Protein folding trajectories can be described quantitatively by one-dimensional diffusion over measured energy landscapes

Krishna Neupane[1], Ajay P. Manuel[1] and Michael T. Woodside[1,2]*

Protein folding features a diffusive search over a multi-dimensional energy landscape in conformational space for the minimum-energy structure[1]. Experiments, however, are usually interpreted in terms of a one-dimensional (1D) projection of the full landscape onto a practical reaction coordinate. Although simulations have shown that folding kinetics can be described well by diffusion over a 1D projection[2,3], 1D approximations have not yet been fully validated experimentally. We used folding trajectories of single molecules held under tension in optical tweezers to compare the conditional probability of being on a transition path[4], calculated from the trajectory[5], with the prediction for ideal 1D diffusion over the measured 1D landscape[6], calculated from committor statistics[7,8]. We found good agreement for the protein PrP (refs 9,10) and for one of the structural transitions in a leucine-zipper coiled-coil[11], but not for a second transition in the coiled-coil, owing to poor reaction-coordinate quality[12]. These results show that 1D descriptions of folding can indeed be good, even for complex tertiary structures. More fundamentally, they also provide a fully experimental validation of the basic physical picture of folding as diffusion over a landscape.

Protein folding is justly renowned for its combinatorial complexity: not only is it driven by a wide range of different and often competing interactions, but there are hundreds or even thousands of degrees of freedom related to the bond angles in the polypeptide chain and the motions of the solvent[1]. The full energy landscape underlying folding thus has a very high dimensionality. Measuring the dynamics in each degree of freedom represents a supreme technical challenge that remains beyond current capabilities. Instead, through necessity, experiments typically monitor the folding dynamics in a much-reduced projection of the full dynamical space, most commonly using a single dimension associated with a convenient observable (for example, radius of gyration, end-to-end extension, and so on), which becomes the collective 'reaction coordinate' used to describe the progress of the folding[13]. The conformational dynamics are then described in terms of diffusion along this reaction coordinate.

Computational simulations suggest that low-dimensional reductions can generally provide a valid description of the folding[14]; indeed, simulations of a variety of small proteins show that kinetic properties such as rates and transition-path times can be accounted for quantitatively by even a 1D projection, with the observed kinetics matching the predictions for diffusion over the 1D energy profile[2,3]. Experimentally, simple 1D approximations have found reasonable empirical success, especially for smaller proteins[10,15], although counterexamples exist that are likely to
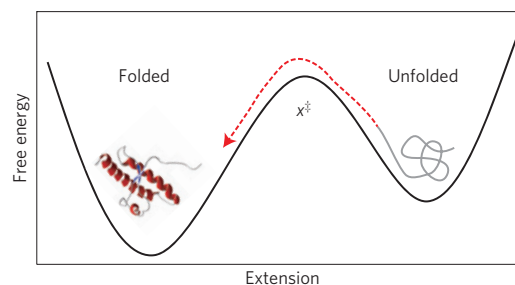

Figure 1 | Transition paths. Transition paths represent the brief part of the trajectory (red) spent crossing the barrier between folded and unfolded states, in contrast to most conformational fluctuations that stay within a potential well and are non-productive (grey).

require multiple dimensions to account for the observed behaviour (for example, proteins with knots[16] or multiple pathways[17]). Nevertheless, there are many potential concerns with simple 1D descriptions. Even if a low-dimensional approximation is valid, a 1D approximation may not be[18,19]. The projection onto the reaction coordinate may also be sub-optimal, incompletely capturing the full dynamics during folding; such 'bad' reaction coordinates may lead to non-Markovian dynamics, poor predictions and incorrect interpretations[20–22]. Reaction-coordinate quality is rarely tested in protein-folding experiments, however, with only a handful of examples published[12]. Moreover, even if a reaction coordinate is known to be good, it has not yet been directly shown that the dynamics along such a coordinate agree quantitatively with 1D diffusion over the measured energy profile.

One way to address this question is through analysis of the transition paths during folding. Transition paths represent the purely reactive portions of the folding trajectory, the fleeting moments when the protein changes from one conformation to another, in contrast to the non-productive fluctuations that comprise most of the trajectory. For two-state folding, where the projected energy profile consists of two wells separated by a barrier, the transition paths are those parts of the trajectory crossing the barrier from one well to the other (Fig. 1). The conditional probability that the molecule is on a transition path at a given reaction-coordinate value, $p(\mathrm{TP}|x)$, provides both a test of reaction-coordinate quality and—if the energy profile is known—whether the dynamics truly reflect 1D diffusion over this profile[4].

