Energy landscape analysis of native folding of the prion protein yields the diffusion constant, transition path time, and rates

Hao Yu[a,1], Amar Nath Gupta[a,1], Xia Liu[a,1], Krishna Neupane[a], Angela M. Brigley[b], Iveta Sosova[b], and Michael T. Woodside[a,b,2]

[a]Department of Physics, University of Alberta, Edmonton, AB, Canada T6G 2E1; and [b]National Institute for Nanotechnology, National Research Council, Edmonton, AB, Canada T6G 2M9

Edited by William A. Eaton, National Institutes of Health – NIDDK, Bethesda, MD, and approved August 2, 2012 (received for review April 12, 2012)

Protein folding is described conceptually in terms of diffusion over a configurational free-energy landscape, typically reduced to a one-dimensional profile along a reaction coordinate. In principle, kinetic properties can be predicted directly from the landscape profile using Kramers theory for diffusive barrier crossing, including the folding rates and the transition time for crossing the barrier. Landscape theory has been widely applied to interpret the time scales for protein conformational dynamics, but protein folding rates and transition times have not been calculated directly from experimentally measured free-energy profiles. We characterized the energy landscape for native folding of the prion protein using force spectroscopy, measuring the change in extension of a single protein molecule at high resolution as it unfolded/refolded under tension. Key parameters describing the landscape profile were first recovered from the distributions of unfolding and refolding forces, allowing the diffusion constant for barrier crossing and the transition path time across the barrier to be calculated. The full landscape profile was then reconstructed from force-extension curves, revealing a double-well potential with an extended, partially unfolded transition state. The barrier height and position were consistent with the previous results. Finally, Kramers theory was used to predict the folding rates from the landscape profile, recovering the values observed experimentally both under tension and at zero force in ensemble experiments. These results demonstrate how advances in single-molecule theory and experiment are harnessing the power of landscape formalisms to describe quantitatively the mechanics of folding.

kinetics | optical trapping | single molecule

Folding free-energy landscapes contain, in principle, all the information needed to describe the conformational dynamics of a protein, from the folding kinetics to the locations of energy barriers and the existence of intermediates or nonnative pathways (1). The competition between potential energy and entropy within the funnel-like landscape of most proteins typically allows the full landscape to be reduced to a one-dimensional free-energy profile along the reaction coordinate (2, 3). Folding rates may then be described using Kramers theory (4) in terms of diffusion across a barrier of height $\Delta G^{\ddagger}$, where the barrier represents the bottleneck formed by the transition state ensemble. Kramers theory provides a physical derivation of the rates in terms of the shape of the energy profile (Fig. 1):

$$k = k_0 \exp(-\Delta G^{\ddagger}/k_{\mathrm{B}}T), \quad \text{where} \quad k_0 = \frac{\sqrt{\kappa_w \kappa_b}}{2\pi k_{\mathrm{B}} T} D, \qquad [1]$$

D is the diffusion constant over the barrier, κ_w is the stiffness (curvature) of the potential well, κ_b is the stiffness of the barrier, and k_B is the Boltzmann constant. The time required to cross over the barrier, the transition path time τ_{tp}, may also be found from the landscape profile (5–7): For an harmonic barrier with $\Delta G^{\ddagger} > 2\ k_{\mathrm{B}}T$ (6),

$$\tau_{\mathrm{tp}} \approx \frac{\ln(2e^{\gamma}\Delta G^{\ddagger}/k_{\mathrm{B}}T)}{D\kappa_b/k_{\mathrm{B}}T} = \frac{\ln(2e^{\gamma}\Delta G^{\ddagger}/k_{\mathrm{B}}T)}{2\pi k_0\sqrt{\kappa_b/\kappa_w}}, \qquad [2]$$

where γ is Euler's constant and the expression becomes exact in the limit of large barrier height. Generally $\tau_{\mathrm{tp}} \ll k^{-1}$, because a significant amount of time is spent diffusing within the potential well before the barrier is actually crossed.

Kramers theory has been widely applied to interpret the time scales for protein conformational dynamics in the context of landscape theory (1, 8–11). Numerous studies have investigated the microscopic protein dynamics encapsulated by D in Eq. **1**, which is important because D effectively describes the roughness of the landscape and sets the fundamental "speed limit" for protein folding (10). The rate of contact formation within short peptides of various lengths has been measured with energy transfer and quenching (12, 13), as has the intrachain diffusion constant for a variety of naturally folding proteins, but only in the denatured or nearly denatured state (11, 14, 15). Importantly, these approaches do not probe chain dynamics under conditions approximating the transition state, where hydrophobic collapse has likely occurred or structure is partly formed. The transition path time has been much less well studied: τ_{tp} has been estimated for protein folding from single-molecule fluorescence trajectories (6, 16), but it has not yet been calculated from folding landscapes. Landscape theories have been used to predict protein folding rates, but only using landscapes derived from simulations (10), because free-energy profiles are difficult to measure.

To demonstrate experimentally the ability of landscape theory to predict folding rates and transition path times, we reconstructed the free-energy landscape for a folding protein using force spectroscopy. Force spectroscopy, in which folding trajectories of a single molecule are observed directly by measuring the end-to-end extension as the molecule unfolds and refolds under the effects of an applied tension, provides a powerful tool for characterizing energy landscapes (17). The height of the energy barrier, its location along the reaction coordinate, and the unloaded unfolding/refolding rate may all be determined from measurements of the distribution of forces required to unfold/refold

Author contributions: M.T.W. designed research; H.Y., A.N.G., X.L., and K.N. performed research; H.Y., A.N.G., X.L., K.N., A.M.B., and I.S. contributed new reagents/analytic tools; H.Y., A.N.G., X.L., K.N., and M.T.W. analyzed data; and H.Y., A.N.G., X.L., K.N., A.M.B., I.S., and M.T.W. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]H.Y., A.N.G., and X.L. contributed equally to this work.

[2]To whom correspondence should be addressed. E-mail: michael.woodside@nrc-cnrc.gc.ca.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1206190109/-/DCSupplemental.

