Supplementary Information for

Unfolded and intermediate states of PrP play a key role in the mechanism of action of an anti-prion chaperone

Rafayel Petrosyan, Shubhadeep Patra, Negar Rezajooei, Craig R. Garen, Michael T. Woodside

Department of Physics, University of Alberta, Edmonton AB, T6G 2E1 Canada

Corresponding author: Michael Woodside
Email: michael.woodside@ualberta.ca

This PDF file includes:

Supplementary Methods
Supplementary References
Figs. S1–S7

Supplementary Methods

ITC analyses: Although ITC data are most commonly fit to models that assume a single set of identical binding sites, our data showed features indicative of two sets of independent binding sites: bi-sigmoidal binding isotherms and two qualitatively different shapes of thermogram peaks in the first and second parts of the titration (1, 2). We fit the isotherms to two different models: one containing a single set of identical binding sites, and the other containing two sets of independent binding sites; each site was described by 3 fitting parameters (N, K_D, H), and each fit also allowed for an enthalpy offset (3). The change in entropy was calculated from the fitting parameters. The two fits are shown in Fig. S1.

The single-site model (Fig. S1, blue) clearly does not capture the shape of the isotherm well at either pH 7 or pH 4.5, whereas the two-site model (Fig. S1, red) fits all the data points in both cases. To verify that the better fit for the two-state model does not simply reflect the larger number of fitting parameters, we calculated the Bayesian information criterion (BIC) and Akaike information criterion (AIC) from the fit for each model. These criteria balance the lower fitting residuals achieved when using more fitting parameters with penalties for the number of fitting parameters used, to assess the relative quality of different fitting models (4, 5). For both criteria, the two-site model was strongly favored, with BIC lower by 27 (pH 7) or 15 (pH 4.5) for the two-site model and AIC lower by 25 (pH 7) or 8 (pH 4.5). The ΔAIC value implies that the single-site model is only 2% as likely as the two-site model at pH 4.5 and under 0.001% at pH 7.

Contour length change analysis: We fit FECs displaying unfolding events with a worm-like chain (WLC) polymer elasticity models, combining independent WLCs in series for the unfolded protein and the DNA handles. An extensible WLC model allowing for enthalpic elasticity (6) was used to describe the handles, whereas an inextensible WLC was used to describe the unfolded protein. The WLC force-dependent extension was given by (7)

$$\frac{x(f)}{L_c} = \frac{4}{3} - \frac{4}{3\sqrt{f+1}} - \frac{10\exp\left(\sqrt[4]{\frac{900}{f}}\right)}{\sqrt{f}\left(\exp\left(\sqrt[4]{\frac{900}{f}}\right)-1\right)^2} + \frac{f^{1.62}}{3.55+3.8f^{2.2}}, \qquad \text{(S1)}$$

where L_c is the contour length, $f = FL_p/k_BT$ is the dimensionless force, L_p is the persistence length, and k_BT is the thermal energy. For the extensible WLC model, an additional term of F/K was added to the right hand side, where K is the enthalpic elasticity. For the DNA handles, L_c was fixed at 990 nm (based on the number of base-pairs used), and the same fitting parameters (L_p ~45 nm and K ~ 1500 pN) were used on all branches of the FEC. For the unfolded protein, L_p was fixed at 0.65 nm (9,10), so that the only fitting parameter was the contour length of unfolded protein.

To verify that PPS did not change the mechanical properties of the DNA handle, we measured pulling curves for DNA handles alone, without any protein, in the presence of 5 μM PPS in solution. Fitting the resulting FECs to the same extensible WLC model returned parameters that did not deviate from those found for measurements of DNA handles in the absence of PPS.