We recently showed how to apply such transition-path analysis to single-molecule force spectroscopy (SMFS) measurements[5]. Single-molecule approaches are particularly well suited to characterizing

[1]Department of Physics, University of Alberta, Edmonton, Alberta T6G 2E1, Canada. [2]National Institute for Nanotechnology, National Research Council, Edmonton, Alberta T6G 2M9, Canada. *e-mail: Michael.Woodside@ualberta.ca
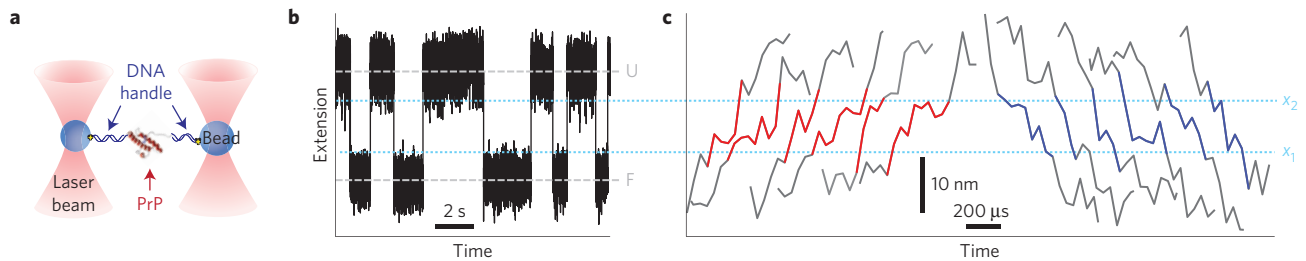
1

Figure 2 | Transition-path identification in force spectroscopy measurements of PrP folding. a, Schematic of force spectroscopy measurements of the prion protein. A single PrP molecule is attached to double-stranded DNA handles bound to beads held in optical traps. **b**, A portion of an equilibrium folding trajectory of PrP under constant force shows characteristic jumps between folded (F: low extension) and unfolded (U: high extension) states. **c**, A set of unfolding (red) and refolding (blue) transition paths passing between x_1 and x_2 (dotted lines) as they cross over the barrier between folded and unfolded states, taken from the complete trajectory.

transition paths[23,24], because the latter are inherently a property of individual molecules. In SMFS, tension is applied to the ends of a single molecule, and its extension—the reaction coordinate—is measured as the conformation fluctuates[25]. Transition paths can be identified clearly because the extension can be measured with high precision. SMFS also provides effective ways for measuring the folding landscape[6]. Applying transition-path analysis to two-state DNA hairpins, the folding dynamics were found to match expectations for 1D diffusion[5]. Proteins pose a greater conceptual challenge for applying 1D descriptions, however, because of their complex tertiary structure.

We first analysed folding trajectories of the prion protein PrP, which was previously shown to have a two-state native folding pathway, although it can also form transient misfolded states[9]. Natively folded PrP molecules attached covalently at each terminus to DNA handles were bound to beads held in high-resolution optical tweezers (Fig. 2a). Trajectories of the molecular extension measured in equilibrium at a constant force of 9–10 pN, near the value at which folded and unfolded states were equally likely, showed multiple transitions between the native and unfolded states (Fig. 2b).

We calculated $p(\mathrm{TP}|x)$ using the Bayesian relation[4]

$$p(\mathrm{TP}|x) = P(x|\mathrm{TP})p(\mathrm{TP})/P(x) \qquad (1)$$

where $P(x)$ is the equilibrium distribution of extension values in the complete trajectory, $P(x|\mathrm{TP})$ is the distribution of extension values along only the transition paths, and $p(\mathrm{TP})$ is the fraction of time spent on transition paths. Transition paths were identified as the parts of the trajectory (Fig. 2c, red and blue) transiting between two boundaries, x_1 and x_2 (Fig. 2c, dotted lines), chosen to bracket the barrier region between the folded and unfolded states. As described previously, this analysis must be corrected for instrumental effects on the measurement (here, the mechanical compliance)[5]. An additional complication in the case of PrP is the presence of misfolded states[9], which contribute to $P(x)$ even though they are excluded from $P(x|\mathrm{TP})$ because they do not transit the full distance between folded and unfolded states. We corrected simultaneously for both compliance effects and misfolded states by replacing $P(x)$ with the probability distribution $P_0(x)$ obtained from the 1D energy profile for PrP folding, calculated from non-equilibrium pulling curves[10]. Here, misfolded states were deselected kinetically by the pulling regime and compliance effects were removed through deconvolution.

