Low Folding Cooperativity of Hp35 Revealed by Single-Molecule Force Spectroscopy and Molecular Dynamics Simulation

Chunmei Lv, Cheng Tan, Meng Qin, Dawei Zou, Yi Cao, and Wei Wang

National Laboratory of Solid State Microstructure and Department of Physics, Nanjing University, Nanjing, People's Republic of China

Supporting Material

1. Experimental details

1.1 AFM data analysis

Worm-like chain model of polymer elasticity (eq. 1) was used to fit the saw-tooth like pattern to extract the persistence length of unfolded HP35 and the mixed form of unfolded HP35 and GB1.

$$F = \frac{k_B T}{p}\left(\frac{1}{4}(1-\frac{x}{Lc})^{-2} - \frac{1}{4} + \frac{x}{Lc}\right) \qquad \text{eq. 1}$$

Where F is the applied force, p is the persistence length, k_B is the Boltzmann constant, Lc is the contour length, and T is the absolute temperature (1, 2).

The experimental data were fitted in Igor Pro (Wavemetrics, Lake Oswego, Oregon, USA) using the built-in data-fitting procedure. Because for p2 fitting, only the data above 70 pN were available, we double-checked the p2 using refolding traces such as the red trace shown in Fig. S2, we did not find any significant differences between the obtained p2 from the two difference data sets. The traces that correspond to earlier detachment (Not all GB1 domains in the polyprotein are unfolded.) or low detaching force are not used for data analysis.

We have also tried data fitting using other polymer elasticity models, such as freely jointed chain model (FJC). We obtained the same trend for Kuhn length as for persistence length.

Estimation of the force acting on the polyprotein after unfolding of the last GB1 domain is a bit tricky. Because we were using very soft cantilever in our experiment, it takes time for the cantilever to recover from bended conformation after the unfolding of a protein domain. Therefore, the force-extension always includes a cantilever relaxation phase, in which the force acting on the polyprotein cannot directly read from the force-extension curve (3, 4). The softer the cantilever, the slower the relaxation is. For example, after the detachment of the polyprotein, the actual force on the polyprotein should be zero. But in the force-extension curve, it only slowly decays to zero following a shallow slope (Fig. 1, Fig. 3, and Fig. S4). To estimate the residual force, we assume after unfolding of GB1, the polyprotein immediately reaches equilibrium. Therefore, the force acting on the polyprotein after the unfolding of a GB1 domain can be estimated based on the WLC fitting to the region of the trace after this unfolding event. The residual force after a GB1 domain unfolding is the force in the WLC fitting trace extrapolating to where the GB1 domain unfolds. From this estimation, the force on the polyprotein where p2 is measured is about 50 pN.

1.2 Estimation of the chance of observing folding/unfolding equilibrium of HP35 in the unfolding traces

According to Bell-Evans Model (Bell et al., Science 1978 and Evans et al., Biophysical Journal 1997), the relationship between folding rate and force for a two-state folder can be described as following:

$$k_f\left(F\right) = k_{f0}\mathrm{e}^{-F\Delta x_f / k_B T} \qquad \text{eq. 2}$$

Where $k_f(F)$ and k_{f0} are folding rates at a given force, F, and at zero force, respectively; Δx_f is the length change of a protein upon folding; k_B is the Boltzmann constant; and T is the absolute temperature.

The folding rate, k_{f0},of HP35 was estimated to be ~1.38×10^6 s^{-1} at 300 K (5). Assuming at a force of 10 pN, the length change upon folding (Δx_f) is less than 3 nm (the end-to-end distance at unfolded state subtracted by the end-to-end distance at folded state and the unfolding distance) (6, 7). Therefore, the folding rate at 10 pN is ~900 s^{-1}. If the pulling speed in AFM experiments is low enough (<400 nm s^{-1}), the fast folding and unfolding equilibrium can be observed in the unfolding traces.)

However, we did not observe any folding events in the unfolding traces even at a pulling speed as low as 10 nm s^{-1}, indicating that HP35 is not a two-state folder.

