

Supporting Information

Yu et al. 10.1073/pnas.1107736109

SI Text

Protein-DNA Construct Design. The amino acid sequence of the recombinant SHaPrP(90–231) protein and corresponding structural and functional elements are displayed in Fig. S1. We note that recombinant SHaPrP(90–231) is able to form PrP^{Sc} in vitro (1).

Protein Purification. After expression, cell pellets were resuspended in buffer A (6M GnHCl, 10 mM Tris, 100 mM NaH₂PO₄, 10 mM reduced glutathione, pH 8.0) and sonicated for 2 min. The lysate was centrifuged, filtered, and then purified by FPLC (GE Healthcare) using a Ni-NTA column. After rinsing the protein bound to the column with buffer A, PrP was refolded on the column using gradient exchange from buffer A to buffer B (10 mM Tris, 100 mM NaH₂PO₄, pH 8.0). Impurities were removed by rinsing on column with buffer C (10 mM Tris, 100 mM NaH₂PO₄, 50 mM imidazole, pH 8.0). The refolded PrP was then eluted with buffer D (10 mM Tris, 100 mM NaH₂PO₄, 500 mM imidazole, pH 5.8), before dialysis into 50 mM sodium phosphate buffer, pH 7.0. Product purity and identity were assessed by SDS/PAGE and Western blotting (Antiprion(109–112) clone 3F4, Millipore).

DNA Handle Attachment. The refolded protein was first reduced with tris(2-carboxyethyl)phosphine (TCEP) in a 100:1 molar ratio with PrP for 30 min. It was then desalted by spin column (Zeba, Thermo Scientific) to remove excess TCEP and activated with 2,2'-dithiodipyridine (Sigma-Aldrich). The CD spectrum of the activated protein confirmed that it remained in the native fold. The activated protein was next reacted with sulfhydryl-labeled DNA handles of two different lengths prepared by PCR. One handle was 798-bp long and labeled at opposite termini by sulfhydryl and biotin, the other 1,261-bp long and terminally labeled with sulfhydryl and digoxigenin. The formation of the correct construct length was verified by gel electrophoresis.

Optical Trap Instrumentation. Two orthogonally polarized laser beams from the same 5W 1,064-nm laser (Spectra-Physics) were used to generate two traps. The position of each trap was controlled independently by one acousto-optic deflector (AOD, AA Optoelectronique) and one electro-optic deflector (EOD, Conoptics) in each axis. The traps were moved apart during all measurements using the EODs only. The stiffness of each trap was controlled by the AODs. Bead positions within the traps were measured by collecting the light from two orthogonally polarized beams at 633 nm scattered by the beads onto independent position-sensitive diodes (Pacific Silicon Sensors). Trap automation and data collection used custom Labview software (National Instruments).

Force-Extension Curves (FECs) of PrP Compared to Reference Construct. Individual FEC measurements of the PrP construct are shown in Fig. S2 for unfolding of PrP (red), where the force on the protein is ramped up until it unfolds, and refolding of the same molecule (blue), where the force on the protein is ramped back down until the protein refolds. The folding transition is seen as a sudden jump in extension and force as the extension of the molecule changes. For comparison, a FEC measurement (black) of the reference construct containing no protein, only DNA handles, shows just the worm-like chain (WLC) elasticity of DNA, without any unfolding features.

FEC Analysis of Contour Length. To determine the change in contour length (ΔL_c) upon unfolding and refolding, the aligned FECs were fit to an extensible WLC model (2):

$$F(x) = \frac{k_B T}{L_p} \left[\frac{1}{4} \left(1 - \frac{x}{L_c} + \frac{F}{K} \right)^{-2} - \frac{1}{4} + \frac{x}{L_c} - \frac{F}{K} \right],$$

where L_p is the persistence length of the polymer, L_c is the contour length, K is the elastic modulus, and k_B is the Boltzmann constant. The crystallographic contour length of an amino acid, $L_c^{aa} = 0.36$ nm/aa (3), was used to convert ΔL_c into the number of amino acids unfolded (n_{aa}), rather than the slightly larger value reported from high-force pulling experiments using an atomic force microscope (4), according to the equation $n_{aa} = (\Delta L_c + \Delta d_T)/L_c^{aa}$. Here we explicitly included the term Δd_T to take into account the fact that as the structure changes, the distance between the two termini of the structured parts of the protein (the points at which force is instantaneously being applied) changes, too. In the case of the unfolded state, $d_T = 0$, whereas for the native state, $d_T = 3.1$ nm is determined from the NMR structure of the protein (5).

