g., $\Delta F_{rms} \sim 15\text{--}25 \text{ pN}$), which makes the accurate estimation of the distribution of unfolding forces, $p(F)$, difficult. For this reason, the $p(F)$ is less commonly analyzed and an analysis of F^* is preferred instead. Analyzing only F^* , however, means that a large amount of the information (e.g., the variance and skew) contained in $p(F)$ is neglected. Theoretical

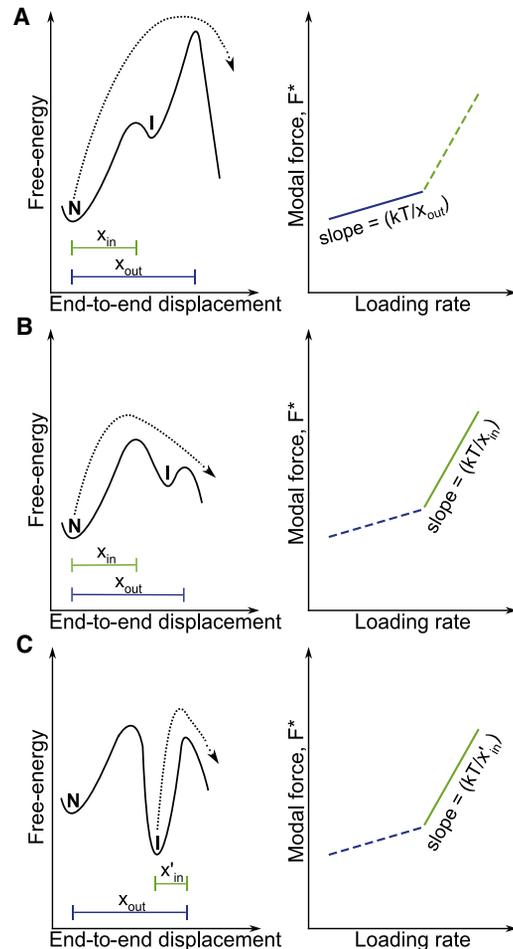


FIGURE 2 One-dimensional free-energy profiles with two sequential barriers. The reaction coordinate is typically taken to be the end-to-end extension of the protein, allowing the force-dependencies, $x_{in/out}$, to be interpreted as the difference in extension between the native and transition states. The resulting force spectrum (i.e., the F^* versus $\ln[R_{F^*}]$ plot) is depicted for each free-energy profile (right, solid line). A key assumption is that the force and end-to-end vectors do not change direction throughout the unfolding transition. (A) At low loading rates, the unfolding of the native state (N) is impeded by the outer barrier, which is located furthest away from the native state (i.e., $x_{out} \text{ \AA}$). (B) At sufficiently high loading rates, the height of the outer barrier becomes low enough such that the inner barrier, located $x_{in} \text{ \AA}$ away from the native state, becomes rate-limiting. This switch in transition barriers shows up as a sharp change in the slope of the force spectrum. (C) If the intermediate is stabilized by the force such that it becomes more stable than the native state, the outer barrier becomes rate-limiting again. In this case, the slope of the force spectrum at high loading rates will be governed by x'_{in} , the distance between the intermediate state (I) and the outer transition state. Therefore, when $x'_{in} = x_{in}$, the change in the force spectrum at high loading rates could be misattributed to the unfolding of the native state over the inner barrier.

analyses of the forced unbinding of receptor-ligand complexes (17,21) have shown that the variation of F^* with $\ln[R_{F^*}]$ is insufficient to discriminate between various kinetic models. For instance, even if the plot of F^* as a function of $\ln[R_{F^*}]$ can be described by the traditional SB model (13), the force-dependence at high pulling velocities need

not correspond to the inner barrier that is encountered when unfolding from the native state; it can be equally consistent with the unfolding of a force stabilized intermediate state over the outer barrier (Fig. 2 C) (17,21).

In this article, we present results from a large number of unfolding simulations of a coarse-grain model of the protein filamin under a broad range of constant forces. These show that a two-path model is physically plausible and produces a deviation from the behavior predicted by one-dimensional models, analogous to that observed experimentally. We also show that the analysis of the distributions of unfolding forces ($p[F]$) contains crucial and exploitable information, and that a proper description of the unfolding of single-domain proteins needs to account for the intrinsic multidimensionality of the underlying free-energy landscape, especially when the applied perturbation is small.

METHODS

Molecular dynamics simulations

All simulations were performed using a structure-based coarse-grain model (22), where each residue is represented by a C_α atom and only native interactions are attractive. The polymeric properties of the chain are governed by an amino-acid-specific dihedral potential. The topology of the model and interaction parameters were based on the lowest energy NMR structure of filamin (PDB ID: 1KSR). The dynamics of the coarse-grain model were simulated using a Langevin integrator with a friction coefficient of 1 ps^{-1} , a timestep of 15 fs, and holonomic constraints on all bonds. Constant forces ranging from 10 to 60 pN were applied to the N- and C-termini of the coarse-grain model. The simulations were performed at 330 K to ensure that good statistics were obtained at small forces (i.e.,

10–20 pN). The unfolding time was defined as the time required for the N-C distance of the coarse-grain model to be $>200 \text{ \AA}$. The average unfolding time, τ_u , was computed using at least 1000 simulations at each force.