1.3 Circular Dichroism (CD)

Far UV CD spectra were recorded on a Jasco-J810 spectropolarimeter flushed with nitrogen gas. The spectra were recorded in a 0.01cm path length cuvette at a scan rate of 50nm/min. The protein samples were measured in 0.1×Tris-HCl at pH 7.4 at a concentration of ~0.2 mg mL^{-1}. For each protein sample, an average of 10 scans was reported and the CD signal was converted into mean residue ellipticity (MRE) using the following equation:

$$\theta_{MRE} = \frac{100\cdot\theta_{obs}}{d\cdot C\cdot\left(n-1\right)} \qquad \text{eq. 3}$$

where θ_{obs} is the observed ellipticity (in deg), d is the path length (in cm), C is the concentration of protein samples (in M), and n is total number of amino acids in the protein. The data reported have been corrected for buffer contributions.

1.4 Molecular dynamics simulation details

Stretching forces were applied to HP35, HP67 and GB1 using constant-velocity SMD protocols. In all the simulations, the Cα atom of the C terminal residue was kept fixed, and the Cα atom of the N terminal residue was pulled along the direction that connects the initial positions of the N terminus and the C terminus.
The structure of HP35 used in the SMD simulation was taken from the Protein Data Bank (code: 1UNC). The structure of HP67 was constructed from its homologous protein (PDB code: 1QQV) using the online modeling tool "SWISS-MODEL SERVER (http://swissmodel.expasy.org)" (8-10) . The structure of GB1 was also taken from the Protein Data Bank (code 1PGA).

Solvation and ionization were performed using the VMD program package (11). The whole protein-water-ion system of HP35, HP67 and GB1 contain 13300, 49811 and 29764 atoms, respectively. All the systems were first energy minimized for

200000 steps. The systems of HP35 and GB1 were then heated to 300 k during an NVT simulation of 1 ns, and equilibrated during an NPT simulation for 10 ns. After the energy minimization, the HP67 system was also heated to 300 k during an NVT simulation of 1 ns. The system was then equilibrated during an NPT simulation of 22 ns, because the structure was artificially built from a homologous model. After all these equilibrium simulations, 25 different conformations were randomly picked out from the trajectory of HP35 and taken as the initial structures for 25 individual SMD simulations. 5 SMD simulations for HP67 and 5 for GB1 were also performed. All the SMD simulations were carried out using the constant-velocity protocol, with pulling velocity $v = 0.1$ m s^{-1}. Constant volume control was used, whereas the constant temperature control was disabled to avoid additional disturbing of the movement of the atoms.

The molecular dynamics simulations were carried out using the NAMD package (version 2.7) with the CHARMM27 force field including CMAP correction to proteins. The TIP3P water model was used for the solvent. The integration time step was set to 2 fs. The non-bonded Coulomb and vdW interactions were calculated with a cutoff using a switching function starting at a distance of 10 Å and reaching zero at 13 Å.

Supporting references:

1. Bustamante, C., J. F. Marko, E. D. Siggia, and S. Smith. 1994. Entropic elasticity of lambda-phage DNA. Science 265:1599-1600.
2. Marko, J. F., and E. D. Siggia. 1995. Stretching DNA. Macromolecules 28:8759-8770.
3. Crampton, N., K. Alzahrani, G. S. Beddard, S. D. Connell, and D. J. Brockwell. 2011. Mechanically unfolding protein L using a laser-feedback-controlled cantilever. Biophys J 100:1800-1809.
4. Dietz, H., and M. Rief. 2004. Exploring the energy landscape of GFP by single-molecule mechanical experiments. Proc Natl Acad Sci U S A 101:16192-16197.
5. Kubelka, J., W. A. Eaton, and J. Hofrichter. 2003. Experimental tests of villin subdomain folding simulations. J Mol Biol 329:625-630.
6. Junker, J. P., F. Ziegler, and M. Rief. 2009. Ligand-dependent equilibrium fluctuations of single calmodulin molecules. Science 323:633-637.
7. Best, R. B., and G. Hummer. 2008. Protein folding kinetics under force from molecular simulation. J Am Chem Soc 130:3706-3707.
8. Arnold, K., L. Bordoli, J. Kopp, and T. Schwede. 2006. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. Bioinformatics 22:195-201.
9. Schwede, T., J. Kopp, N. Guex, and M. C. Peitsch. 2003. SWISS-MODEL: An automated protein homology-modeling server. Nucleic Acids Res 31:3381-3385.
10. Guex, N., and M. C. Peitsch. 1997. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. Electrophoresis 18:2714-2723.
11. Phillips, J. C., R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale, and K. Schulten. 2005. Scalable molecular dynamics with NAMD. J Comput Chem 26:1781-1802.