Constant-Force Measurements. Constant force was verified during passive force clamp measurements using the force recorded in the harmonic trap, as described previously (6). As shown in Fig. S3, the force did not jump between two different levels as would be expected if the force were changing between the folded and unfolded states (7); instead, only fluctuations were observed. Importantly, no transients were observed in the force, as always occur during measurements using a force clamp based on an active feedback loop (8). To demonstrate that the force is truly constant, the force measured simultaneously with the extension record (Fig. S3A) was partitioned into two subsets corresponding to the time spent respectively in the folded (blue) and unfolded (red) states, as determined from the extension record. Histograms of these two subsets of the force record display Gaussian behavior, with the centers differing by less than 0.07 pN (Fig. S3B).

Alignment of Folding and Unfolding Transitions at Constant Force. To search for short-lived intermediate states within the transition from N to U (or U to N), multiple transitions were aligned and averaged to reduce the effects of Brownian motion and determine the shape (extension vs. time) of the folding and unfolding transitions accurately. The transitions were aligned by fitting each transition to the logistic function, an analytic approximation to the Heaviside step function that matches the shape of the transition reasonably well:

$$X(t) = \frac{1}{1 + \exp[-\alpha(t - t_0)]}, \quad [\text{S1}]$$

where X is the end-to-end extension, t_0 is the time at the center of the transition, and α represents the slope of the transition (as $\alpha \rightarrow \infty$, $X(t) \rightarrow \Theta(t)$, the Heaviside step function). The records were aligned on the t_0 fit values for each transition (Fig. S5).

A total of 3,364 folding transitions recorded from different molecules at a variety of forces (within approximately 1 pN of the force at which PrP was poised equally between U and N) were aligned this way, as were 3,318 unfolding transitions. All folding transitions were then averaged, as were the unfolding transitions. No asymmetry was observed between the shapes of the average folding and unfolding transitions: The time-reversed unfolding

curve matches the folding curve within experimental uncertainty (Fig. 2B).

Instrumental Response Function. The response of the optical trap to a step function signal [Fig. S6 (blue)] was determined using a reference construct containing no protein, just the DNA handles attached together via disulfide bonds. The reference construct was held between the traps, the traps were moved suddenly ($<1 \mu\text{s}$), and the resulting motion of the reference construct was measured with the detectors. Such measurements reflect the intrinsic filtering effects of all aspects of the instrumentation (including the effects of the beads and handles); they can be used to determine the effective transfer function that smoothes the actual extension change in the protein according to $O(t) = X(t) \otimes H(t)$, where $O(t)$ is the observed extension record, $X(t)$ is the actual extension time series of the protein itself, and $H(t)$ is the transfer function of the instrument. 200 measurements of the response of the optical trap detectors were averaged as described above for the unfolding/refolding transitions. The amplitude of the trap motion was made equal to the extension change between U and N at constant force (20 nm). Approximating the fast impulse used to measure the response as a true step function, the response is described well if the transfer function is Lorentzian [Fig. S6 (red)].

Modeling the Effect of an Intermediate State. To model the shape of the folding transition expected at constant force if an obligate intermediate I were present, we generated waveforms containing step functions from U to I and then from I to N. Step functions were used as an approximation for the fast (μs -scale) transition time for protein folding (9), which is much faster than the time response of the optical trap (approximately 100 μs , see Fig. S6). The lifetime in state I was chosen randomly from an exponential distribution with mean value τ . Two hundred such curves were generated, convolved with the Lorentzian transfer function to match the shape that would be expected in the measurement (illustrated in Fig. S7A) and then averaged as for the experimentally measured transitions. This procedure simulated the average transition that would be measured for a given value of intermediate lifetime τ and extension change between U and I, Δx_{UI} .