The τ_u as a function of force was fitted to the Dudko-Hummer-Szabo (DHS) rate equation (see Eq. 4 below), the Saffman-Delbrück (SB) model, and the two-path kinetic scheme (Fig. 3 B) by nonlinear least-squares. The SB model is

$$\tau_u(F) = \tau_{\text{in}}(0)\exp[-Fx_{\text{in}}/k_B T] + \tau_{\text{out}}(0) \times \exp[-Fx_{\text{out}}/k_B T],$$

where $\tau_{\text{in/out}}(0)$ is the time needed to unfold over the inner (i.e., in) or outer (i.e., out) barrier at zero force, and $x_{\text{in/out}}$ represents the force-dependence of the height of the inner or outer barrier. For the two-path model, the unfolding time for each unfolding pathway (i.e., starting from either A1 or A2) was modeled using the Bell model (7),

$$k(0)^{-1}\exp(-Fx_u/k_B T),$$

where x_u is the force-dependence of the unfolding time. The equilibrium between the states A1 and A2 was modeled assuming linear force-dependence of the relevant free-energy, i.e.,

$$\Delta G_{\text{eq}}(F) = \Delta G_{\text{eq}}(0) - Fx_{\text{eq}},$$

where x_{eq} is the force-dependence of the equilibrium free-energy.

Fitting of the unfolding force distributions

There were 66, 34, 91, 107, 231, 149, and 199 unfolding force measurements at pulling velocities of 1–4000 nm/s, respectively. The distribution

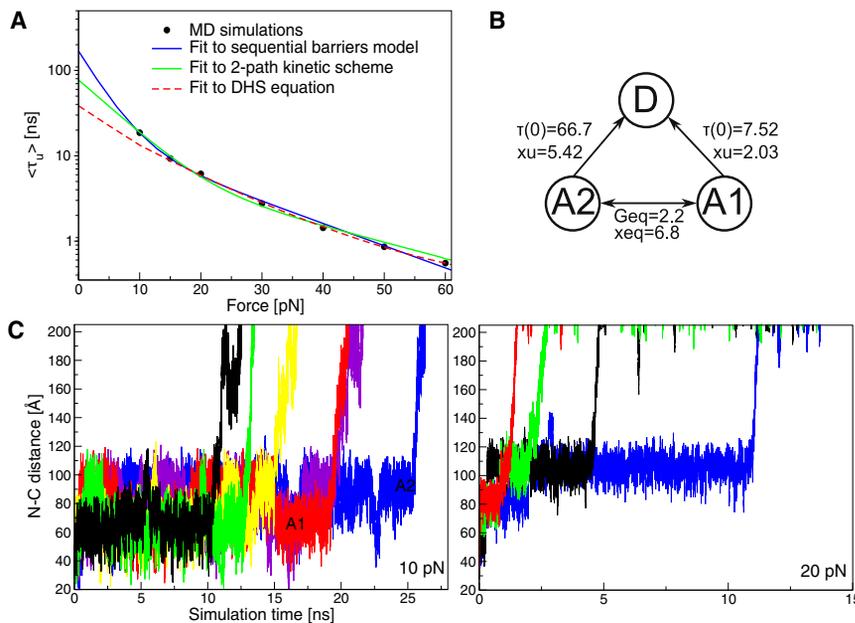


FIGURE 3 (A) The average unfolding time, τ_u , of a coarse-grain model of ddFLN4 as a function of the applied force at 330 K. (Dark-dashed red online) Fit of τ_u to the DHS equation, Eq. 4, over the entire force range. The DHS parameters are $\tau_{\text{DHS}}(0) = 38.5 \text{ ns}$, $G^\ddagger = 6.5 k_B T$, and $x_u = 4.92 \text{ \AA}$. (Dark continuous blue online) Fit to the SB model. The parameters for the inner and outer barriers are $\tau_{\text{in}}(0) = 17.7 \text{ ns}$ and $x_{\text{in}} = 2.7 \text{ \AA}$; $\tau_{\text{out}}(0) = 150.8 \text{ ns}$ and $x_{\text{out}} = 12.9 \text{ \AA}$, respectively. (Light continuous green online) Fit to the kinetic scheme shown in panel B. At forces $<30 \text{ pN}$, a significant convex curvature is observed. Please refer to Methods for details of the fitting procedure. (B) The kinetic scheme used to fit τ_u shown in panel A. The kinetic scheme was inspired by the rapid equilibrium between A1 (centered $\sim 70 \text{ \AA}$) and A2 (centered $\sim 95 \text{ \AA}$) observed in the time series seen in the time series in panels C and D. The $\tau(0)$ are in units of nanoseconds; G_{eq} and G^\ddagger are in units of $k_B T$; and x_u and x_{eq} are in units of \AA . (C) (Left) The time series of the N-C distance of the coarse-grain filamin model at 10 pN. Unfolding can be initiated either from A1 (red, black, and green online) or from A2 (yellow, blue, and purple online). (Right) The time series of the N-C distance of the coarse-grain filamin model at 20 pN. The fast switching between A1 and A2 is lost, and unfolding proceeds mainly from A2.