Supporting Table:

Table S1 Summary of the persistence lengths of polyproteins obtained from force-extension traces and their corresponding structural elements.

	p (mean±S.D, nm)	**Corresponding elements and their relative contour length**
p1 of $(GB1\text{-}HP35)_8$ [Figure 3d]	0.72±0.15	unfolded HP35 (102 nm), folded GB1(21 nm), other unstructured linker (10 nm)
p2 of $(GB1\text{-}HP35)_8$ [Figure 3h]	0.49±0.09	unfolded HP35 (102 nm), unfolded GB1 (163 nm) , other unstructured linker (10 nm)
p1 of $(GB1\text{-}HP67)_8$ [Figure 3e]	0.62±0.15	unfolded HP67 (195 nm) , folded GB1 (21 nm), other unstructured linker(10 nm)
p2 of $(GB1\text{-}HP67)_8$ [Figure 3i]	0.41±0.11	unfolded HP67 (195 nm), unfolded GB1 (163 nm), other unstructured linker (10 nm)
p1 of $(GB1)_8$ [Figure 3f]	0.56±0.14	folded GB1 (21 nm), other unstructured linker (10 nm)
p2 of $(GB1)_8$ [Figure 3j]	0.41±0.08	unfolded GB1 (163 nm), other unstructured linker (10 nm)

Supporting Figures:

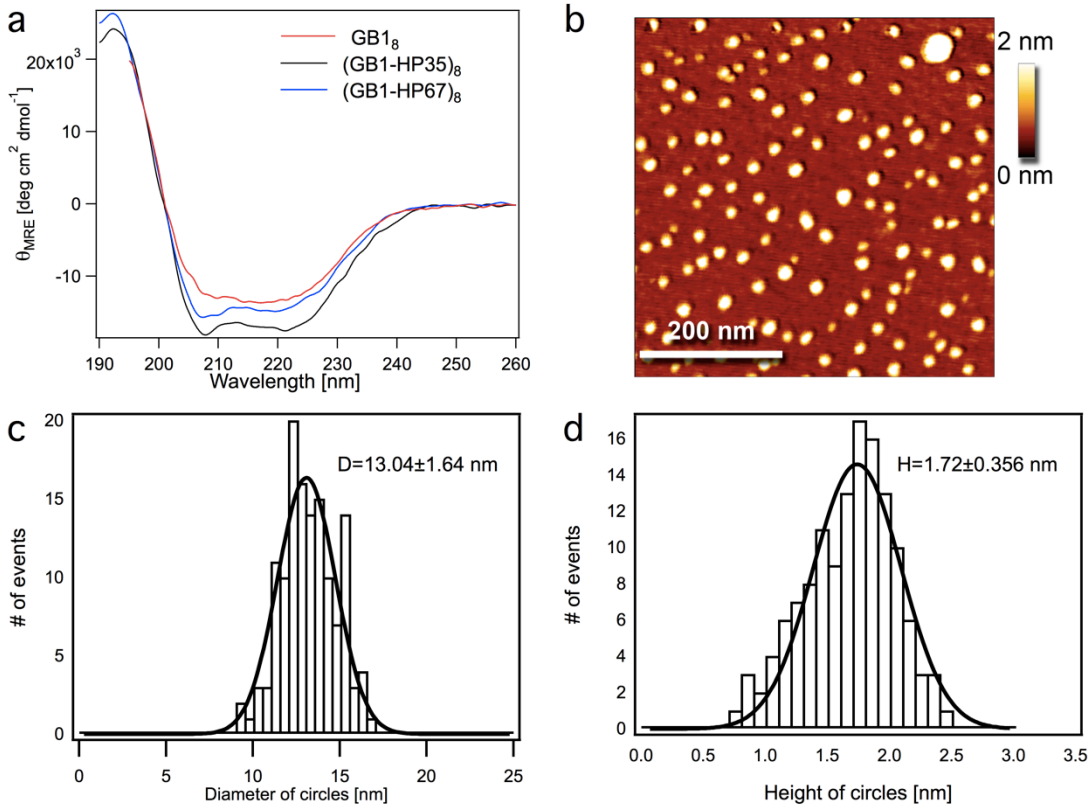

Fig. S1 The structure and morphology of the protein samples used in single molecule AFM experiments. a) CD spectra of polyprotein GB1$_8$, (GB1-HP35)$_8$ and (GB1-HP67)$_8$. b) an AFM image of (GB1-HP35)5 sample obtained in air using intermittent-contact mode. c) and d) the histograms of the diameter and the height of the (GB10HP35)$_8$ molecules. The broadening effect from cantilever tip radius has been subtracted in diameters reported in c). Such uniform distribution suggests that HP35 are properly folded in the polyprotein.