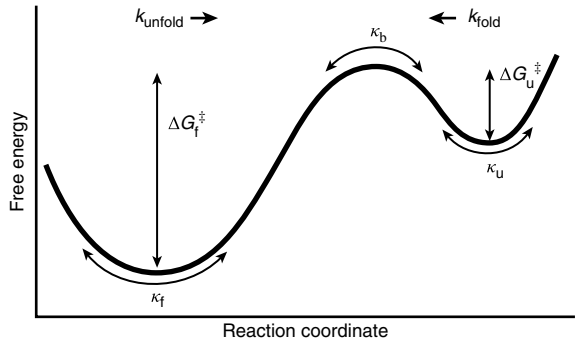

Fig. 1. Kramers theory of reaction rates. The rates for folding (k_{fold}) and unfolding (k_{unfold}) can be calculated for a given shape of the energy landscape in terms of the diffusion constant over the energy barrier (*D*), the height of the energy barriers ($\Delta G_u^{\ddagger}$ from the unfolded state and $\Delta G_f^{\ddagger}$ from the folded state, respectively), and the stiffness of the potential wells (κ_u for the unfolded state and κ_f for the folded state, respectively) and barrier (κ_b).

the molecule as the force is ramped up/down (9, 18). Such key parameters may also be obtained from equilibrium measurements of the extension under constant force (19). Full one-dimensional free-energy profiles may even be reconstructed, from either equilibrium probability distributions of the molecular extension (20, 21) or fluctuation theorem analysis of force-extension curves (FECs) using the Hummer-Szabo formalism (22–24).

We combined these methods to reconstruct the free-energy landscape for native folding of PrP, a highly conserved, membrane-associated protein. PrP folds rapidly into its native structure (25), but it also forms stable, nonnative structures which propagate infectiously, causing transmissible spongiform encephalopathies (26). Both the structure of the misfolded infectious state and the mechanisms by which it forms and propagates remain controversial, motivating efforts to understand better what properties give rise to PrP's unusual behavior. From the parameters describing the energy landscape, we calculated the diffusion constant for barrier crossing, the average transition time for folding and unfolding, and finally the folding and unfolding rates.

Results

To measure the folding of single PrP molecules, recombinant hamster PrP with its unstructured N terminus truncated at residue 90 was covalently attached to "handles" made of double-stranded DNA, which were in turn bound to beads held in a high-resolution dual-beam optical trap as shown in Fig. 2*A* (inset) (27, 28). We first recorded FECs where PrP unfolded from its native state. The traps were moved apart to ramp up the force until PrP unfolded. Representative unfolding FECs (Fig. 2*A*, red) reveal a nonlinear rise in force with extension arising from the elasticity of the handles, interrupted by a sudden increase in extension and concomitant drop in force as PrP unfolded in a single step. The contour length change from folded (N) to unfolded (U) was determined by fitting the FECs to two wormlike chains (WLCs) in series, one for the DNA handles and one for the protein (28). The value obtained from fitting 3,250 unfolding FECs, 34.1 ± 0.4 nm (all errors: standard error on the mean), agrees well with the 34.3 nm expected from the NMR structure of the native state (29), confirming that the PrP was indeed natively folded before each pull.

The distribution of unfolding forces, $p(F)$, can be described in terms of key features of the folding landscape (9):

$$p(F) \propto \frac{k(F)}{r} \exp\left\{\frac{k_{unfold}}{\Delta x^{\ddagger} r} - \frac{k(F)}{\Delta x^{\ddagger} r}\left(1 - \frac{\Delta x^{\ddagger} F}{\Delta G^{\ddagger}}\nu\right)^{1-1/\nu}\right\}, \quad \text{where} \qquad [3]$$

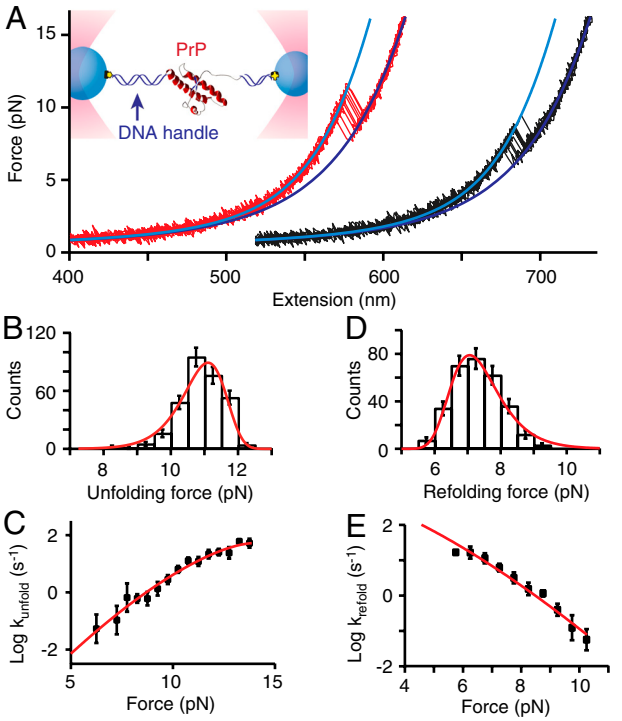

Fig. 2. Force-extension curves of PrP unfolding. (*A*) DNA handles attached to each end of a PrP molecule are bound to beads held under tension in an optical trap (inset). Ramping up the force produces a FEC: The DNA handles stretch until the protein unfolds abruptly, here in a two-state process. 10 representative unfolding FECs (red) and refolding FECs (black) are shown (curves offset for clarity). Curved lines represent WLC fits to the folded (cyan) and unfolded (blue) states. (*B*) The distribution of unfolding forces fits well to [**3**], yielding the barrier height for unfolding and location along the reaction coordinate. (*C*) The unfolding rate as a function of force obtained from the FECs (18) fits well to Eq. **4**, yielding the same parameters. (*D*, *E*) The same fits for the refolding rate and refolding force distribution yield the barrier height and location for refolding.

$$k(F) = k_{unfold}\left(1 - \frac{\Delta x^{\ddagger} F}{\Delta G^{\ddagger}}\nu\right)^{1/\nu-1} \exp\left\{\frac{\Delta G^{\ddagger}}{k_B T}\left[1 - \left(1 - \frac{\Delta x^{\ddagger} F}{\Delta G^{\ddagger}}\nu\right)^{1/\nu}\right]\right\} \qquad [4]$$

is the force-dependent unfolding rate, k_{unfold} is the unfolding rate at zero force, $\Delta x^{\ddagger}$ is the distance to the transition state from the folded state, $\Delta G^{\ddagger}$ is the energy barrier height from the folded state, r is the loading rate, and ν parameterizes the shape of the barrier ($\nu = 2/3$ for a linear-cubic potential as assumed here). The measured unfolding force distributions (Fig. 2*B*) were well fit by Eq. **3**. We also calculated the force-dependent unfolding rates (Fig. 2*C*) from the survival times during the pulling measurements (18), and fit them to Eq. **4**. Based on these two sets of fits, we found that $\Delta x^{\ddagger} = 9 \pm 1$ nm from the native state, $\log k_{unfold} = -6 \pm 1\ s^{-1}$, and $\Delta G^{\ddagger} = 64 \pm 6$ kJ/mol from the native state.