of unfolding forces, $p(F)$, at each pulling velocity was determined by constructing a histogram of the measured unfolding forces using a bin-width of $\sim 4\text{--}5$ pN. The bin-width was chosen according to the Freedman-Diaconis rule, i.e.,

$$\text{bin width} \approx 2[IQ]n^{-1/3},$$

where IQ is the interquartile range. The root-mean-square fluctuation in the unfolding force due to the fluctuations of the cantilever,

$$\Delta F_{\text{rms}} \approx \sqrt{k_c \times k_B T} \approx 5 \text{ pN},$$

was of similar magnitude. A total of 56 data-points (i.e., eight points per pulling velocity) was used for the fitting procedure.

To fit the $p(F)$ to the various kinetic models, we calculated the theoretical force distributions according to the Evans-Ritchie formalism (6),

$$p(F) = \frac{k_u(F)}{R(F, L, v)} \exp \left[- \int_0^F \frac{k_u(f)}{R(f, L, v)} df \right], \quad (1)$$

where F is the force, $k_u(F)$ is the unfolding rate function that describes a particular kinetic model, and $R(F, L, v)$ is the loading rate.

The loading rate, $R(F, L, v)$ was calculated by modeling the polymeric spacer as a wormlike chain (23)

$$R(F, L, v) = v \left[\frac{1}{\kappa_c} + \frac{2\beta L l_p (1 + \beta F l_p)}{3 + 5\beta F l_p + 8(\beta F l_p)^{5/2}} \right]^{-1}, \quad (2)$$

where $\beta = 1/k_B T$; v is the pulling velocity (1–4000 nm/s); $\kappa_c = 6$ pN/nm is the cantilever spring constant; $l_p = 0.5$ nm is the persistence length; and L values represent the contour lengths: 79 nm (200–4000 nm/s) (24), 98 nm (20 nm/s) (5), and 92 nm (1 nm/s) (5).

For the two-path kinetic scheme, $k_u(F)$ is given by

$$k_u(F) = \frac{K_{\text{eq}}(F)k_{A2-D}(F) + k_{A1-D}(F)}{K_{\text{eq}}(F) + 1}, \quad (3)$$

where $k_{A1-D}(F)$ and $k_{A2-D}(F)$ are the unfolding rate constants for each of the two pathways, and $K_{\text{eq}}(F)$ is the equilibrium constant between states A1 and A2.

The rate constants $k_{A1-D}(F)$ and $k_{A2-D}(F)$ were modeled using the DHS rate equation—specifically the cusp model (10),

$$k(F) = k(0) \left[1 - \frac{F x_u}{2G^\ddagger(0)} \right] \times \exp \left[\frac{G^\ddagger(0)}{k_B T} \left\{ 1 - \left(1 - \frac{F x_u}{2G^\ddagger(0)} \right)^2 \right\} \right], \quad (4)$$

where $k(0)$ is the usual Kramer's unfolding rate in the absence of force, and $G^\ddagger(0)$ and x_u are the zero force barrier height and transition state distance, respectively. We assumed

$$k(0) = 10^7 \exp[G^\ddagger(0)/k_B T],$$

thus reducing the DHS rate equation to a two-parameter model.

The equilibrium constant $K_{\text{eq}}(F)$ between the states A2 and A1 is given by

$$-k_B T \ln [K_{\text{eq}}(F)] = G_{\text{eq}}(0) - F x_{\text{eq}} - \frac{F^2}{2} [\Delta \kappa], \quad (5)$$

where

$$G_{\text{eq}}(0) = G_{A2}(0) - G_{A1}(0) \text{ and } x_{\text{eq}} = x_{A2} - x_{A1}$$

is the difference in the free-energies and extensions of A2 and A1 at zero force, respectively; and $\Delta \kappa$ is a parameter that accounts for the force-dependence of the extensions of the states (8,11).

The best-fit parameters for each kinetic model were determined by fitting the $p(F)$ from all seven pulling velocities simultaneously (unless indicated otherwise). The best-fitting parameters were determined by minimizing the sum of squared deviation from the experimental $p(F)$, and a visual inspection of the resulting fits. Parameter space was explored using a two-step procedure programmed in MATLAB (The MathWorks, Natick, MA):

Step 1. A broad search of parameter space using a genetic algorithm; and Step 2. A fine optimization of the best-fitting parameters from Step 1 using the Nelder-Mead algorithm.