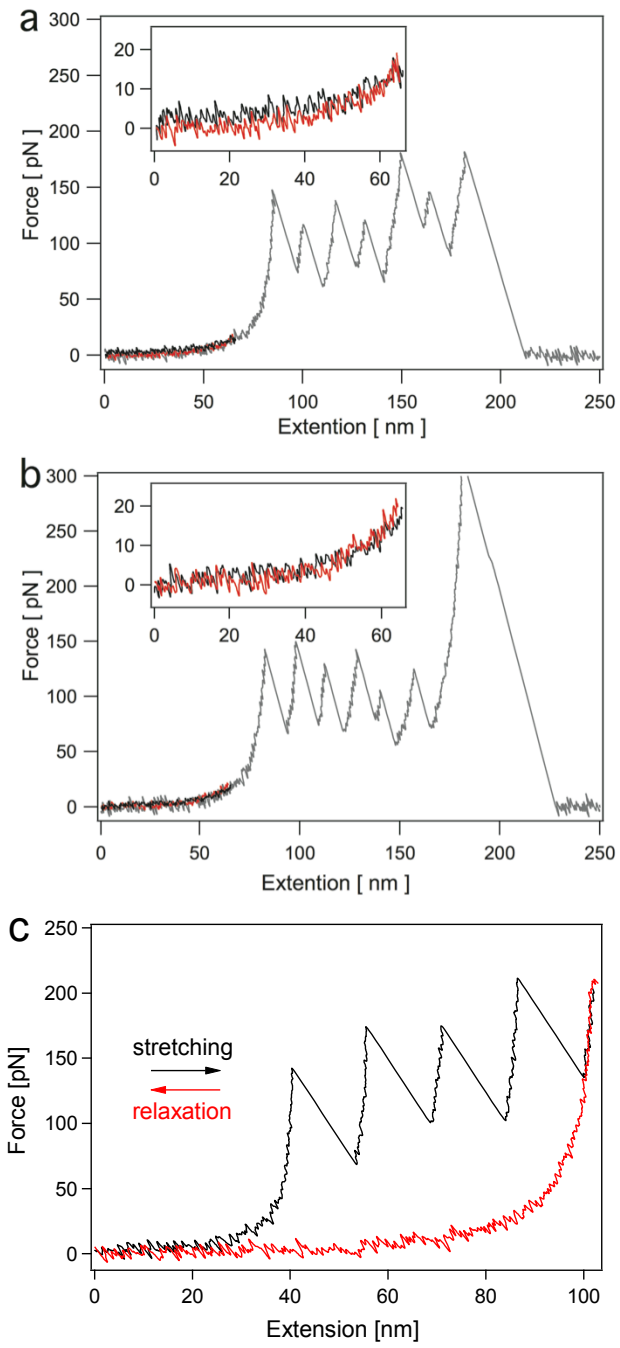

Fig. S2 Refolding traces of $(GB1\text{-}HP35)_8$ at a pulling speed of 400 nm s^{-1}. a) and b) Two additional refolding traces of HP35 (the stretching force is kept low during refolding cycles.). Inset shows the average trace obtained from 4 refolding cycles on the same molecule. No obvious hysteresis is observed between stretching and relaxation traces. c) Stretching-relaxation cycles of $(GB1\text{-}HP35)_8$ with higher stretching force. The hysteresis between stretching and relaxation traces comes from the unfolding of GB1.