Fig. 2C compares simulations with $\Delta x_{UI} = 7 \text{ nm}$ (the extension change from U to M1) and $\tau = 0, 20, 50,$ and $100 \mu\text{s}$ (respectively yellow, red, cyan, purple) to the experimental data (black). Fig. S7B repeats the calculation with $\Delta x_{UI} = 10.0 \text{ nm}$ (the extension change from U to M2) and the same lifetimes. The experimental data are clearly inconsistent with an obligate intermediate having Δx_{UI} close to these values unless the intermediate lifetime is less than approximately 50–100 μs , thereby ruling out the possibility that M1 and M2 (which have longer lifetimes) might be on the native folding pathway.

Finally, Fig. S7C and D repeat the same calculations, this time simulating what would be expected for an on-pathway intermediate close to the native state, such as would occur if β -strand 1 were to detach (10). If all the N-terminal residues up to E146 were detached in state I (i.e., everything up to helix 1), then the expected extension change between N and I, Δx_{NI} , would be $\Delta x_{NI} = 4.5 \text{ nm}$; if only the residues up to R136 detached, then it would be $\Delta x_{NI} = 2.3 \text{ nm}$. The area around position 139 is known from NMR experiments on bovine PrP to be relatively stable to chemical denaturation (11). Simulations with $\tau = 0, 20, 50, 100,$ and $150 \mu\text{s}$ (respectively yellow, red, cyan, purple, and green) in Fig. S7C and D show that the experimental data are inconsistent with such an intermediate unless its lifetime is less than approximately 100 μs .

Optical Trap Point Spread Function. To determine the point spread function (PSF) of the optical trap for extension measurements, i.e., the distribution of extensions that would be expected for a

construct of fixed length, we measured extension records for the reference constructs described above (DNA handles attached together without any protein between them). These reference constructs were identical to the PrP constructs except that they lacked PrP. Extension records measured at constant force with a 50 kHz bandwidth for approximately 100 s displayed distributions that were almost Gaussian but partly asymmetric (Fig. 4A).

For a harmonic potential with fixed stiffness, the PSF should be Gaussian (7, 12). The asymmetry we observe is due to the anharmonicity of the trapping potential used for the passive force clamp (6): The amplitude of the Brownian fluctuations in the position of the bead is here sufficiently large that the bead explores a significant portion of the anharmonic portion of the potential well. The trap stiffness experienced by this bead therefore varies across the distribution of bead positions. At high displacements from the trap center (corresponding to low molecular extensions) the bead visits regions of negative stiffness, thereby decreasing the effective stiffness of the system (traps plus molecule), whereas at low displacements (corresponding to high molecular extensions) the bead visits regions of positive stiffness, increasing the effective system stiffness. The low-extension side of the Gaussian distribution expected for a harmonic trap is therefore stretched out, whereas the high-extension side is compressed.

These effects were modeled phenomenologically by describing the PSF as a Gaussian-like function with an extension-dependent width parameter:

$$P(x) = \begin{cases} A \exp\left[-\frac{(x-x_0)^2}{2(\sigma-c|x-x_0|)^2}\right], & x > x_0 \\ A \exp\left[-\frac{(x-x_0)^2}{2(\sigma+c|x-x_0|)^2}\right], & x < x_0. \end{cases} \quad [\text{S2}]$$

Here c represents the extension-dependent stretching or compression of the distribution width, and σ represents the extension-independent component of the width. This function fits the data very well, leaving effectively no residual other than counting noise (Fig. 4A).

Fitting Histograms to the PSF. The coefficients c and σ for the PSF [Eq. S2] were determined empirically from measurements of the reference compound and thereafter treated as fixed parameters when fitting the PrP extension histograms as in Fig. 4. To average over statistical counting noise, the residuals from fits of the histograms (as in Figs. 4 and 5) were smoothed with a boxcar filter using a 2.5 nm window. This window was smaller than the half-width of the PSF (approximately 3 nm) and hence would not obscure true signals.

Analysis of On-Pathway Intermediates and Limits of Detection. An intermediate state I on the pathway from U to N might be obligate (e.g., if there is only one pathway from U to N) or nonobligate (e.g., if there is more than one pathway from U to N), as illustrated in Fig. S4A. Regardless of the nature of the intermediate, in measurements where the system is in thermodynamic equilibrium (such as constant-force records of long duration), all possible pathways between accessible states must be sampled. Hence any on-pathway state I must in general be observed in three ways: (i) as a “step” between U and N as the molecule folds or unfolds; (ii) as a “spike” up from N, when N unfolds transiently into I and then refolds; and (iii) as a spike down from U, when U folds transiently into I that then unfolds again.