Separately, FECs were also measured while ramping down the force to refold the protein. Behavior similar to the unfolding FECs was seen, with the same change in contour length, as expected (Fig. 2*A*, black). The refolding force distributions (Fig. 2*D*) and force-dependent refolding rates (Fig. 2*E*) were fit to Eqs. **3** and **4**, analogously to the data from the unfolding FECs, replacing k_{unfold} with k_{fold} (the folding rate at zero force). The results from 1,062 refolding curves were $\Delta x^{\ddagger} = 3.0 \pm 0.6$ nm from the unfolded state, $\log k_{fold} = 3.9 \pm 0.2\ s^{-1}$, and $\Delta G^{\ddagger} = 5 \pm 3$ kJ/mol from the unfolded state.

Using these landscape parameters, we evaluated the transition path time from Eq. **2**, assuming that $\kappa_b \approx \kappa_w$ (6) and making use of the fact that the Kramers prefactor can be expressed in terms of the fitting parameters as $k_0 = k_{\mathrm{fold/unfold}} \exp(\Delta G^{\ddagger}/k_{\mathrm{B}}T)$. The result for the unfolding transition was $\tau_{\mathrm{tp}} = 4 \times 10^{0\pm1}$ μs, where the uncertainty arises mainly from the dependence of k_0 on $\Delta G^{\ddagger}$ and hence is in the exponent. For the refolding transition, the landscape parameters implied a similar value, $\tau_{\mathrm{tp}} = 5 \times 10^{0\pm0.5}$ μs. Given that τ_{tp} is expected to be the same regardless of the direction of the transition (5), we averaged the results to obtain $\tau_{\mathrm{tp}} = 5 \times 10^{0\pm0.4}$ μs. The landscape parameters from the FEC analysis also allowed the diffusion constant for barrier crossing to be determined: For the linear-cubic potential profile being assumed,

$$D = \frac{\pi}{3}\left[\frac{k_{\mathrm{fold/unfold}}(\Delta x^{\ddagger})^2}{\Delta G^{\ddagger}/k_{\mathrm{B}}T}\right]\exp\left(\frac{\Delta G^{\ddagger}}{k_{\mathrm{B}}T}\right), \qquad [5]$$

using k_{unfold} in the case of the unfolding results and k_{fold} in the case of refolding results, along with the respective barrier locations and heights. As for τ_{tp}, the values were calculated independently for unfolding and refolding. For the unfolding transition, $D = 6 \times 10^{-13\pm1}$ m^2/s, whereas for refolding, essentially the same result was found, $D = 3 \times 10^{-13\pm0.6}$ m^2/s. Because the values are again expected to be the same, we averaged them to obtain $D = 4 \times 10^{-13\pm0.5}$ m^2/s.

This analysis assumed a specific shape for the landscape profile (i.e., linear-cubic). To verify that this assumption was reasonable, we reconstructed the full profile of the landscape from the FECs using the Hummer-Szabo formalism (24). Landscape reconstructions based on high-bandwidth equilibrium measurements of the extension at constant force or trap separation typically provide higher spatial resolution (20, 22). However, equilibrium trajectories present interpretation difficulties for PrP, because in addition to native folding they also contain transitions into short-lived, nonnative states accessible only from the unfolded state (28, 30), which would distort the reconstructed landscape. In contrast, unfolding FECs always go first from N to U, with subsequent refolding into the misfolded states suppressed by virtue of the nonequilibrium conditions caused by the fast force-ramp rate. As a consequence, landscapes reconstructed from unfolding FECs should reflect primarily the properties of the native pathway, with minimal contribution from the nonnative pathways.

The free-energy profile at zero force was calculated from each set of FECs measured under the same conditions, as described previously (22, 24). The resulting profiles are dominated by the stretching energy of the handles in the folded (Fig. 3*A*, cyan) and unfolded (Fig. 3*A*, blue) states, as expected (22). To see more clearly the potential wells for the native and unfolded states, the profiles were tilted at a force of $F = 9.1$ pN (Fig. 3*B*, red: Individual reconstructions, black: average). The barrier between the wells is indeed dominated by the protein, as seen by comparing the reconstructed landscape to the energy "profile" that consists solely of the WLC energies of the folded and unfolded states (Fig. S1).

The profile in Fig. 3*B* indicates a two-state system with a barrier roughly midway between the native and unfolded states under 9 pN of tension. As an independent test of whether this overall picture is correct, we analyzed equilibrium measurements of the extension at constant force, in the range approximately 8–10 pN. Previous work has shown that the native folding pathway did not include any intermediate state (28, 30). Representative records at different forces (Fig. 4*A*) were filtered to remove the short-lived off-pathway states (28) and the lifetimes of the two states were determined by threshold analysis (19). The lifetimes were single-exponentially distributed at each force (Fig. 4*B*), as expected for a transition with a single rate-determining barrier. The folding and unfolding rates varied exponentially with force

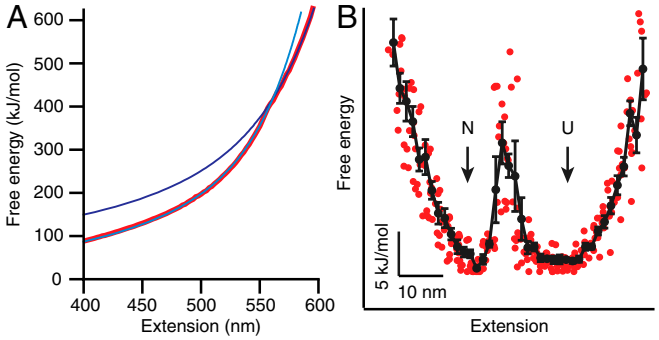

Fig. 3. Free-energy profile reconstruction for PrP folding. (*A*) The molecular free-energy profile at zero force reconstructed from a set of FECs by the Hummer-Szabo method (red) is dominated by the energies of the handles in the folded (cyan) and unfolded (blue) states. (*B*) Reconstructions from different sets of FECs (red) were tilted to 9.1 pN and averaged (black). Error bars show standard error.

(Fig. 4*C*) and were well fit, over the limited force range sampled, by the Bell formula (31), $k \propto \exp(\pm F\Delta x^{\ddagger}/k_{\mathrm{B}}T)$. These fits yielded a barrier 10.4 ± 0.6 nm from the folded state at approximately 8–10 pN, in good agreement with the location of the barrier in Fig. 3*B*. The rates from constant force measurements also agreed reasonably well with the rates obtained from the FECs (Fig. 2). The force producing equal folding and unfolding rates, $F_{1/2}$, was 9.1 ± 0.1 pN.