The result for $p(\mathrm{TP}|x)$ is highly peaked, reaching a maximum value of about 0.45 (Fig. 3a, black). These features are indicative of a good reaction coordinate[3]; the two states are well resolved along the coordinate, and the protein is very likely to be found on a transition path in the region between the states, leading to a highly peaked $p(\mathrm{TP}|x)$; ideally, $p(\mathrm{TP}|x)$ should reach 0.5 at the barrier between

the states. Indeed, the location of the energy barrier, $x^{\ddagger}$ (Fig. 3a, dashed line), found from the reconstructed landscape (Fig. 3a, blue), was very close to the peak in $p(\mathrm{TP}|x)$, well within the resolution of the reconstruction.

Having established the quality of extension as a reaction coordinate, we next tested whether the statistics of the transition paths were well described by 1D diffusion over the measured landscape. In the case of ideal diffusion[4], one should have $p(\mathrm{TP}|x) = 2p_{\mathrm{fold}}(x)[1 - p_{\mathrm{fold}}(x)]$, where $p_{\mathrm{fold}}(x)$ is the committor, the probability that when the molecule starts at x it will reach the folded state before the unfolded state[7]. For a two-state system, $p_{\mathrm{fold}}(x)$ is approximately 0 near the unfolded state, 1 near the folded state, and 1/2 at the top of the barrier. In the case of diffusive dynamics along a 1D energy profile $G(x)$, and assuming for simplicity a constant diffusion coefficient[4,8], $p_{\mathrm{fold}}(x)$ is given by[8]

$$p_{\mathrm{fold}}(x) = \int_x^{x_{\mathrm{u}}} \mathrm{d}x' \, e^{\beta G(x')} \Big/ \int_{x_{\mathrm{f}}}^{x_{\mathrm{u}}} \mathrm{d}x' \, e^{\beta G(x')} \qquad (2)$$

Using the result (Fig. 3b, orange) to calculate $\Phi(x) = 2p_{\mathrm{fold}}(x)[1 - p_{\mathrm{fold}}(x)]$ (Fig. 3a, orange), we found that the $\Phi(x)$ agreed surprisingly well with $p(\mathrm{TP}|x)$: the location, height and width of the two peaks were all very similar, well within the limits of experimental uncertainty. The folding dynamics are thus well described by 1D diffusion over the measured landscape, the central result of this work. We confirmed this result using an alternative approach, calculating $p_{\mathrm{fold}}(x)$ directly from the extension trajectory rather than from the 1D landscape (see Methods): for 1D diffusion, both methods should yield the same result[8]. Indeed, the landscape p_{fold} (Fig. 3b, orange) agreed very well with the trajectory p_{fold} (Fig. 3b, black), confirming that 1D diffusion over the reconstructed landscape describes the dynamics well.

The notion that 1D approximations are plausible was supported by previous work showing that PrP folding kinetics were consistent with Kramers' theory over several orders of magnitude[10], and that the kinetics of other proteins were similarly consistent with 1D models[15]. Our new results provide a deeper and more direct test of protein folding as a diffusive search over an energy landscape, showing that not only the kinetics but more importantly the statistics of the transition paths—the most important parts of the folding trajectories—match predictions for 1D diffusion over the measured landscape. The quantitative nature of the agreement is remarkable, given the size and complexity of the structure being formed: 104 amino acids forming 3 helices, 2 strands and multiple loops.