An example of this behavior is seen when measuring the extension at constant force (Fig. S4B) of a simple DNA hairpin with a sequence designed to generate an obligate intermediate in force-spectroscopy measurements (13). The intermediate is clearly seen as a step between U and N, as a spike up from N, and as a spike down from U. Such results are qualitatively different from the PrP measurements shown in Figs. 1 and 3.

The failure to observe any steps between U and N, whether directly in the records or indirectly via the average of the protein folding/unfolding transitions (as described above), indicates that an obligate on-pathway intermediate does not exist, unless its lifetime is less than the time resolution of the measurements (estimated from the simulations above at approximately 50–100 μ s). Even an intermediate located as little as 2 nm from U or N is ruled out. A rare, nonobligate intermediate is more difficult to rule out by observing the transitions between N and U: Its effects might be lost when averaging the transitions because they are overwhelmed by the transitions containing no intermediate. However, this would only be true if the intermediate lifetime is short compared to the averaging window needed to reduce the Brownian noise significantly. For example, states with lifetimes approximately 0.5 ms or greater can be readily observed directly from extension records filtered at 0.5 ms. In particular, states M1 and M2 (with lifetimes estimated at approximately 0.3–0.5 ms) are clearly detected in the records displayed in Fig. 1C and Fig. 3A.

To place an upper bound on the occurrence of any rare, non-obligate intermediate, we analyzed the behavior in state N. First, we examined the constant-force records for spikes indicating transient unfolding from N to I, especially for evidence of a state I that is well separated from N in distance and has a lifetime of approximately 0.5 ms or longer. Considering transient unfolding from N to I with Δx_{NI} approximately 10 nm or more (i.e., a large extension change), we saw no evidence for such transitions at the level of one event per 300 s (upper bound based on the typical minimum record length), leading to an estimate of approximately 0.002% for the maximum occupancy of such a state.

Because this analysis would not reveal states with very short lifetimes, especially if they involve only a small distance change, we next examined the histograms of the extension values recorded when the protein was nominally in state N. Any partially unfolded intermediate with a different extension than N would show up as an extra peak in the histogram, due to transitions from N to I and back. After fitting the extension histograms in state N to the PSF (as in Fig. 2C), no peaks were observed in the residuals with $\Delta x_{NI} < 10$ nm having an area greater than approximately 0.01% of the total histogram area. The small peaks that did remain in the residual were almost certainly noise, as their location changed from one histogram to the next (Fig. S8A). Averaging over all the residuals to look for any peak occurring consistently at any given extension, we obtain the estimate of approximately 0.001% for the maximum occupancy of a nonobligate intermediate.

Misfolded State Fit Residuals. The quality of the fits for states M1 and M2 is demonstrated by the small residuals left after fitting the high-extension data in the constant-force trajectories (i.e., the nominally unfolded state) to the PSFs for U, M1, and M2 (Fig. S8B).

Folding Rate Estimations. The average folding rate from U to N was measured directly from the extension records. The average folding rate from U to M3 was also estimated directly from the extension records. However, transitions into states M1 and M2 were difficult to identify directly from the extension records, as they were both very short-lived and had low occupancy. We therefore estimated the folding rates for M1 and M2 by dividing the total time spent in each state (determined from the area of the fits to the extension histograms as in Fig. 4C) by the lifetime of the states estimated from the extension records (as in Fig. 3A, *Inset*) to determine the number of transitions into M1 or M2, and then dividing the total time spent in the unfolded state by the number of transitions. The average lifetimes were estimated as being approximately 0.3–0.5 ms for both states. Given the average occupancies for M1 and M2 (0.6% and 0.1%, respectively), this

resulted in rates of approximately 15 s^{-1} and approximately 3 s^{-1} , respectively.