The reconstructed profile in Fig. 3*B* includes not only the intrinsic PrP free-energy profile but also the effects of the compliant DNA handles and beads, which smooth the PrP landscape profile (22). Attempts to calculate D and τ_{tp} directly from this profile therefore produce incorrect results: Using the curvatures of the wells and barrier from Fig. 3 and the folding/unfolding rates at 9.1 pN (Fig. 4*C*), $D = 6 \times 10^{-15\pm0.4}$ m^2/s and hence $\tau_{\mathrm{tp}} = 3 \times 10^{2\pm0.4}$ μs. Both of these values are two orders of magnitude slower than those obtained from the FEC kinetic analysis above; τ_{tp} is also one order slower than the upper limit of approximately 40 μs found by direct examination of the trajectories (Fig. S2). To remove the smoothing effects of the instrument, we used the point-spread function (PSF) of the instrument measured from a construct consisting of beads and handles alone, without protein (Fig. S3), to perform an iterative nonlinear deconvolution of the landscape profile. This approach has been applied previously to landscape profiles obtained via equilibrium measurements (20, 21), but not yet using the Hummer-Szabo approach (22, 23). The deconvolved landscape at 9.1 pN is shown in Fig. 5 (Fig. 5*A*, red); the associated extension probability distribution and residual error are shown in Fig. S4.

We first tested the consistency of the deconvolved profile with the results found previously from FECs and constant force measurements. At 9.1 pN tension, the barrier is located approximately 8.5 nm from the folded state and approximately 11.5 nm from the unfolded state (with an error of one bin width, or 1.5 nm), in good agreement with the $\Delta x^{\ddagger}$ values from constant-force trajectories. Tilting the profile back to zero force (Fig. 5*B*, red), we compared it to the results of the FEC kinetic analysis. The barriers previously found at zero force for leaving the folded state (Fig. 5*B*, black) and the unfolded state (Fig. 5*B*, blue) agree well with the barrier in the reconstruction, within error (the uncertainty in the barrier height after deconvolution is estimated at ±5 kJ/mol), validating the reconstruction and deconvolution. In addition, the shape of the landscape supports the earlier assumption of a linear-cubic profile for determining D and τ_{tp}.

Next, we tilted the deconvolved profile over a wide range of forces and calculated the folding and unfolding rates predicted by Kramers theory [Eq. **1**]. Because the rates at all forces depend on the same diffusion constant D, this effectively amounts to fitting

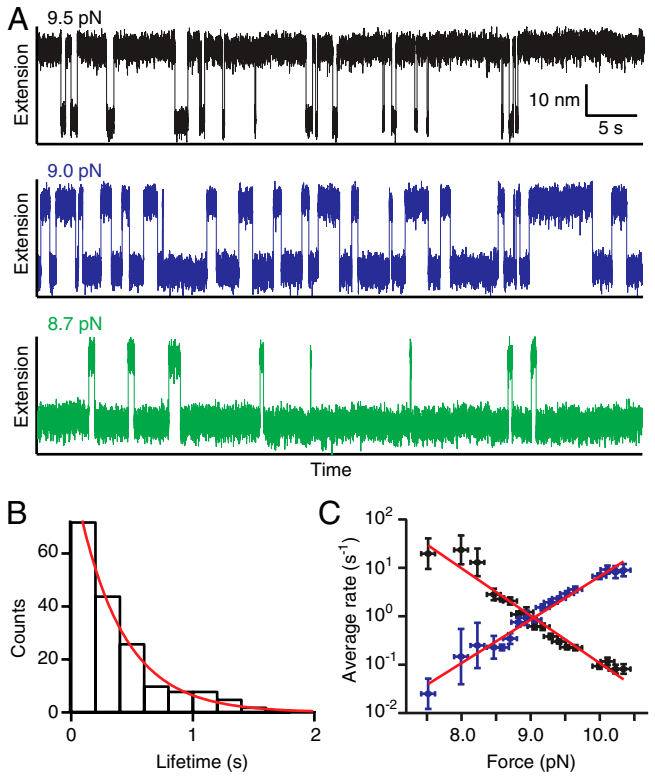

Fig. 4. Constant-force trajectories of PrP folding. (*A*) Extension records at constant force for three different forces show abrupt jumps as PrP unfolds/refolds in a two-state process. (*B*) The distribution of lifetimes for N and U are single-exponential. Here the distribution and exponential fit are shown for N at 9.4 pN. (*C*) The rates for folding (black) and unfolding (blue) vary exponentially with force, crossing at force $F_{1/2}$. Error bars show standard error.

globally all the force-dependent rates in Fig. 2 *C* and *E*, and 4*C*. The global fit yielded $D = 8 \times 10^{-11\pm1}$ m^2/s, consistent within error with the result from the FEC kinetic analysis. The predicted rates (Fig. 6, red: folding, cyan: unfolding) recapitulate the force-dependent rates found from both the FEC kinetic analysis (Fig. 6, black: folding, purple: unfolding) and the constant-force measurements (Fig. 6, grey: folding, blue: unfolding), over four orders of magnitude. They also agree well, within error, with the rates found at zero force (Fig. 6, brown: folding, green: unfolding) from chemical denaturation studies (25).

Finally, we used the fitted value of D to recalculate τ_{tp} from the deconvolved profile. The result, $\tau_{tp} = 2 \times 10^{-2\pm1}$ μs, is somewhat faster than the previous result from the kinetic analysis, but consistent given the relatively large uncertainty.

Discussion

The results presented above show that the protein folding-energy landscape profile reconstructed by measuring single molecules under tension can be used to predict key properties, ranging from commonly studied observables such as the folding rates to rarely characterized measures such as the transition path time, as well as fundamental quantities such as the diffusion constant. To assess the validity of these results, we compared them to independent measurements of PrP and other proteins using different techniques. Considering first the rates, the folding rate predicted at zero force from Fig. 5*B*, $5 \times 10^{3\pm1}$ s^{-1}, agrees well with the rate found for PrP from ensemble chemical denaturation measurements, approximately 4×10^3 s^{-1} (25). This agreement between ensemble and single-molecule results supports the notion that the same barrier is being probed in the two different measurements, despite the different modes of denaturation (32). The Kramers

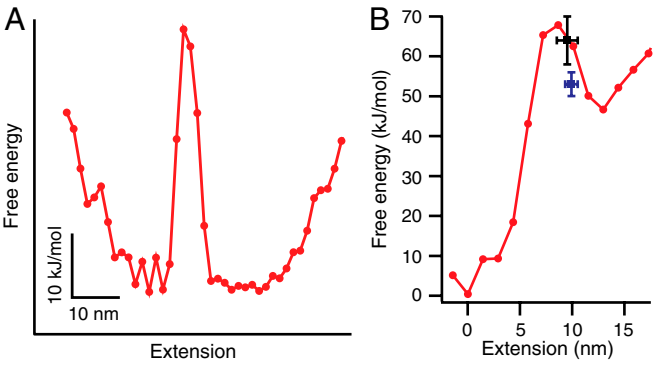

Fig. 5. Energy landscape deconvolution. (*A*) The free-energy profile at $F_{1/2}$ after deconvolution. (*B*) The deconvolved free-energy profile tilted to zero force (red) agrees with the barrier height and location from the kinetic FEC analysis of [3] and [4] (black: unfolding, blue: refolding).

prefactor k_0 calculated at zero force, $10^{7\pm1}$ s^{-1} from Fig. 5 or $10^{5\pm0.5}$ s^{-1} from the FEC kinetic analysis, is also consistent with the range found experimentally in previous work: 10^6–10^8 s^{-1} for contact formation in unfolded polypeptides and proteins (11, 12), though it is lower than a previous estimate of 10^9 s^{-1} from unfolding ubiquitin by atomic force microscopy (33).