To test whether a similar result holds for other proteins, we analysed equilibrium folding trajectories of a coiled-coil leucine zipper, which in contrast to PrP exhibited three-state behaviour with an obligate intermediate[11] (Fig. 4a). Treating the folding as sequential two-state transitions, previous work found that
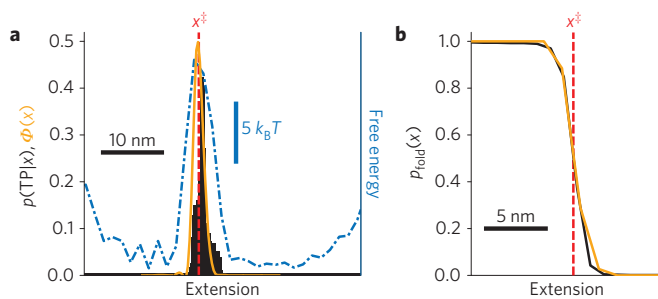
Figure 3 | Transition-path and committor analyses of PrP folding reveal 1D diffusion. a, The conditional transition-path probability $p(\mathrm{TP}|x)$ (black) peaks at a value of ~ 0.45 at the location of the barrier (red), as determined from the energy landscape (blue). $p(\mathrm{TP}|x)$ agrees very well with $\Phi(x)$ (orange) calculated from the committor $p_{\mathrm{fold}}(x)$, indicating that the dynamics are well described by 1D diffusion. **b**, $P_{\mathrm{fold}}(x)$ calculated from the energy landscape using equation (2) (orange) agrees with $p_{\mathrm{fold}}(x)$ calculated directly from equilibrium extension trajectories using equation (3) (black), confirming 1D diffusion. Both pass through 1/2 near the position of the barrier (red).
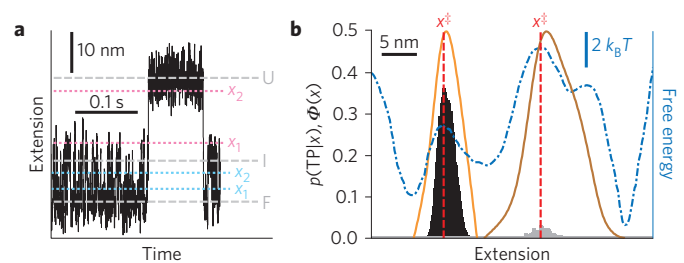

Figure 4 | Transition-path analysis of the leucine zipper. a, Portion of the folding trajectory, showing three states: folded (F), intermediate (I) and unfolded (U). Boundaries x_1 and x_2 (dotted lines) define the transition paths for the F↔I (cyan) and I↔U (purple) transitions. **b**, $p(\mathrm{TP}|x)$ for F↔I (black) agrees reasonably well with $\Phi(x)$ (orange), confirming 1D diffusion, but for I↔U, $p(\mathrm{TP}|x)$ (grey) does not agree with $\Phi(x)$ (brown). The low peak for $p(\mathrm{TP}|x)$ indicates that extension is not a good reaction coordinate for this transition.

end-to-end extension was not a good reaction coordinate for the I↔U transition[12], making this protein an interesting test case for transition-path analysis. We defined boundaries x_1 and x_2 for each transition as above, calculating $p(\mathrm{TP}|x)$ from equation (1) for each transition (Fig. 4b, black: F↔I, grey: I↔U), using the compliance-deconvolved[11] distribution $P_0(x)$, as well as $\Phi(x)$ from equation (2) (Fig. 4b, orange: F↔I, brown: I↔U), using the deconvolved landscape (Fig. 4b, blue). Reasonable agreement was found for F↔I, confirming that its folding dynamics are well described by diffusion over the measured energy profile. However, the test failed for I↔U: there were more non-reactive fluctuations into the barrier region than expected (even after accounting for compliance effects), depressing $p(\mathrm{TP}|x)$. Extension was thus not a good reaction coordinate for this transition, as found previously[12].

The I↔U transition provides a counterexample where diffusion over the measured landscape does not describe the observed dynamics well, highlighting the importance of reaction-coordinate quality. Whereas the reaction coordinate can be engineered in computations to ensure optimal low-dimensional descriptions of the dynamics[13,14], in experiments it is imposed by the choice of assay, without any particular privilege; here, for example, the applied force does not ensure that the reaction coordinate is always good. Changing the pulling axis may permit reaction-coordinate optimization[22], but such an optimization has never been done experimentally.

The ability to capture the folding dynamics on a single dimension is usually understood intuitively as indicating a clear separation of timescales between a single slow coordinate that dominates the behaviour and faster dynamics along all other coordinates[18,26] (although this explanation is not formally dispositive[21]). An important implication is that the transition paths probably funnel through a single, well-defined region of phase space acting as the transition-state ensemble (multiple pathways with different diffusivities would be likely to prevent quantitative agreement with 1D diffusion), suggesting that the transition-state ensemble can be identified in a physically meaningful way[22].