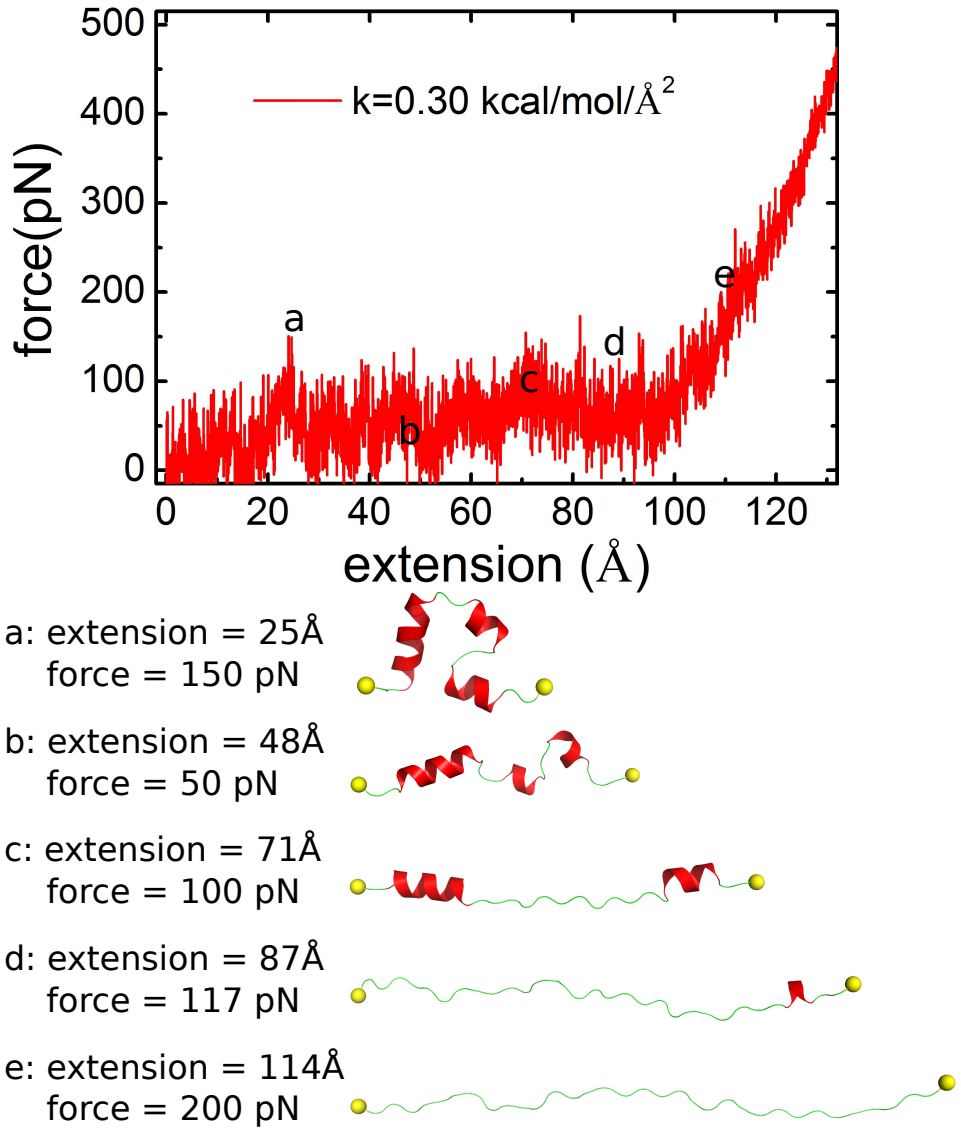

Fig. S3 Molecular dynamics simulation of HP35 unfolding up to a stretching force as high as 400 pN. Instead of showing a clear unfolding transition, the forces fluctuate around 30-100 pN spanning a long extension regime up to ~10 Å. This corresponds to gradually lose of its secondary structure. The secondary structure completely disappear at ~104 Å, at a pulling force of 120-150 pN. A few representative snapshots at various extensions and forces are depicted below the unfolding trace. In the simulation, the Cα atom of the C terminal residue was kept fixed, and the Cα atom of the N terminal residue was pulled in the simulation. The secondary structures are identified using their characteristic dihedral angles.