The force-dependence of the rates depends primarily on the distance to the barrier, $\Delta x^{\ddagger}$: very small $\Delta x^{\ddagger}$ leads to "brittle" behavior where the rate changes little with force, whereas larger $\Delta x^{\ddagger}$ results in "compliant" behavior with a strong force-dependence. PrP folding is clearly compliant, similar to proteins like helical coiled-coils (20) or calmodulin (34), but contrasting with the folding of many other proteins like titin (35), ubiquitin (33), and GFP (36). This result suggests that the key interactions holding the structure together are located far from the points at which force is applied, allowing large extension fluctuations to occur before the protein reaches the transition state. Given that each unfolded amino acid extends approximately 0.22 nm at $F_{1/2}$, $\Delta x^{\ddagger}$ corresponds to approximately 40 amino acids unfolded. Such a non-compact transition state was recently suggested by phi-analysis of PrP folding (37), which showed a structural nucleus developing between the helices 2 and 3 of the native fold (Fig. S5). One possible transition state consistent with both the single-molecule and phi-analysis results would retain the core of the molecule (most of helix 2 and 3, the adjacent loop and β-strand 2) while unfolding β-strand 1 and helix 1. This picture is also supported by NMR studies showing that β-strand 1 is relatively unstable (38), but

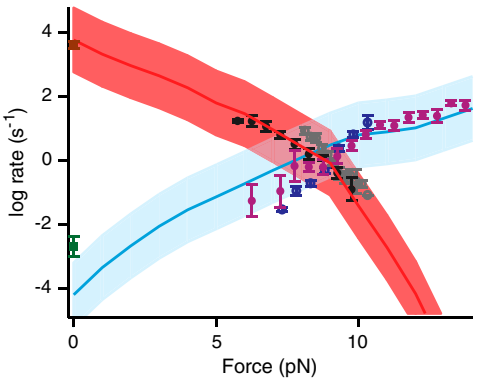

Fig. 6. Comparison between experimental and predicted rates. The rates for folding and unfolding found from FEC analysis (black: folding; purple: unfolding) and constant-force measurements (grey: folding; blue: unfolding) agree within error with the rates calculated directly from the deconvolved landscape profile by Kramers theory, when tilted at various forces (red: folding; cyan: unfolding). Shaded areas represent error in calculated rates. Zero-force folding (brown) and unfolding (green) rates are from chemical denaturation measurements (25).

additional experiments, possibly combining phi-analysis with single-molecule measurements, will be required to explore the nature of the transition state in more detail.*

The most interesting aspect of the landscape analysis is the ability to quantify properties such as the diffusion constant over the barrier and the transition path time, which are extremely difficult to measure by any method. For the diffusion constant, an average of the results from the two methods used above (kinetic analysis and landscape deconvolution) yields the final result $D = 1 \times 10^{-12\pm0.4}$ m^2/s. This value is similar to the range reported from fluorescence studies of intrachain diffusion in unfolded proteins: Typically $D \sim 10^{-10}$–10^{-11} m^2/s (11, 14, 15), although in one case as low as 10^{-13} m^2/s (39). Crucially, however, our result characterizes diffusion across the barrier itself, the critical parameter in Kramers theory. In contrast, previous work described diffusion within the unfolded state only, which may be different (8). The value we find is at the low end of the range, as would be expected because any structure formed near the transition state would presumably slow the configurational diffusion relative to the unstructured chain. We are not aware of any other published measurements of D for barrier crossing in protein folding, for more direct comparison. However, our results are consistent with a previous analysis of force spectroscopy measurements of titin unfolding (40, 41), which implied a value $D = 1 \times 10^{-14\pm2}$ m^2/s (although D was not explicitly calculated).

Turning to the transition path time, an average of the results from the kinetic analysis and landscape deconvolution yields $\tau_{tp} = 2 \times 10^{0\pm0.4}$ μs. This value is much faster than the characteristic time scale for folding at zero force, approximately 100 μs, but in excellent agreement with the only direct measurements of τ_{tp} for folding proteins, which found $\tau_{tp} \sim 2$ μs for the WW domain of the formin binding protein and $\tau_{tp} < 10$ μs for the protein GB1 (16). Our result also agrees very well with an atomistic equilibrium molecular dynamics simulation in explicit solvent, which found $\tau_{tp} = 0.5 \pm 0.1$ μs for the FiP WW domain (42), corresponding to $\tau_{tp} \sim 1.5$ μs after viscosity correction (16). Interestingly, PrP is twice as large as GB1 and three times as large as the WW domains, but still has a similar transition path time, suggesting that τ_{tp} is relatively insensitive to protein size. The topology of PrP is also quite distinct from that of the other proteins, suggesting that τ_{tp} is also insensitive to topology. In contrast, both protein size (43, 44) and the topology of the native fold (45) are known to play important roles in determining the folding rate.

In conclusion, these results demonstrate the power of single-molecule methods for quantifying the fundamental properties driving folding reactions. They also represent a stringent test of the underlying landscape theories, showing that landscape theory can be used to understand force spectroscopy measurements in a comprehensive, unified way. Not only can the location and height of free-energy barriers be determined, and indeed the full profile of the landscape, but these can even be used to predict folding and unfolding rates over many orders of magnitude, as well as the values of elusive quantities such as the diffusion constant and transition path time. A similar analysis could be applied to force spectroscopy measurements for many other proteins, pointing the way towards increasingly quantitative experimental applications of folding landscape theories.