It will be instructive to apply transition-path analysis more widely, to understand better the limits of 1D descriptions. It will be particularly interesting to analyse proteins exhibiting evidence of multiple competing pathways[17], distributions of barriers[27] and dynamic disorder[19], forms of 'anomalous' diffusion such as subdiffusion of the backbone[28], or particularly complex structures such as knots[16], to obtain a quantitative look at how 1D descriptions

break down (and the quality of the reaction coordinate in these cases). In addition to surveying different proteins, transition-path analysis will also be valuable for going beyond the current study, which was limited to equilibrium measurements under tension, to test whether the result depends on the mode of denaturation (for example, force, temperature, chemical denaturant), the probe used, or other measurement conditions (for example, equilibrium versus non-equilibrium), all of which may alter key folding properties such as the dominant pathways and barriers and hence the effective dimensionality[29,30]. Such studies should help establish how widely 1D landscapes can be applied and under what conditions 1D descriptions fail.

Methods

Methods and any associated references are available in the online version of the paper.

References

1. Bryngelson, J. D. & Wolynes, P. G. Spin glasses and the statistical mechanics of protein folding. *Proc. Natl Acad. Sci. USA* **84,** 7524–7528 (1987).
2. Socci, N. D., Onuchic, J. N. & Wolynes, P. G. Diffusive dynamics of the reaction coordinate for protein folding funnels. *J. Chem. Phys.* **104,** 5860–5868 (1996).
3. Zheng, W. & Best, R. B. Reduction of all-atom protein folding dynamics to one-dimensional diffusion. *J. Phys. Chem. B* **119,** 15247–15255 (2015).
4. Best, R. B. & Hummer, G. Reaction coordinates and rates from transition paths. *Proc. Natl Acad. Sci. USA* **102,** 6732–6737 (2005).
5. Neupane, K., Manuel, A. P., Lambert, J. & Woodside, M. T. Transition-path probability as a test of reaction-coordinate quality reveals DNA hairpin folding is a one-dimensional diffusive process. *J. Phys. Chem. Lett.* **6,** 1005–1010 (2015).
6. Woodside, M. T. & Block, S. M. Reconstructing folding energy landscapes by single-molecule force spectroscopy. *Annu. Rev. Biophys.* **43,** 19–39 (2014).
7. Du, R., Pande, V. S., Grosberg, A. Y., Tanaka, T. & Shakhnovich, E. S. On the transition coordinate for protein folding. *J. Chem. Phys.* **108,** 334–350 (1998).
8. Chodera, J. D. & Pande, V. S. Splitting probabilities as a test of reaction coordinate choice in single-molecule experiments. *Phys. Rev. Lett.* **107,** 098102 (2011).
9. Yu, H. *et al*. Direct observation of multiple misfolding pathways in a single prion protein molecule. *Proc. Natl Acad. Sci. USA* **109,** 5283–5288 (2012).
10. Yu, H. *et al*. Energy landscape analysis of native folding of the prion protein yields the diffusion constant, transition path time, and rates. *Proc. Natl Acad. Sci. USA* **109,** 14452–14457 (2012).
11. Gebhardt, J. C. M., Bornschloegl, T. & Rief, M. Full distance-resolved folding energy landscape of one single protein molecule. *Proc. Natl Acad. Sci. USA* **107,** 2013–2018 (2010).
12. Morrison, G., Hyeon, C., Hinczewski, M. & Thirumalai, D. Compaction and tensile forces determine the accuracy of folding landscape parameters from single molecule pulling experiments. *Phys. Rev. Lett.* **106,** 138102 (2011).