Unfolding force distribution analyses: As done previously for studies of PrP measured in the absence of PPS (9), we attempted to use the model of Dudko *et al*. (10) to fit the distribution of unfolding forces. This model assumes single, well-defined native state and a linear-cubic profile for the free-energy barrier, yielding the result

$$p(F)=\frac{k(F)}{r}\exp\left\{\frac{k_0}{\Delta x^{\ddagger} r}-\frac{k(F)}{\Delta x^{\ddagger} r}\left(1-\frac{2\Delta x^{\ddagger} F}{3\Delta G^{\ddagger}}\right)^{-1/2}\right\}, \text{ where} \tag{S2}$$

$k(F)=k_0\left(1-\frac{2\Delta x^{\ddagger} F}{3\Delta G^{\ddagger}}\right)^{1/2}\exp\left\{\frac{\Delta G^{\ddagger}}{k_B T}\left[1-\left(1-\frac{2\Delta x^{\ddagger} F}{3\Delta G^{\ddagger}}\right)^{3/2}\right]\right\}$ is the force-depending unfolding rate, k_0 is the unfolding rate at zero force, $\Delta x^{\ddagger}$ is the distance from the folded state to the energy barrier for unfolding (transition state), $\Delta G^{\ddagger}$ is the barrier height, and r is the loading rate (here, 35 pN/s).

To account for the effects of heterogeneous initial states on the unfolding force distribution we used a model that was proposed recently (11). Eq. S3 gives the unfolding force probability density function for this model:

$$p(F)=\frac{k_0}{r}\exp\left[\frac{F\Delta x^{\ddagger}}{k_{\mathrm{B}}T}\right]\left(1+\frac{k_0 k_{\mathrm{B}} T\Delta}{r\Delta x^{\ddagger}}\left(\exp\left[\frac{F\Delta x^{\ddagger}}{k_{\mathrm{B}}T}\right]-1\right)\right)^{-\frac{\Delta+1}{\Delta}}. \tag{S3}$$

Here k_0 is the average zero-force unfolding rate for the heterogeneous ensemble of starting states, $\Delta x^{\ddagger}$ is the average distance to the transition state for the ensemble of starting states, and Δ is a parameter reflecting the extent of heterogeneity.

Clustering analysis: The unfolding events in the presence of PPS were clustered in the ΔL_c-unfolding force plane using a variational Gaussian mixture algorithm. The quality of the clustering was tested using the silhouette index, which quantifies how well each point belongs to the cluster to which it is assigned on a scale of −1 to 1 (12, 13). The global silhouette index of 0.41 indicated reasonable clustering. To account for uncertainties in cluster assignments when creating unfolding force distributions for individual clusters, data-points were weighted by the probability of belonging to that cluster, as calculated using the naive Bayes classifier function in Mathematica.

Docking analysis: To determine the most likely binding locations of PPS, we carried out docking computations of dimeric PPS (using the structure from PubChem) with the structured C-terminal domain of hamster PrP^C (PDB 1B10). Docking was done using AutoDock 4.2 (14). Four independent computations were carried out, each generating a set of docked poses ranked by binding energy. The top 5 predicted binding poses from each of four independent computations, illustrated in Fig. S7, primarily involve binding to helices 1 and 2 (contacts with residues R156, H187, and K194), or binding to the middle of helix 3 (contacts with residues Y128, N181, K185, and G186).

Supplementary References:

1. C. Li, *et al.*, Cupric Ions Induce the Oxidation and Trigger the Aggregation of Human Superoxide Dismutase 1. *PLoS ONE* **8**, e65287 (2013).

2. A. L. Pey, T. Majtan, J. M. Sanchez-Ruiz, J. P. Kraus, Human cystathionine β-synthase (CBS) contains two classes of binding sites for S-adenosylmethionine (SAM): complex regulation of CBS activity and stability by SAM. *Biochem. J.* **449**, 109–121 (2013).

3. C. A. Brautigam, H. Zhao, C. Vargas, S. Keller, P. Schuck, Integration and global analysis of isothermal titration calorimetry data for studying macromolecular interactions. *Nat. Protoc.* **11**, 882–894 (2016).

4. G. Schwarz, Estimating the dimension of a model. *Ann. Stat.* **6**, 461–464 (1978).

5. H. Akaike, A new look at the statistical model identification. *IEEE Trans. Automat. Contr.* **19**, 716–723 (1974).

6. M. D. Wang, H. Yin, R. Landick, J. Gelles, S. M. Block, Stretching DNA with Optical Tweezers. *Biophys. J.* **72**, 1335–1346 (1997).

7. R. Petrosyan, Improved approximations for some polymer extension models. *Rheol. Acta* **56**, 21–26 (2017).

8. E. A. Shank, C. Cecconi, J. W. Dill, S. Marqusee, C. Bustamante, The folding cooperativity of a protein is controlled by its chain topology. *Nature* **465**, 637–640 (2010).