FECs of the C179A/C214A Mutant. FEC measurements of the C179A/C214A mutant (2,410 curves from three molecules) were used to determine the contour length change from folded to unfolded (Fig. S10). Fits to the extensible WLC model as described for the wild-type protein yielded $\Delta L_c = 34 \pm 1$ nm, the same value as for wild-type PrP. If the mutant had formed a nonnative fold, we would expect to observe a different length change. The excellent agreement with ΔL_c expected for the native structure (34.3 nm) indicates that the mutant forms the same fold as wild-type PrP in these measurements.

Aggregation of the C179A/C214A Mutant Compared to Wild-Type PrP.

We monitored the aggregation of C179A/C214A PrP and wild-type PrP using a turbidity assay, similar to a previously published method (14). Briefly, 40 μ M protein in 10 mM phosphate buffer (pH 7.0) was incubated for a week at 37 °C while being agitated at 300 rpm, and the turbidity was measured periodically using a microplate reader via the optical density at 420 nm. The mutant aggregated during sample preparation before the first turbidity measurement was made [Fig. S9 (red)]. The same turbidity was observed with and without reducing agent present (4 mM TCEP). Wild-type PrP aggregated more slowly: When reduced with TCEP, aggregation occurred over the course of 1 d, but without TCEP the protein was stable for more than 1 wk [Fig. S9 (blue)].

Disulfide Bond. Wild-type SHaPrP has a disulfide bond between C179 and C214 (5). This bond is present in our protein samples after purification and refolding, as shown by a Raman peak at 533 cm^{-1} that is in the $510\text{--}540 \text{ cm}^{-1}$ range specific to S-S bonds (15). However, we observe no signature of the disulfide bond in the force-spectroscopy measurements: The full contour length of the structured domain unfolds in a single step at a force much lower than typically needed to break disulfide bonds, without any intermediate state at the distance expected from the NMR structure for unfolding everything up to the disulfide (22.2 nm). It therefore seems unlikely that the disulfide is present during the force-spectroscopy measurements. It is possible the bond does not re-form properly after reduction to attach the DNA handles, or it might be broken while handling the beads at the outset of the measurement, when very large but uncalibrated forces are typically applied. The oxygen-scavenging system that protects against oxidative damage of the protein and DNA creates a nonoxidizing environment, reducing the likelihood of re-forming a disulfide bond.

Free Energy of Misfolded States. Equilibrium measurements allow the free energy differences between states to be determined directly from the ratio of occupancies of the states, using the Boltzmann probability (16):

$$\ln(P_M/P_U) = -k_B T (\Delta G_{\text{fold}} + \Delta G_{\text{stretch}} - F \cdot \Delta x_{MU}), \quad [\text{S3}]$$

where P_M is the probability of finding state M , P_U is the probability of finding state U , k_B is the Boltzmann constant, ΔG_{fold} is the folding equilibrium free energy of M relative to U , $\Delta G_{\text{stretch}}$ is the free energy required to stretch out the unfolded amino acids under tension, F is the applied tension, and Δx_{MU} is the extension change between M and U . P_M/P_U was obtained from the extension histograms and plotted as a function of force for each misfolded state (Fig. S11). The plot was then fitted to Eq. S3, treating Δx_{MU} as a fixed parameter whose value was determined from the extension histograms. The stretching energy, estimated from the WLC model, was then subtracted from the zero-force intercept of the fit to obtain the folding energy of the misfolded state at zero force.

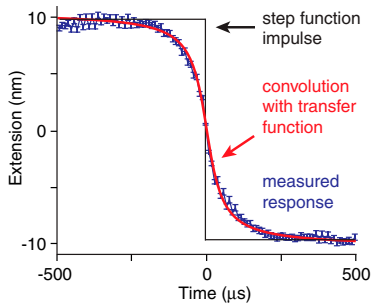


Fig. 56. Instrument response function. The average instrument response (blue) to a step impulse function (black) is well-approximated by the convolution of the impulse function with a Lorentzian transfer function (red).