*It was recently suggested from a study of apomyoglobin that compliant folding may be indicative of a molten globule rather than the native structure (49). We do not believe that we are observing a molten globule state for PrP, for several reasons. First, PrP was held at low force (<2 pN) for 1–2 s between unfolding events, more than sufficient time to form the native structure given the very fast folding rate (25). Similar conditions in the apomyoglobin study were indeed claimed to lead to the native structure; the molten globule was only observed when insufficient refolding time was allowed at low force (e.g., in constant-force measurements). In contrast to this previous study, however, we did not see any change in $\Delta x^{\ddagger}$ for unfolding when comparing FECs to constant force measurements. Second, there is independent support for an extended transition state from phi-analysis (37). Third, the agreement between the folding rate predicted from the landscape at zero force and the measured rate strongly suggests that the same barrier is being probed in the pulling measurements and ensemble denaturation studies.

Methods

Sample Preparation. Truncated, His-tagged, wild-type hamster PrP(90–231) was engineered with cysteine residues at each terminus, expressed in *Escherichia coli*, purified by affinity FPLC, and refolded on-column as described (28). Protein identity was checked by Western blotting, and folding into the native structure verified by CD spectroscopy. Double-stranded DNA handles were prepared and attached to PrP, and the resulting chimeras were bound to polystyrene beads (28). Reference constructs used to measure the PSF were made by the same method but omitting PrP. All samples were placed in 50 mM MOPS, pH 7.0, with 200 mM KCl and an oxygen scavenging system (8 mU/μL glucose oxidase, 20 mU/μL catalase, 0.01% w/v D-glucose) for measurement in the optical trap. Under these conditions, the internal disulfide in the PrP molecules was reduced (28).

Force Spectroscopy Measurements. Measurements were performed on a custom high-resolution optical trap in a temperature-controlled room (28) at 20.0 ± 0.1 °C. For FECs, the traps had stiffness 0.3 pN/nm and 0.9 pN/nm and were moved apart at rates of 20–230 nm/s. Data were sampled at 20 kHz and filtered online at the Nyquist frequency with an 8-pole Bessel filter. Sets of FECs from the same molecule under the same conditions were aligned offline to correct for small amounts of instrumental drift. Constant-force data were measured with a passive force clamp (46), with trap stiffness 0 pN/nm and 0.3 pN/nm. Data were sampled at 20 kHz or 50 kHz, filtered online as for FECs, and median-filtered offline in a 5-ms window to remove the short-lived off-pathway states (28). The native and unfolded states were separated by a threshold set to the midpoint between the two peaks in the extension histograms (19). Trap stiffnesses were calibrated (47) separately for FEC and constant-force measurements.

Landscape Reconstruction. The free-energy landscape profile of the protein plus handles was reconstructed using the Hummer-Szabo formalism as previously described (22). The profile was calculated for each set of FECs containing at least 100 pulls on a given molecule at a given pulling speed, then all profiles were averaged. Bin widths for the reconstructions were chosen to ensure sufficient data within each bin to define the work for that bin reliably. During averaging, the profile was rebinned with 1.5 nm width. The landscape was not reconstructed from constant-force extension distributions as in Fig. 4*A* (21) because of the existence of frequently-sampled but short-lived off-pathway states (28). These off-pathway states are not visible in Fig. 4*A* due to the data filtering.

Point-Spread Function. PSF for deconvolution was measured using a reference construct consisting of two DNA handles connected together without any protein. This reference construct of fixed, known length was held under tension at a constant trap separation, and its extension was measured under the same conditions as for the FECs for 100–120 s. The extension distribution (Fig. S3) was almost Gaussian, but slightly asymmetric as described previously (28). Because the free-energy profile being deconvolved was reconstructed from FECs in which the force depended on the molecular extension, the PSF was measured over a range of forces spanning the typical forces used during unfolding.

Free-Energy Profile Deconvolution. The smoothing of the reconstructed free-energy profile due to elastic compliance was removed by pointwise nonlinear iterative deconvolution as described previously (20, 21). The instrument PSF, $S(x)$, smoothed the true extension probability function to produce the measured extension probability, $P(x)$, hence the deconvolution was performed on $P(x)$. The free-energy profile, $\Delta G(x)$, was then recovered from $P(x)$ by $\Delta G(x) = -k_B T \cdot \ln(P(x))$. The deconvolution was done pointwise because the PSF width varied with force, which itself varied with extension (Fig. 2*A*). An initial solution $P^0(x)$ was first calculated from the average profile reconstruction at $F_{1/2}$. The true distribution function for the protein alone was then approached iteratively, with the $(k+1)$th iteration at molecular extension a given by (20, 21):

$$P^{k+1}(x) = P^k(x) + r[P^k(x)] \times \{P(x) - S^a(x) \otimes P^k(x)\}, \quad [6]$$

where $S^a(x)$ is the PSF corresponding to the average force at extension a. The relaxation function $r[P^k(x)] = r_0(1 - 2 \cdot |P^k(x) - 0.5|)$ constrained the solution to remain within the boundaries $0 \le P^k(x) \le 1$ required for a

physical probability function, with the amplitude r_0 controlling the speed of convergence. Using $r_0 = 2$, the solution converged after approximately 40,000 iterations. To reduce artifactual fluctuations in the deconvolution due to measurement noise, $P^0(x)$ was smoothed in a 3 nm window; $S^a(x)$ was smoothed in the same way to compensate for this additional smoothing (48). The centre of the low-extension probability distribution peak, giving the location of the folded state, was used as the origin for the profile at zero force. The energy required to stretch the unfolded protein under tension was found from integrating the protein-only WLC curve and subtracted from the profile at zero force.

ACKNOWLEDGMENTS. We thank Attila Szabo and members of the Woodside lab for helpful discussions. This work was supported by funds from PrioNet Canada, Alberta Prion Research Institute, Alberta Innovates (AI) Technology Solutions, AI Health Solutions, Natural Sciences and Engineering Research Council, and the National Institute for Nanotechnology.