To make the search of parameter space more efficient, all parameters were constrained to be in the following bounds: 1–30 $k_B T$ for free-energies, and 1–30 Å for x_u and x_{eq} .

RESULTS AND DISCUSSION

Molecular dynamics simulations as a proof of principle

The pitfalls of solely relying on the relationship between F^* and $\ln[R_{F^*}]$ to extract the features of the free-energy landscapes of proteins are illustrated in this section by molecular dynamics simulations of a coarse-grain model of filamin. In Fig. 3 A, the average unfolding time, τ_u , is shown as a function of the applied force. It is clear that the unfolding kinetics of the coarse-grain model exhibits two distinct regimes: the force-dependence of τ_u is greater at small forces (10–20 pN) than at larger forces (30–60 pN). This is analogous to the convex curvature observed for filamin experimentally (Fig. 1), because the force-dependence of τ_u and the dependence of F^* on R_{F^*} are related through the parameter

$$x_{\text{in/out}} \sim (d \ln[\tau_u] / dF)^{-1} \sim dF^* / d \ln[R_{F^*}].$$

It should be noted that the steep curvature at low forces cannot be accounted for by the DHS equation, i.e., Eq. 4 (Fig. 3 A, *dashed line/red online*). As will be clear later, the main reason for the convex curvature is the complex multistate unfolding kinetics of the filamin model. To clearly illustrate the underlying physical picture, we will use the Bell model as a simple parameterization of the rate constant in the rest of the simulation analysis.

As shown in Fig. 3 A (*dark continuous line/blue online*), the simulation results can be reproduced by the SB model. Inspection of the time series of the N-C distance at 10 pN (Fig. 3 C, *left*) and 20 pN (Fig. 3 C, *right*), however, reveal

that the SB model is not consistent with the unfolding mechanism of the coarse-grain model. The time series at 10 pN shows that before unfolding, the system equilibrates rapidly between two states, A1 and A2, centered at N-C distances of ~ 70 Å and 95 Å, respectively. Full unfolding then proceeds from either A1 or A2. At 20 pN, the fast interconversion between A1 and A2 is lost and unfolding occurs predominantly from the more extended state, A2. There are thus two parallel unfolding pathways available to the system in the small force regime (Fig. 3 C and Fig. 4). In contrast, the SB model predicts that the system will only unfold via one pathway in the small force regime (Fig. 2). Taken together, the two sets of time series suggest that the unfolding kinetics of the system can be explained by a three-state kinetic scheme with parallel unfolding pathways (Fig. 3 B). The kinetic scheme is able to reproduce the force-dependence of τ_u relatively well (Fig. 3 A, *light continuous line/green online*), but with rate parameters (i.e., $\tau_u(0)$ and x_u) that are distinctly different from those emerging from the SB model. It should be noted that without the insights provided by the simulations, the SB model would have been deemed the superior model due to 1), its parsimony—it only requires four parameters compared to six for the two-path kinetic scheme; and 2), its slightly better overall fit to τ_u .

Although a coarse-grain model of a protein was used in the simulations, the model is realistic in the sense that it captures key proteinlike features—a defined native structure, and native interactions that vary depending on the type of interactions observed in the experimentally deter-

mined structure. Thus, while the simplified model used here may not reproduce all the experimentally characterized properties of a particular protein, it is perfectly valid for looking at the general properties of proteins. In this regard, the simulations show that the curvature observed experimentally for filamin is physically possible for an Immunoglobulin-like domain, and that an alternative kinetic model can account for the convex curvature at low pulling velocities.

Unlike the SB model, there is no requirement in the two-path kinetic scheme for the unfolding pathways to be confined to a single reaction coordinate, or for unfolding to be a sequential process. The proposed kinetic scheme is thus a multidimensional generalization of the SB model that is consistent with the new view (25) of protein folding/unfolding.

Fitting the experimentally determined distribution of unfolding forces

Given a limited number of data-points in typical plots of $\ln[\tau_u]$ versus F or F^* versus $\ln[R_{F^*}]$ and the noise inherent in single-molecule measurements, the SB model is almost guaranteed to provide a good fit to any set of data consisting of two quasilinear regimes. The apparent good agreement between the data and the SB model may then lead to a potentially misleading picture of the underlying free-energy landscape. The situation, however, is not entirely dire. Using unfolding force data obtained for filamin in a broad range of pulling velocities (1–4000 nm/s) (5), we show that the distribution