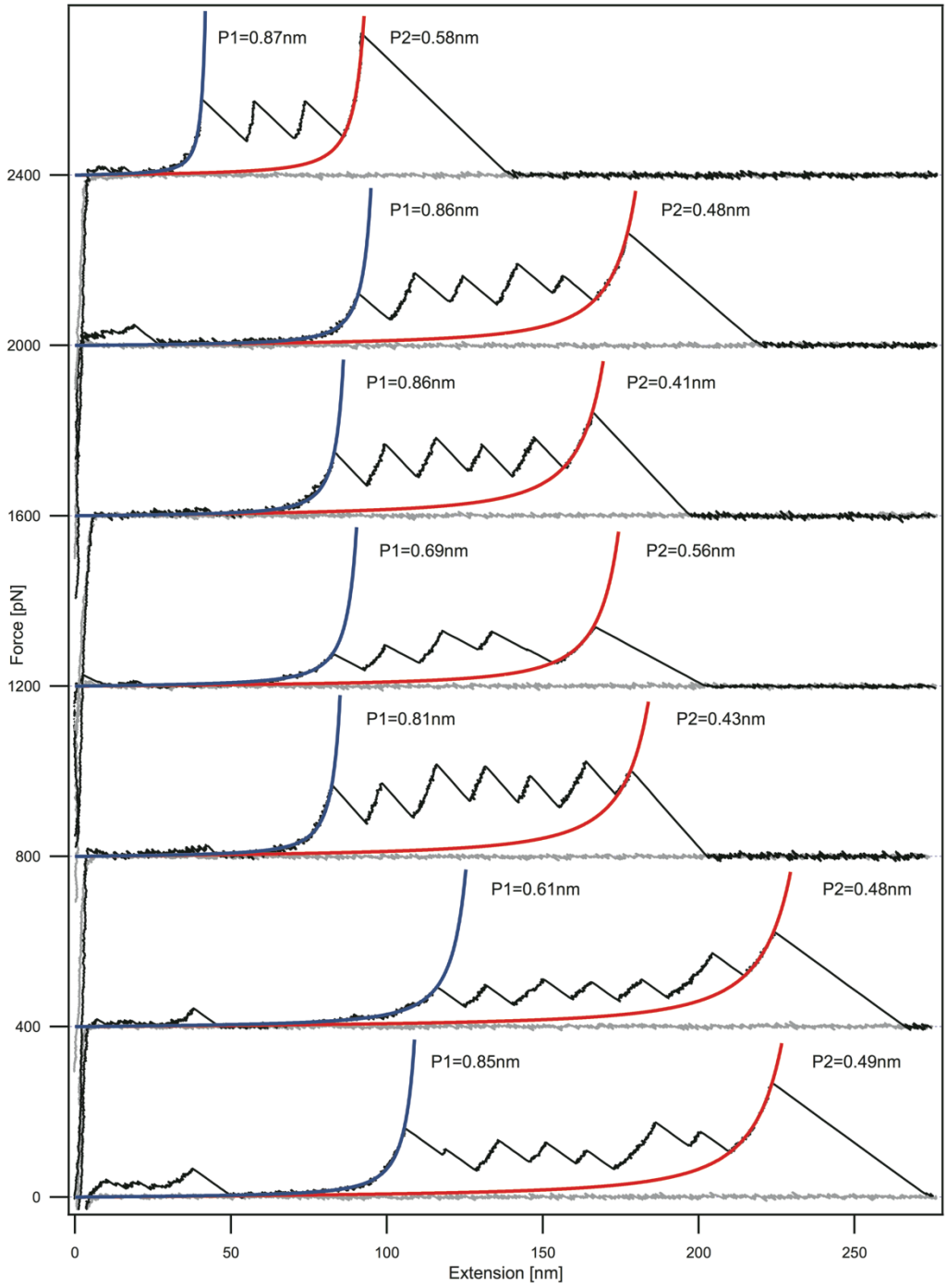

Fig. S4 Additional unfolding traces of $(GB1\text{-}HP35)_8$ at a pulling speed of 400 nm s^{-1}. Blue and red lines correspond to WLC fitting. Clearly, the persistence of unfolded HP35 (p1) is longer than that of the mixed form of unfolded HP35 and GB1 (p2).

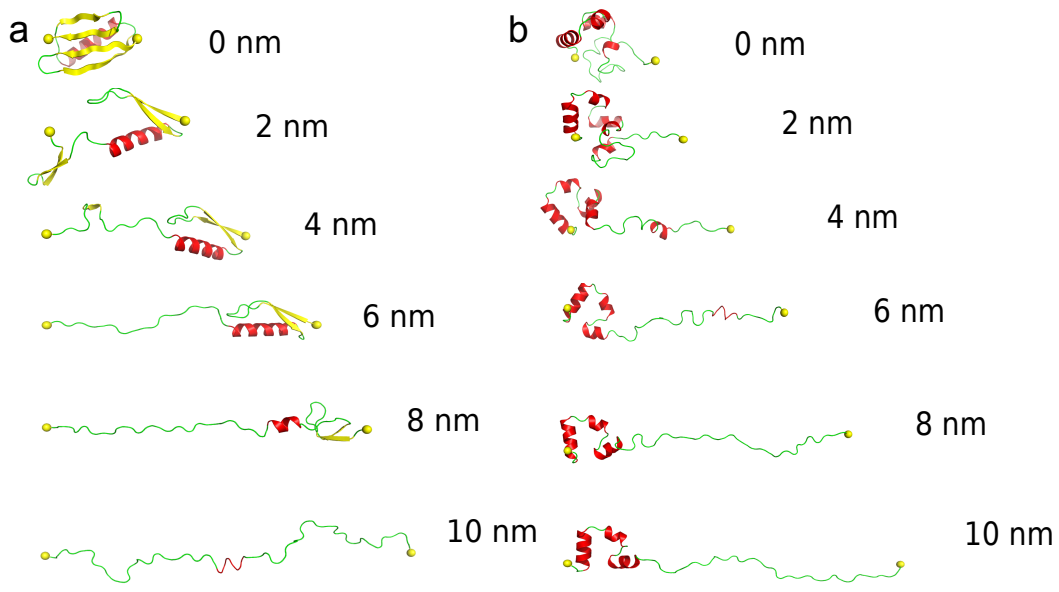

Fig. S5 Representative snapshots of GB1 (a) and HP67 (b) during the SMD simulations at constant pulling velocity (v = 0.1 m s^{-1}). In all the simulations, the Cα atom of the C terminal residue was kept fixed, and the Cα atom of the N terminal residue was pulled in the simulation.