13. Best, R. B. & Hummer, G. Diffusion models of protein folding. *Phys. Chem. Chem. Phys.* **13,** 16902–16911 (2011).
14. Das, P., Moll, M., Stamati, H., Kavraki, L. E. & Clementi, C. Low-dimensional, free-energy landscapes of protein-folding reactions by nonlinear dimensionality reduction. *Proc. Natl Acad. Sci. USA* **103,** 9885–9890 (2006).
15. Cellmer, T., Henry, E. R., Hofrichter, J. & Eaton, W. A. Measuring internal friction of an ultrafast-folding protein. *Proc. Natl Acad. Sci. USA* **105,** 18320–18325 (2008).
16. Kamitori, S. A real knot in protein. *J. Am. Chem. Soc.* **118,** 8945–8946 (1996).
17. Udgaonkar, J. B. Multiple routes and structural heterogeneity in protein folding. *Annu. Rev. Biophys.* **37,** 489–510 (2008).
18. Suzuki, Y. & Dudko, O. K. Single-molecule rupture dynamics on multidimensional landscapes. *Phys. Rev. Lett.* **104,** 048101 (2010).
19. Hyeon, C., Hinczewski, M. & Thirumalai, D. Evidence of disorder in biological molecules from single molecule pulling experiments. *Phys. Rev. Lett.* **112,** 138101 (2014).
20. Plotkin, S. S. & Wolynes, P. G. Non-Markovian configurational diffusion and reaction coordinates for protein folding. *Phys. Rev. Lett.* **80,** 5015–5018 (1998).
21. Makarov, D. E. Interplay of non-Markov and internal friction effects in the barrier crossing kinetics of biopolymers: insights from an analytically solvable model. *J. Chem. Phys.* **138,** 014102 (2013).
22. Dudko, O. K., Graham, T. G. W. & Best, R. B. Locating the barrier for folding of single molecules under an external force. *Phys. Rev. Lett.* **107,** 208301 (2011).
23. Chung, H. S., McHale, K., Louis, J. M. & Eaton, W. A. Single-molecule fluorescence experiments determine protein folding transition path times. *Science* **335,** 981–984 (2012).
24. Neupane, K. *et al.* Transition path times for nucleic acid folding determined from energy-landscape analysis of single-molecule trajectories. *Phys. Rev. Lett.* **109,** 068102 (2012).
25. Ritchie, D. B. & Woodside, M. T. Probing the structural dynamics of proteins and nucleic acids with optical tweezers. *Curr. Opin. Struct. Biol.* **34,** 43–51 (2015).
26. Berezhkovskii, A. M. & Zitserman, V. Y. Activated rate processes in a multidimensional case. A new solution of the Kramers problem. *Phys. Stat. Mech. Appl.* **166,** 585–621 (1990).
27. Brujic, J., Hermans, R. I., Walther, K. A. & Fernandez, J. M. Single-molecule force spectroscopy reveals signatures of glassy dynamics in the energy landscape of ubiquitin. *Nature Phys.* **2,** 282–286 (2006).
28. Milanesi, L. *et al.* Measurement of energy landscape roughness of folded and unfolded proteins. *Proc. Natl Acad. Sci. USA* **109,** 19563–19568 (2012).
29. Guinn, E. J., Jagannathan, B. & Marqusee, S. Single-molecule chemo-mechanical unfolding reveals multiple transition state barriers in a small single-domain protein. *Nature Commun.* **6,** 6861 (2015).
30. Hyeon, C., Morrison, G., Pincus, D. L. & Thirumalai, D. Refolding dynamics of stretched biopolymers upon force quench. *Proc. Natl Acad. Sci. USA* **106,** 20288–20293 (2009).

Author contributions

M.T.W. designed the research; K.N. and A.P.M. performed the research; all authors wrote the paper.

Additional information

Supplementary information is available in the online version of the paper. Reprints and permissions information is available online at www.nature.com/reprints. Correspondence and requests for materials should be addressed to M.T.W.

Competing financial interests

The authors declare no competing financial interests.

Methods

Sample preparation and measurement. Samples of truncated hamster prion protein, PrP(90–231), were expressed, purified, refolded and attached covalently at each terminus to double-stranded DNA handles roughly 1 kilobase in length as described previously[9]. Protein–DNA chimaeras were bound specifically to 600-nm- and 810-nm-diameter polystyrene beads labelled with avidin and anti-digoxigenin, respectively. Samples were placed in 50 mM Mops, pH 7.0, 200 mM KCl and an oxygen scavenging system[9] for measurement using a custom dual-trap optical tweezers apparatus described previously[31]. Extension trajectories were measured at equilibrium under a constant force of 9–10 pN, maintained by a passive force clamp to avoid artefacts in the transition region[32], sampled at 50 kHz or 20 kHz and filtered online at the Nyquist frequency.

Folding trajectories of the leucine zipper were generously provided by C. Gebhardt and M. Rief. The sample preparation and measurement conditions have been described previously[11]. Briefly, the construct (consisting of three tandem repeats of the GCN4 leucine zipper) was attached to labelled DNA handles as done for PrP. Protein–DNA chimaeras were bound specifically to polystyrene beads held in dual-trap optical tweezers. Extension trajectories were measured in phosphate-buffered saline at equilibrium with a constant trap position, using a pretension such that all three states were occupied, sampled at 100 kHz and filtered at 20 kHz. Note that the lowest-extension state in these trajectories is not the fully folded native state of the leucine zipper, which unfolds at a much lower force, but rather a partially folded intermediate. For simplicity of labelling, however, we treat it here as the folded state under tension.