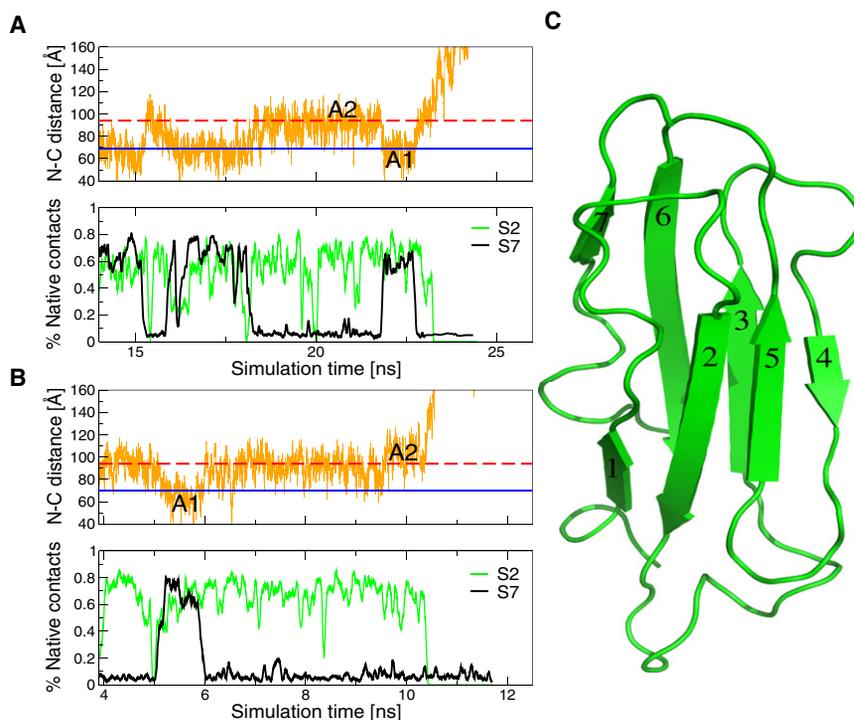


FIGURE 4 The time series of the fraction of native contacts between pairs of neighboring strands at 10 pN. (*Light/green online and dark/black online*) Lines represent the nativeness of strands 2 and 7, respectively (see panel C). Strand 1 unfolds immediately upon force application while the other strands (i.e., S3–S6) track the behavior of strand 2 or stay native ($\sim 80\%$) until full unfolding occurs. (A) A representative trajectory unfolding from state A1, characterized by the cooperative unfolding of strands 2–7. State A1 consists of conformations that have strand 1 distorted. (B) A representative trajectory unfolding from state A2. State A2 consists of conformations that have strands 1 and 7 unfolded. (C) The structure of the lowest energy NMR structure of filamin (PDB ID: 1KSR) with the strands numbered in order of the amino-acid sequence.

of unfolding forces, $p(F)$, contains additional information regarding the underlying free-energy landscape.

As mentioned in the Introduction, the plot of F^* as a function of $\ln[R_{F^*}]$ for filamin revealed a convex curvature at low pulling velocities (Fig. 1) that could be adequately described by the SB model (Fig. 1, *dashed lines*) as well as a three-state kinetic scheme with two distinct pathways (Fig. 1, *continuous line/red online*). If based solely on Fig. 1, it is difficult to argue against the SB model because it requires fewer parameters, and fits F^* slightly better at pulling velocities of 1–20 nm/s. From the fits of F^* to the SB model, the force-dependence of the outer barrier is $x^{F^*}_{\text{out}} = 21.6 \text{ \AA}$. This means that at pulling velocities of 1–20 nm/s, filamin unfolds from the native state over a barrier characterized by $x^{F^*}_{\text{out}} = 21.6 \text{ \AA}$ (Fig. 2). Fitting the $p(F)$ at pulling velocities of 1 nm/s and 20 nm/s however, reveal that the $p(F)$ at these pulling velocities cannot be simultaneously described by the same x^{pf}_{out} , in contradiction to the SB model. The $p(F)$ at 20 nm/s cannot be reproduced by using the parameters, $k(0)$ and x^{pf}_{out} , determined at 1 nm/s (Fig. 5, *dotted line/magenta online*). Furthermore, the $x^{pf}_{\text{out}} = 11 \text{ \AA}$ estimated from fitting the $p(F)$ at 1 nm/s is one-half of that estimated from fitting F^* . In contrast, the two-path kinetic scheme (Fig. 5, *last panel*) is able to describe the $p(F)$ of all seven pulling velocities simultaneously. The above results show that despite the good fits to F^* , the SB is inconsistent with the $p(F)$, while the two-path kinetic scheme fits both the $p(F)$ and F^* well.

As a control, we fitted the $p(F)$ to a model where the $A1 \rightarrow D$ pathway in the two-path kinetic scheme is ignored, and the $A2 \rightarrow D$ pathway is modeled as before using the DHS rate equation. The resulting model is a variant of the traditional SB model where unfolding can also be initiated from the intermediate state at high pulling velocities.