9. H. Yu, *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. U.S.A.* **109**, 14452–14457 (2012).

10. O. K. Dudko, G. Hummer, A. Szabo, Intrinsic Rates and Activation Free Energies from Single-Molecule Pulling Experiments. *Phys. Rev. Lett.* **96**, 108101 (2006).

11. M. Hinczewski, C. Hyeon, D. Thirumalai, Directly measuring single-molecule heterogeneity using force spectroscopy. *Proc. Natl. Acad. Sci. U.S.A.* **113**, E3852–E3861 (2016).

12. P. J. Rousseeuw, Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics* **20**, 53–65 (1987).

13. M. Brun, *et al.*, Model-based evaluation of clustering validation measures. *Pattern Recognition* **40**, 807–824 (2007).

14. G. M. Morris, *et al.*, Automated Docking Using a Lamarckian Genetic Algorithm and an Empirical Binding Free Energy Function. *J. Comput. Chem.* **19**, 1639–1662 (1998).

15. S. Sen Mojumdar, *et al.*, Partially native intermediates mediate misfolding of SOD1 in single-molecule folding trajectories. *Nat. Commun.* **8**, 1881 (2017).

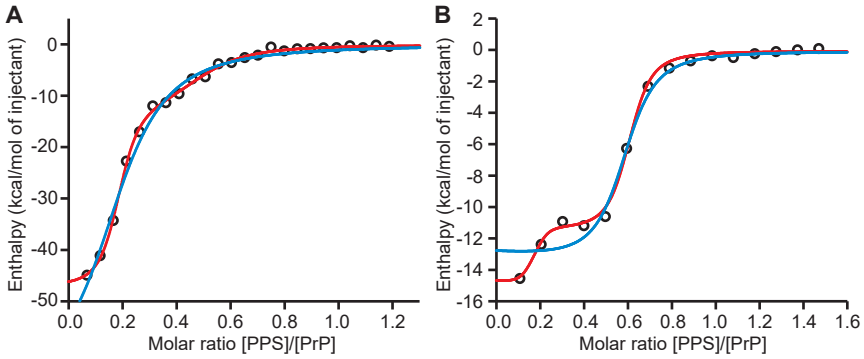

Fig. S1: Comparison of one-site and two-site models for ITC fitting. (A) Fitting the binding isotherm for ITC measurements at pH 7 to a single-site model (blue) does not capture the bi-sigmoidal shape of the isotherm, whereas fitting to a two-site model (red) does. (B) The same applies to measurements at pH 4.5.

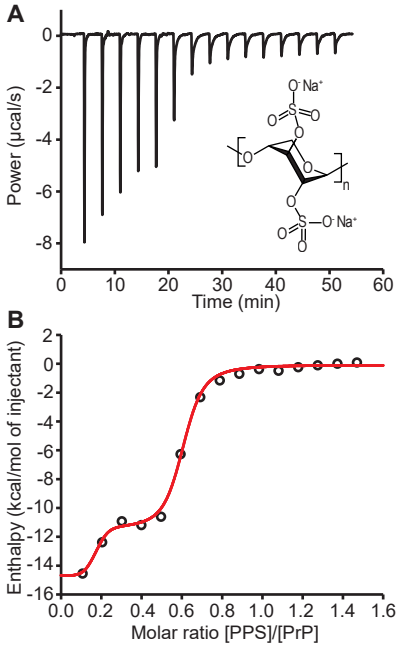

Fig. S2: ITC measurements of PPS binding to PrP at pH 4.5. (A) ITC thermogram. (B) Binding isotherm (black) fit to model with two independent binding sites (red).