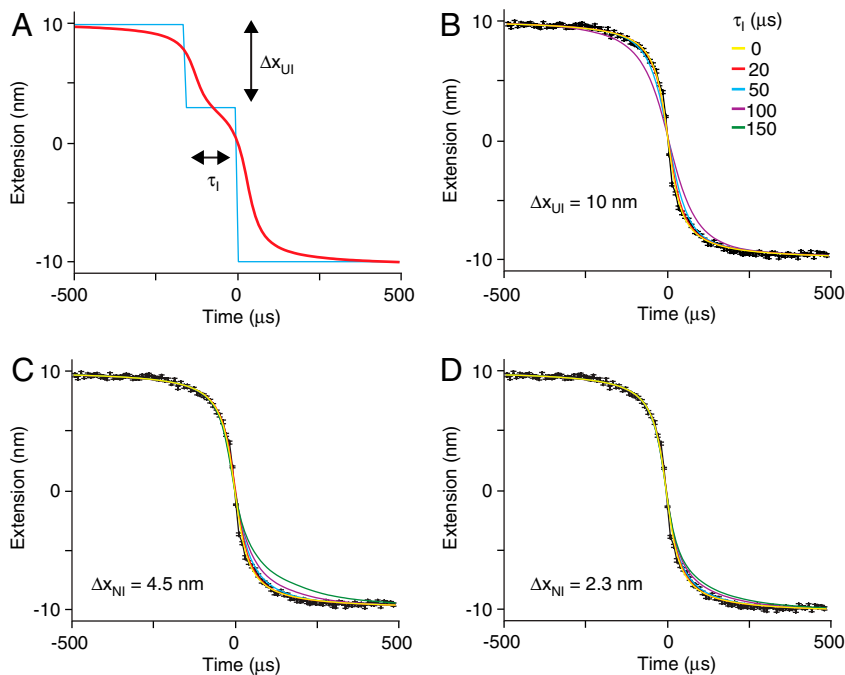


Fig. 57. Simulation of intermediate states. (A) A step function from U to I and then from I to N was generated to simulate the presence of an intermediate distance Δx_{UI} from U (cyan). This was convolved with the Lorentzian transfer function to add in the filtering effect of the instrument (red), and then averaged for exponentially distributed lifetimes with mean value τ_i . (B)-(D) Simulated transitions involving obligate intermediate states with varying lifetimes τ_i , located 10 nm from U (B), 4.5 nm from N (C), and 2.3 nm from N (D), respectively. Black: average measured transition; yellow, red, cyan, purple, green: simulated transitions with $\tau = 0, 20, 50, 100,$ and $150 \mu\text{s}$ respectively.

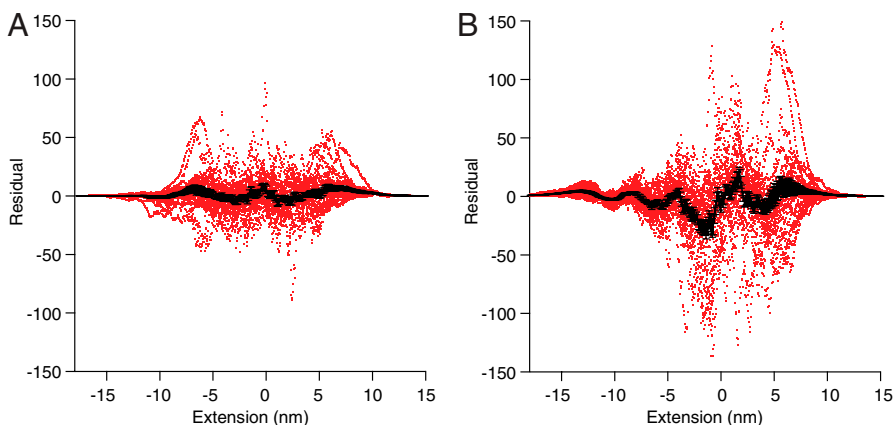


Fig. 58. Residuals for extension histogram fits. (A) The residuals from fitting the low-extension peak of the extension histograms (N in Fig. 1C) to the PSF of the trap, as in Fig. 4B, are shown for 22 constant-force records (red) containing over 29 million data points in total. All residuals were aligned on the peak of the original histogram. The amplitudes and locations of any peaks in the residuals fluctuate randomly. The average of all aligned residuals (black) is featureless, indicating that there is no other state present. Error bars indicate standard error on the mean. (B) The residuals from fitting the high-extension peak of the extension histograms (as in Fig. 4C) to PSFs for states U, M1, and M2 are shown for 22 constant-force records (red) containing over 86 million data points in total. Again, the average (black) is featureless, indicating the excellence of the fits.