The fits to the SB variant (Fig. 5, *dashed lines/blue online*) are clearly not as good as those to the two-path kinetic scheme (Fig. 5, *continuous lines/red online*), especially at 20–200 nm/s where the fitted distributions are significantly displaced from the $p(F)$. In 250 iterations of the fitting procedure (see Methods), a large proportion (28%) returned, numerically, very similar parameter estimates for the SB variant (Fig. 5). Significantly, further local optimization (i.e., Nelder-Mead minimization) of poorer fits yielded the same parameters, indicating that the parameter values that we obtained for the SB variant are likely to be optimal. Taken together, these results suggest that the additional unfolding pathway is crucial to the good fit of $p(F)$ generated by the two-path model, and not a redundant feature of the kinetic scheme.

CONCLUSIONS

Force spectroscopy is providing unique insights into the free-energy landscapes of proteins thanks to the single-molecule feature of the technique and the quantifiable effect of the perturbation applied on the free-energy landscape. Forced unfolding experiments are most often interpreted by fitting the data to two-state models. Such analyses relate the features (e.g., x_u) of the underlying free-energy landscape to the dependence of the modal unfolding force (F^*) on the loading rates. Most often an approximately logarithmic dependence of F^* on the cantilever retraction speed is observed, which is what the Bell model (the most widely used model) predicts. Increasingly, however, mechanical probes are revealing complexity in the free-energy landscapes of proteins (26–29) that cannot be easily accounted-for by the conventional kinetic models.

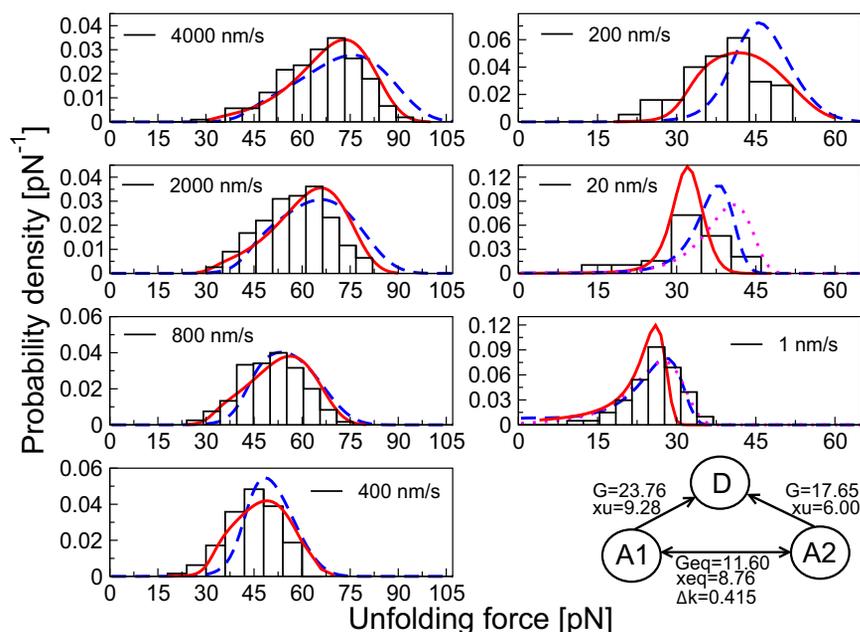


FIGURE 5 The distribution of unfolding forces at each pulling velocity (*black histograms*). (*Continuous lines/red online*) Fits to the kinetic scheme shown in the last panel (G_{eq} and G are in units of $k_B T$; x_{eq} and x_u are in units of \AA ; Δk is in units of $\text{\AA}/\text{pN}$). (*Dashed lines/blue online*) Fits to a kinetic model (denoted as the SB variant in the main text) where the pathway $A1 \rightarrow D$ is absent ($G_{\text{eq}} = 7.45 k_B T$, $x_{\text{eq}} = 1.01 \text{ \AA}$, $x_u = 4.2 \text{ \AA}$, $G = 15.75 k_B T$, and $\Delta k = 0.303 \text{ \AA}/\text{pN}$). The dotted line (*magenta online*) at 1 nm/s is a fit to the unfolding force histogram using the rate equations of Bell (7) ($k[0] = 2 \times 10^{-4} \text{ s}^{-1}$, $x^{pf}_{\text{out}} = 11 \text{ \AA}$). The dotted line at 1 nm/s cannot be seen clearly, because it overlaps perfectly with the dashed line. The dotted line (*magenta online*) at 20 nm/s is a fit to the unfolding force histogram calculated using the parameters determined at 1 nm/s.