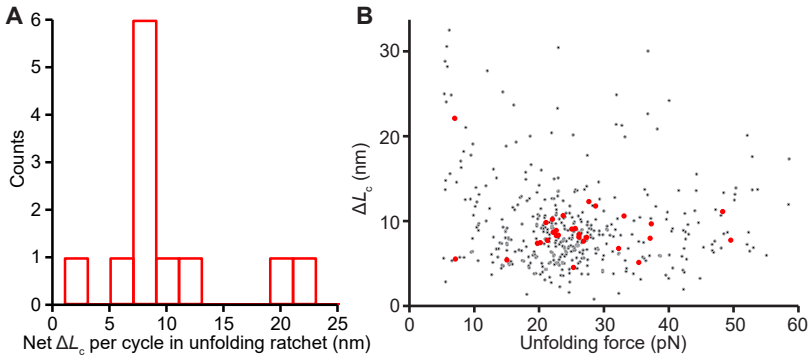

Fig. S3: Analysis of ratchet-like unfolding events. (A) Net change in contour length per unfolding-refolding cycle during ratchet-like unfolding. (B) The properties of the unfolding transitions (force, ΔL_c) during ratchet-like events (red) were somewhat more tightly clustered than all unfolding transitions (grey), with almost all having ΔL_c ~ 6–12 nm.

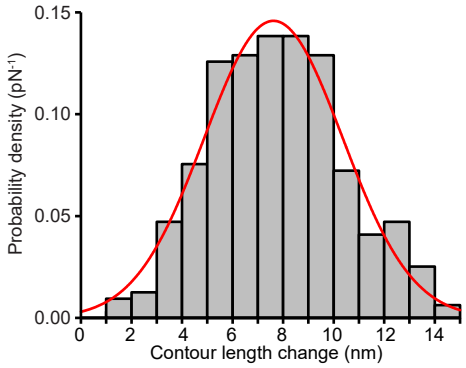

Fig. S4: Contour-length changes in the main cluster of unfolding events. Fitting the distribution of ΔL_c values (black) to a Gaussian (red) returns a standard deviation of 2.7 nm, notably higher than the standard deviation of ~1.1 nm observed for single well-defined states in previous measurements using the same instrumentation (9, 15), indicating the presence of more than one state.

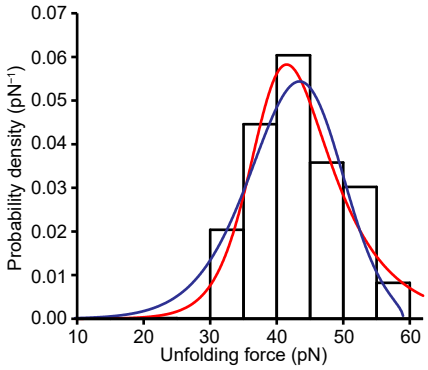

Fig. S5: Analysis of cluster of high-force transitions. The force distribution for the high-force cluster (Fig. 5B, red) fit better to the heterogeneous model (red) than the model with a single initial state (blue), returning a heterogeneity parameter of $\Delta = 2.2 \pm 0.9$.

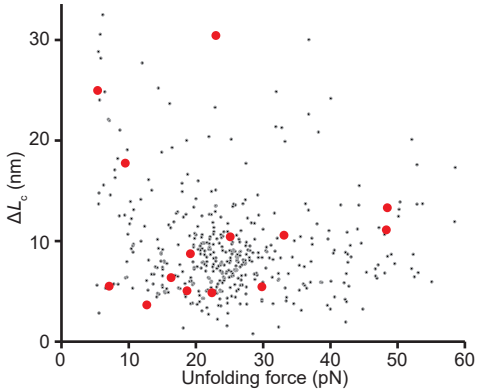

Fig. S6: Analysis of first pulling events from each molecule. The events observed when each molecule was unfolded for the first time (red) were just as heterogeneous as the other unfolding transitions observed subsequently (grey).

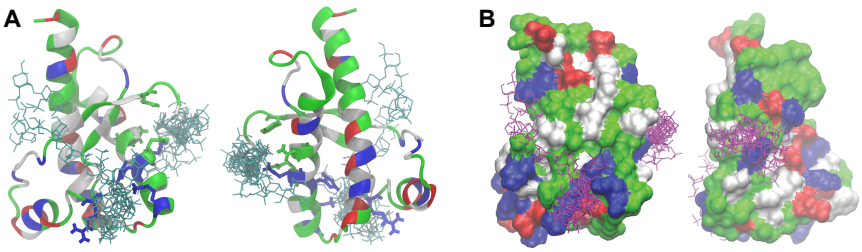

Fig. S7: Models of PPS binding to PrP. (A) Ribbon diagram and (B) space-filling model views of PrP showing the top 5 poses for PPS dimers bound to natively folded PrP found from each of four docking calculation, viewed from two perspectives. Basic residues are shown in blue, acidic in red, and polar in green.