Transition-path analysis. Transition paths were identified as the parts of the trajectory traversing between two boundaries, x_1 and x_2, respectively near the folded and unfolded states, chosen so as to allow the transition paths to be identified clearly while excluding most of the trajectory spent on non-productive attempts to cross the barrier. They were therefore located on the shoulders of the peaks in $P(x)$ corresponding respectively to the folded and unfolded states, on the side of the peaks towards the barrier region, at the inflection points of Gaussian fits to the peaks in $P(x)$. In the case of the leucine zipper, the two sequential transitions were analysed separately, as independent two-state transitions.

For PrP, the compliance-corrected distribution $P_0(x)$ was found as the equilibrium distribution expected from Boltzmann's formula using the energy landscape for native folding calculated from non-equilibrium force–extension curves via the Hummer–Szabo formalism[33,34], after deconvolution of compliance effects[10]. The resolution of the landscape reconstruction for PrP was 1.4 nm (ref. 10). For the leucine zipper, $P_0(x)$ was found by empirical deconvolution of $P(x)$ using the measured point-spread function, taking into account the position-dependence of the point-spread function arising from the constant-trap-position measurement modality[11]. To maintain the normalization of $p(\text{TP}|x)$, $p(\text{TP})$ was multiplied by $\int_{x_1}^{x_2} P_0(x)/P(x)\mathrm{d}x$, to correct for the fraction of the statistical weight in the transition region that was induced by the instrumental compliance, as described previously[5]. For PrP, $p(\text{TP}|x)$ was calculated for 3,759 transitions; for the leucine zipper, 32,689 F↔I transitions and 283 I↔U transitions were analysed.

Committor analysis. For calculating $\Phi(x)$, the splitting probability was determined from the energy profile measured for each protein (Figs 3a and 4b, blue) via equation (2). The landscape after deconvolution was used in each case, to avoid artefacts from compliance effects[35]. The boundaries x_f and x_u were chosen to be near the folded and unfolded peaks in $P(x)$; the result was insensitive to the precise choice of boundary location[35].

To calculate $p_{\text{fold}}(x)$ empirically from the extension trajectory, for comparison with the landscape-derived committor (Fig. 3b), we used[8]

$$p_{\text{fold}}(x_0) = \int_0^T \delta(x_0 - x(t))c(t)\mathrm{d}t \Big/ \int_0^T \delta(x_0 - x(t))\mathrm{d}t \qquad (3)$$

where δ is the Dirac delta function and the function $c(t)$ is 1 if, in the interval after time t, the trajectory hits the folded state (represented by an absorbing boundary x_f) before it hits the unfolded state (at x_u); otherwise it is 0. As the misfolded states in the trajectory for PrP can alter p_{fold} calculated from the trajectory but are very short-lived[9], we minimized their influence by median-filtering the trajectory in a 1-ms window before calculating p_{fold}. The result was relatively insensitive to the filter window size, in the range from ~0.5 to 2 ms (Supplementary Fig. 1).

References

31. Neupane, K., Yu, H., Foster, D. A. N., Wang, F. & Woodside, M. T. Single-molecule force spectroscopy of the add adenine riboswitch relates folding to regulatory mechanism. *Nucleic Acids Res.* **39,** 7677–7687 (2011).
32. Greenleaf, W. J., Woodside, M. T., Abbondanzieri, E. A. & Block, S. M. Passive all-optical force clamp for high-resolution laser trapping. *Phys. Rev. Lett.* **95,** 208102 (2005).
33. Hummer, G. & Szabo, A. Free energy reconstruction from nonequilibrium single-molecule pulling experiments. *Proc. Natl Acad. Sci. USA* **98,** 3658–3661 (2001).
34. Gupta, A. N. *et al.* Experimental validation of free-energy-landscape reconstruction from non-equilibrium single-molecule force spectroscopy measurements. *Nature Phys.* **7,** 631–634 (2011).
35. Manuel, A. P., Lambert, J. & Woodside, M. T. Reconstructing folding energy landscapes from splitting probability analysis of single-molecule trajectories. *Proc. Natl Acad. Sci. USA* **112,** 7183–7188 (2015).