Here, by using a combination of atomistic simulations and an analysis of the distributions of unfolding forces ($p[F]$) for the protein filamin, which was previously shown to exhibit anomalous unfolding kinetics (5), we attempted to illustrate that an analysis solely based on F^* may give a misleading picture of the free-energy landscape. Specifically, we demonstrated that while kinetic models may fit the modal or average value of the unfolding force/time to good approximation, they may dramatically violate the whole distribution of forces. For filamin, we find that the $p(F)$ can be well described by a model with two unfolding pathways, but not by a model with sequential barriers. At sufficiently low pulling velocities or forces, it is likely that a protein unfolds through both its forced and unforced pathway (11,18). In this regard, previous work (20,30,31) has shown that the forced and unforced pathways are distinct. In addition, extensive simulations of a coarse-grain model of filamin show that the two-path model is physically plausible for a small single-domain protein, and can produce a curvature similar to that observed experimentally.

Recently, it was elegantly demonstrated that anomalous kinetic signatures can also be explained by the smooth deformation of a single unfolding pathway by the applied force (32). In this scenario, the force changes the properties of the ground state and/or transition state resulting in the alteration of the force-dependence of the unfolding pathway. We note that the approach used here, which is analogous to a master-equation-type description of the kinetics, is compatible with the single-path picture. The interpretation of the kinetic equations in terms of the structure of the underlying free-energy landscape, however, is more subtle; e.g., how many coordinates are required to adequately describe the free-energy landscape? Are there two competing, coexisting transition barriers; or is there a single, albeit kinetically distinct, pathway at different forces? We previously showed that the experimental data for ddFLN4 can be fitted to a kinetic model with two pathways with very good agreement (5). However, the currently available experimental data does not allow a distinction between a free-energy landscape with two coexisting pathways and one with a single pathway. To better understand the free-energy landscape of this rather simple and very common protein fold, further experiments as well as all-atom simulations at experimental pulling velocities would be required.

Indeed, the correspondence of any model with reality cannot be guaranteed due to: 1), the use of simplifying assumptions; 2), the impossibility of examining all possible kinetic models; 3), the lack of detailed structural information due to the vast difference in timescales between full-atom molecular-dynamics simulations and mechanical unfolding experiments; and/or 4), the difficulty in performing the relevant experiments.

The conclusions drawn here should not be taken as a criticism of the validity of existing analyses (e.g., the Bell model), which are often justified in the context of the exper-

iments (e.g., comparing between mutants). Instead, the results here illustrate the currently underexploited potential of force spectroscopy techniques, particularly the wealth of information contained in the $p(F)$.

The main problem in the analysis of the $p(F)$ is the need to obtain large amounts of data (e.g., 1000 unfolding events per pulling velocity) in a large range of pulling velocities (stability and force noise). However, with the advent of improved instruments and techniques (33,34), such problems are likely to be circumvented in the near future. Despite such difficulties, current datasets may already be sufficient to discriminate between kinetic models and/or suggest alternative interpretations, which may serve as the basis of future investigations.

We thank Peter Rein ten Wolde for comments on a previous version of the manuscript and Godfrey Beddard, Dave Brockwell, and Neal Crampton for stimulating discussions.

E.P. and Z.T.Y. are supported by the Wellcome Trust and the University of Leeds. M.S. was supported in the framework of the Elite Network of Bavaria by the International Graduate School-NanoBioTechnology (grant No. IDK-NBT). M.R. was supported by an SFB 863 Grant of Deutsche Forschungsgemeinschaft.

REFERENCES

- Li, L., H. H. Huang, ..., J. M. Fernandez. 2005. Mechanical unfolding intermediates observed by single-molecule force spectroscopy in a fibronectin type III module. *J. Mol. Biol.* 345:817–826.
- Peng, Q., and H. Li. 2008. Atomic force microscopy reveals parallel mechanical unfolding pathways of T4 lysozyme: evidence for a kinetic partitioning mechanism. *Proc. Natl. Acad. Sci. USA.* 105:1885–1890.
- Schwaiger, I., A. Kardinal, ..., M. Rief. 2004. A mechanical unfolding intermediate in an actin-crosslinking protein. *Nat. Struct. Mol. Biol.* 11:81–85.
- Garcia-Manyes, S., L. Dougan, ..., J. M. Fernandez. 2009. Direct observation of an ensemble of stable collapsed states in the mechanical folding of ubiquitin. *Proc. Natl. Acad. Sci. USA.* 106:10534–10539.
- Yew, Z. T., M. Schlierf, ..., E. Paci. 2010. Direct evidence of the multidimensionality of the free-energy landscapes of proteins revealed by mechanical probes. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 81:031923.
- Evans, E., and K. Ritchie. 1997. Dynamic strength of molecular adhesion bonds. *Biophys. J.* 72:1541–1555.
- Bell, G. I. 1978. Models for the specific adhesion of cells to cells. *Science.* 200:618–627.
- Sachs, F., and H. Lecar. 1991. Stochastic models for mechanical transduction. *Biophys. J.* 59:1143–1145.
- Hummer, G., and A. Szabo. 2003. Kinetics from nonequilibrium single-molecule pulling experiments. *Biophys. J.* 85:5–15.
- Dudko, O. K., G. Hummer, and A. Szabo. 2006. Intrinsic rates and activation free energies from single-molecule pulling experiments. *Phys. Rev. Lett.* 96:108101.
- Yew, Z. T., S. Krivov, and E. Paci. 2008. Free-energy landscapes of proteins in the presence and absence of force. *J. Phys. Chem. B.* 112:16902–16907.
- Evans, E. 2001. Probing the relation between force, lifetime, and chemistry in single molecular bonds. *Annu. Rev. Biophys. Biomol. Struct.* 30:105–128.
- Merkel, R., P. Nassoy, ..., E. Evans. 1999. Energy landscapes of receptor-ligand bonds explored with dynamic force spectroscopy. *Nature.* 397:50–53.