1. Plotkin SS, Onuchic JN (2002) Understanding protein folding with energy landscape theory. Part I: Basic concepts. *Q Rev Biophys* 35:111–167.
2. Oliveberg M, Wolynes PG (2005) The experimental survey of protein-folding energy landscapes. *Q Rev Biophys* 38:245–288.
3. Thirumalai D, O'Brien EP, Morrison G, Hyeon C (2010) Theoretical perspectives on protein folding. *Annu Rev Biophys* 39:159–183.
4. Hanggi P, Talkner P, Borkovec M (1990) Reaction-rate theory—50 years after Kramers. *Rev Mod Phys* 62:251–341.
5. Chaudhury S, Makarov DE (2010) A harmonic transition state approximation for the duration of reactive events in complex molecular rearrangements. *J Chem Phys* 133:034118.
6. Chung HS, Louis JM, Eaton WA (2009) Experimental determination of upper bound for transition path times in protein folding from single-molecule photon-by-photon trajectories. *Proc Natl Acad Sci USA* 106:11837–11844.
7. Hummer G (2004) From transition paths to transition states and rate coefficients. *J Chem Phys* 120:516–523.
8. Best RB, Hummer G (2010) Coordinate-dependent diffusion in protein folding. *Proc Natl Acad Sci USA* 107:1088–1093.
9. Dudko OK, Hummer G, Szabo A (2006) Intrinsic rates and activation free energies from single-molecule pulling experiments. *Phys Rev Lett* 96:108101.
10. Kubelka J, Hofrichter J, Eaton WA (2004) The protein folding 'speed limit'. *Curr Opin Struct Biol* 14:76–88.
11. Nettels D, Gopich IV, Hoffmann A, Schuler B (2007) Ultrafast dynamics of protein collapse from single-molecule photon statistics. *Proc Natl Acad Sci USA* 104:2655–2660.
12. Krieger F, Fierz B, Bieri O, Drewello M, Kiefhaber T (2003) Dynamics of unfolded polypeptide chains as model for the earliest steps in protein folding. *J Mol Biol* 332:265–274.
13. Lapidus LJ, Eaton WA, Hofrichter J (2000) Measuring the rate of intramolecular contact formation in polypeptides. *Proc Natl Acad Sci USA* 97:7220–7225.
14. Chen Y, Parrini C, Taddei N, Lapidus LJ (2009) Conformational properties of unfolded HypF-N. *J Phys Chem B* 113:16209–16213.
15. Hagen SJ, Hofrichter J, Szabo A, Eaton WA (1996) Diffusion-limited contact formation in unfolded cytochrome C: Estimating the maximum rate of protein folding. *Proc Natl Acad Sci USA* 93:11615–11617.
16. Chung HS, McHale K, Louis JM, Eaton WA (2012) Single-molecule fluorescence experiments determine protein folding transition path times. *Science* 335:981–984.
17. Borgia A, Williams PM, Clarke J (2008) Single-molecule studies of protein folding. *Annu Rev Biochem* 77:101–125.
18. 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.
19. Woodside MT, et al. (2006) Nanomechanical measurements of the sequence-dependent folding landscapes of single nucleic acid hairpins. *Proc Natl Acad Sci USA* 103:6190–6195.
20. Gebhardt JC, Bornschlogl T, Rief M (2010) Full distance-resolved folding energy landscape of one single protein molecule. *Proc Natl Acad Sci USA* 107:2013–2018.
21. Woodside MT, et al. (2006) Direct measurement of the full, sequence-dependent folding landscape of a nucleic acid. *Science* 314:1001–1004.
22. Gupta AN, et al. (2011) Experimental validation of free-energy-landscape reconstruction from non-equilibrium single-molecule force spectroscopy measurements. *Nat Phys* 7:631–634.
23. Harris NC, Song Y, Kiang C-H (2007) Experimental free energy surface reconstruction from single-molecule force spectroscopy using Jarzynski's equality. *Phys Rev Lett* 99:068101.
24. Hummer G, Szabo A (2001) Free energy reconstruction from nonequilibrium single-molecule pulling experiments. *Proc Natl Acad Sci USA* 98:3658–3661.
25. Wildegger G, Liemann S, Glockshuber R (1999) Extremely rapid folding of the C-terminal domain of the prion protein without kinetic intermediates. *Nat Struct Mol Biol* 6:550–553.
26. Caughey B, Baron GS, Chesebro B, Jeffrey M (2009) Getting a grip on prions: Oligomers, amyloids, and pathological membrane interactions. *Annu Rev Biochem* 78:177–204.
27. Neupane K, Yu H, Foster DA, Wang F, Woodside MT (2011) Single-molecule force spectroscopy of the add adenine riboswitch relates folding to regulatory mechanism. *Nucleic Acids Res* 39:7677–7687.
28. Yu H, et al. (2012) Direct observation of multiple misfolding pathways in a single prion protein molecule. *Proc Natl Acad Sci USA* 109:5283–5288.
29. James TL, et al. (1997) Solution structure of a 142-residue recombinant prion protein corresponding to the infectious fragment of the scrapie isoform. *Proc Natl Acad Sci USA* 94:10086–10091.
30. Hoffmann A, Woodside MT (2011) Signal-pair correlation analysis of single-molecule trajectories. *Angew Chem Int Ed* 50:12643–12646.
31. Bell GI (1978) Models for the specific adhesion of cells to cells. *Science* 200:618–627.
32. Dudko OK, Graham TGW, Best RB (2011) Locating the barrier for folding of single molecules under an external force. *Phys Rev Lett* 107:208301.
33. Popa I, Fernandez JM, Garcia-Manyes S (2011) Direct quantification of the attempt frequency determining the mechanical unfolding of ubiquitin protein. *J Biol Chem* 286:31072–31079.
34. Junker JP, Ziegler F, Rief M (2009) Ligand-dependent equilibrium fluctuations of single calmodulin molecules. *Science* 323:633–637.
35. Rief M, Gautel M, Oesterhelt F, Fernandez JM, Gaub HE (1997) Reversible unfolding of individual titin immunoglobulin domains by AFM. *Science* 276:1109–1112.
36. Dietz H, Rief M (2004) Exploring the energy landscape of GFP by single-molecule mechanical experiments. *Proc Natl Acad Sci USA* 101:16192–16197.
37. Hart T, et al. (2009) Folding kinetics of the human prion protein probed by temperature jump. *Proc Natl Acad Sci USA* 106:5651–5656.
38. Julien O, Chatterjee S, Thiessen A, Graether SP, Sykes BD (2009) Differential stability of the bovine prion protein upon urea unfolding. *Protein Sci* 18:2172–2182.
39. Waldauer SA, Bakajin O, Lapidus LJ (2010) Extremely slow intramolecular diffusion in unfolded protein L. *Proc Natl Acad Sci USA* 107:13713–13717.
40. Carrion-Vazquez M, et al. (1999) Mechanical and chemical unfolding of a single protein: A comparison. *Proc Natl Acad Sci USA* 96:3694–3699.
41. Hummer G, Szabo A (2003) Kinetics from nonequilibrium single-molecule pulling experiments. *Biophys J* 85:5–15.
42. Shaw DE, et al. (2010) Atomic-level characterization of the structural dynamics of proteins. *Science* 330:341–346.
43. Ivankov DN, et al. (2003) Contact order revisited: Influence of protein size on the folding rate. *Protein Sci* 12:2057–2062.
44. Li MS, Klimov DK, Thirumalai D (2004) Thermal denaturation and folding rates of single domain proteins: Size matters. *Polymer* 45:573–579.
45. Baker D (2000) A surprising simplicity to protein folding. *Nature* 405:39–42.
46. Greenleaf WJ, Woodside MT, Abbondanzieri EA, Block SM (2005) Passive all-optical force clamp for high-resolution laser trapping. *Phys Rev Lett* 95:208102.
47. Neuman KC, Block SM (2004) Optical trapping. *Rev Sci Instrum* 75:2787–2809.
48. Jansson PA (1997) *Deconvolution of Images and Spectra* (Academic Press, New York), 2nd Ed.
49. Elms PJ, Chodera JD, Bustamante C, Marqusee S (2012) The molten globule state is unusually deformable under mechanical force. *Proc Natl Acad Sci USA* 109:3796–3801.