14. Thomas, W. E., E. Trintchina, ..., E. V. Sokurenko. 2002. Bacterial adhesion to target cells enhanced by shear force. *Cell*. 109:913–923.
15. Marshall, B. T., M. Long, ..., C. Zhu. 2003. Direct observation of catch bonds involving cell-adhesion molecules. *Nature*. 423:190–193.
16. Evans, E., A. Leung, ..., C. Zhu. 2004. Mechanical switching and coupling between two dissociation pathways in a P-selectin adhesion bond. *Proc. Natl. Acad. Sci. USA*. 101:11281–11286.
17. Bartolo, D., I. Derényi, and A. Ajdari. 2002. Dynamic response of adhesion complexes: beyond the single-path picture. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 65:051910.
18. West, D. K., P. D. Olmsted, and E. Paci. 2006. Mechanical unfolding revisited through a simple but realistic model. *J. Chem. Phys.* 124:154909.
19. Best, R. B., E. Paci, ..., O. K. Dudko. 2008. Pulling direction as a reaction coordinate for the mechanical unfolding of single molecules. *J. Phys. Chem. B*. 112:5968–5976.
20. Williams, P. M., S. B. Fowler, ..., J. Clarke. 2003. Hidden complexity in the mechanical properties of titin. *Nature*. 422:446–449.
21. Derényi, I., D. Bartolo, and A. Ajdari. 2004. Effects of intermediate bound states in dynamic force spectroscopy. *Biophys. J.* 86:1263–1269.
22. Karanicolas, J., and C. L. Brooks, 3rd. 2003. The structural basis for biphasic kinetics in the folding of the WW domain from a formin-binding protein: lessons for protein design? *Proc. Natl. Acad. Sci. USA*. 100:3954–3959.
23. Dudko, O. K., G. Hummer, and A. Szabo. 2008. Theory, analysis, and interpretation of single-molecule force spectroscopy experiments. *Proc. Natl. Acad. Sci. USA*. 105:15755–15760.
24. Schlierf, M., and M. Rief. 2006. Single-molecule unfolding force distributions reveal a funnel-shaped energy landscape. *Biophys. J.* 90:L33–L35.
25. Dill, K. A., and H. S. Chan. 1997. From Levinthal to pathways to funnels. *Nat. Struct. Biol.* 4:10–19.
26. Forman, J. R., Z. T. Yew, ..., J. Clarke. 2009. Non-native interactions are critical for mechanical strength in PKD domains. *Structure*. 17:1582–1590.
27. Serquera, D., W. Lee, ..., L. S. Itzhaki. 2010. Mechanical unfolding of an ankyrin repeat protein. *Biophys. J.* 98:1294–1301.
28. Brujic, J., R. I. Hermans, ..., J. M. Fernandez. 2006. Single-molecule force spectroscopy reveals signatures of glassy dynamics in the energy landscape of ubiquitin. *Nat. Phys.* 2:282–286.
29. Dougan, L., J. Li, ..., J. M. Fernandez. 2009. Single homopolypeptide chains collapse into mechanically rigid conformations. *Proc. Natl. Acad. Sci. USA*. 106:12605–12610.
30. Paci, E., and M. Karplus. 2000. Unfolding proteins by external forces and temperature: the importance of topology and energetics. *Proc. Natl. Acad. Sci. USA*. 97:6521–6526.
31. Klimov, D. K., and D. Thirumalai. 2000. Native topology determines force-induced unfolding pathways in globular proteins. *Proc. Natl. Acad. Sci. USA*. 97:7254–7259.
32. Suzuki, Y., and O. K. Dudko. 2010. Single-molecule rupture dynamics on multidimensional landscapes. *Phys. Rev. Lett.* 104:048101.
33. Schlierf, M., F. Berkemeier, and M. Rief. 2007. Direct observation of active protein folding using lock-in force spectroscopy. *Biophys. J.* 93:3989–3998.
34. Liu, R., S. Garcia-Manyes, ..., J. M. Fernandez. 2009. Mechanical characterization of protein L in the low-force regime by electromagnetic tweezers/evanescent nanometry. *Biophys. J.* 96:3810–3821.