Supporting Information

Yu et al. 10.1073/pnas.1206190109

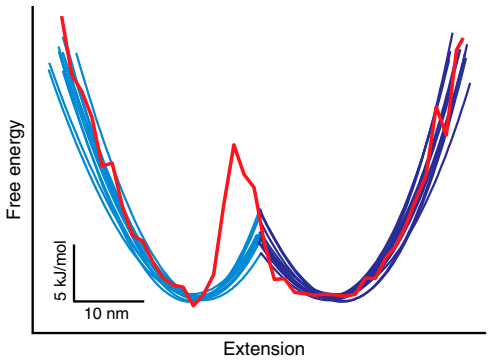

Fig. S1. Comparison of reconstructed landscape to WLC energies. The outer walls of the potential wells of the reconstructed landscape (red) match the energies of the WLC fits to the folded (cyan) and unfolded (blue) states, but the barrier is quite different, indicating that the latter is dominated by the properties of the protein unfolding. The WLC stretching energies were found by integrating the FEC fits as in Fig. 3*A*, then tilted under a force of $F_{1/2}$ for comparison to the reconstructed landscape, showing only the lowest free energy at each extension.

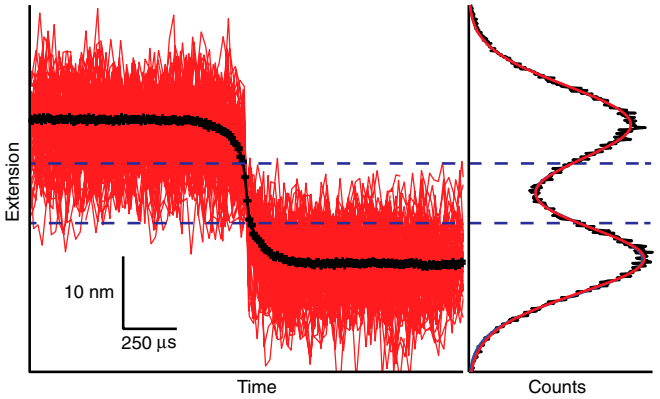

Fig. S2. Direct estimate of an upper bound for the transition time. Segments of the unfiltered constant-force extension trajectories, centered around the folding transitions, were extracted from the full trajectories, aligned on the center of the transition (red: 100 representative transitions), and averaged to remove Brownian fluctuations (black: average of over 3,300 transitions). An upper bound for the transition time of approximately 40 μs was estimated from the average time required to move between the inflection points of the folded and unfolded distributions (dashed lines).

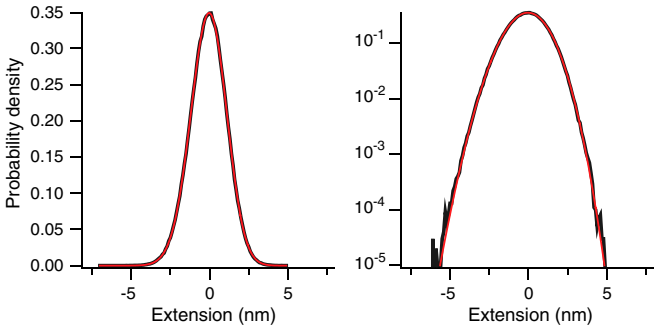

Fig. S3. Point-spread function. A typical point-spread function (black), measured under the same conditions as the FECs using the reference construct consisting of DNA handles only, shows a quasi-Gaussian profile, well fit (red) by a phenomenological model from Yu et al. (1). The log scale graph (right) shows the slight asymmetry in the point-spread function.

1 Yu H, et al. (2012) Direct observation of multiple misfolding pathways in a single prion protein molecule. *Proc Natl Acad Sci USA* 109:5283–5288.

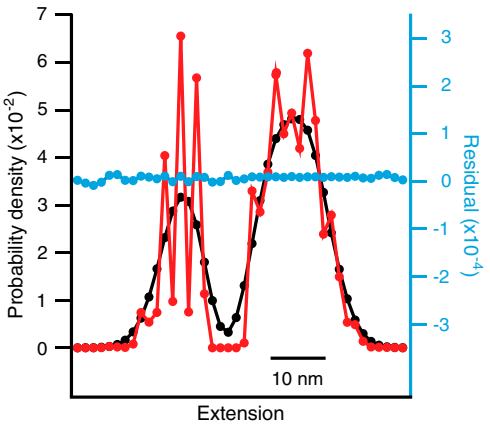

Fig. S4. Extension probability deconvolution. The probability density of the extension at $F_{1/2}$ after deconvolution (red) of the initial solution (black). There is little residual error (cyan). We note that the width of the folded state distribution is only somewhat reduced after deconvolution. This may reflect structural dynamics of the nominally folded state, such as fluctuations arising from the relatively unstable β-strand 1 (1).

1 Julien O, Chatterjee S, Thiessen A, Graether SP, Sykes BD (2009) Differential stability of the bovine prion protein upon urea unfolding. *Protein Sci* 18:2172–2182.

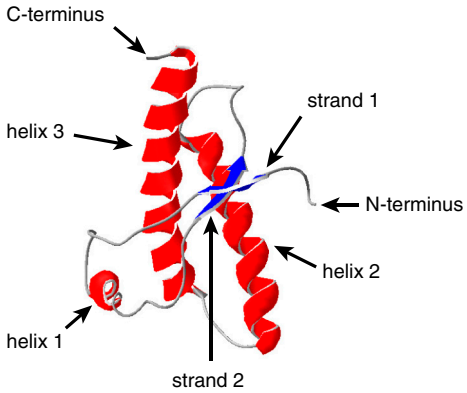

Fig. S5. Native structure of hamster PrP. The native structure (1) consists of 2 short β-strands near the N terminus (strands 1 and 2), 1 short α-helix (helix 1), and two long α-helices (helices 2 and 3) at the C-terminal end of the protein.

1 James TL, et al. (1997) Solution structure of a 142-residue recombinant prion protein corresponding to the infectious fragment of the scrapie isoform. *Proc Natl Acad Sci USA* 94:10086–10091.