Supporting Material

Mechanical anisotropy of ankyrin repeats

Whasil Lee, Xiancheng Zeng, Kristina Rotolo, Ming Yang, Christopher J. Schofield, Vann Bennett, Weitao Yang, Piotr E. Marszalek

Supporting information text

The analysis of native contacts during transient refolding of D34

When a transient pair of antiparallel helices is formed (Fig. 5A, middle panel), 11 contacts are formed, and also 11 contacts are formed when a transient loop-hairpin structure is created. Both processes shorten the polypeptide chain by 2 nm and the structures are inherently unstable and can easily open due to thermal fluctuations. If H2 of R16 had refolded to the stack, it would have formed only 12 contacts (marked by a yellow/blue arrow in Fig. 5A), but would have shortened the chain by 3 nm. This number of contacts is comparable to the number of possible native contacts formed in a pair of antiparallel helices, or a loop-hairpin structure elsewhere along the chain. The refolding at the stack interface competes with refolding elsewhere when the polypeptide chain relaxes, but neither is stable enough to resist the thermal fluctuations. During that time window between 973 and 1010 ns, the transient structures formed one by one, prohibiting the H2 of R16 from refolding because the transient structures required less contraction length than R16-H2 (2 nm vs. 3 nm). Between 1010 and 1012 ns, two transient structures in R13-R14 opened simultaneously, providing an opportunity for the entire R16 (H1, H2 and the loop-hairpin) to refold in a single step forming a very large number of contacts.

MATERIALS AND METHODS

DNA cloning and protein expression

Our chimeric polyprotein $(I27)_3$-NI6C construct consisted of NI6C, a consensus Ankyrin repeat (AR) protein, and three I27 domains flanking the N-termini. The sequences of the capping and internal repeats of NI6C were adapted from (1) (Fig. 1B). The NI6C gene was inserted into the Poly-I27 pRSETa vector [a kind gift from Jane Clark (2)] using KpnI and MluI restriction sites. The engineered plasmids were transformed into E. coli C41 (DE3) (Lucigen catalog # 89027), and the $(I27)_3$-NI6C protein was expressed using IPTG induction. The harvested cells were lysed and purified by using a Ni-NTA affinity column (GE Healthcare catalog # 17-5268-02). The gene of D34 (residues 402-827, sequence in Fig. 1A) was inserted into pET-28a plasmid which contains His-tag. The engineered plasmids were transformed into E. coli BL21, and expressed using IPTG induction. The expressed proteins were purified by using a Ni-NTA affinity column followed by size exclusion chromatography. Our chimeric polyprotein $(I27)_3$-D34-$(I27)_3$ construct consisted of D34 and three I27 domains each flanking the N- and C-termini. The D34 gene was inserted into the Poly-I27 pRSETa vector [a kind gift from Jane Clark (2)] using KpnI and NheI restriction sites, and the 8th I27 module was replaced by strep-tag and a STOP codon.. The engineered plasmids were transformed into E. coli C41 (DE3) (Lucigen catalog # 89027), and the construct was expressed using IPTG induction. The harvested cells were lysed and purified by using a Ni-BTA affinity column followed by strep-tag column (IBA #2-1202-0250). Proteins were determined to be greater than 95% pure by SDS-PAGE analysis. All purified proteins were dialyzed in 150 mM NaCl, 1mM EDTA, 2mM TCEP (Thermo Scientific, product # 77720) and 10 mM Tris buffer (ph 7.4).

Single molecule force spectroscopy

50 µl of a diluted protein solution (1~5 µg/ml) was incubated on a substrate for 20 minutes, rinsed once by adding and removing 50 µl of the buffer to remove protein molecules that did not adhere to the substrate surface. Both nickel-NTA functionalized glass substrates (3-5) and clean glass substrates were used for AFM experiments. All stretching measurements were carried out on custom-built AFM instruments (3, 6, 7) equipped with an AFM

detector head (Veeco Metrology Group), and high-resolution piezoelectric stages with vertical resolution of 0.1 nm (Physik Instrumente). The spring constant, k_c of each AFM cantilever was calibrated in solution using the energy equipartition theorem as described (8). Force-extension measurements were performed in solution using biolevers (Veeco, $k_c \approx 6$ pN/nm) or microlevers (Veeco, $k_c \approx 18$ pN/nm) at pulling speeds of 0.03 ~ 0.1 nm/ms at room temperature. For AFM force spectroscopy measurements, proteins were picked up by an AFM tip by gently contacting the sample substrate at forces lower than 1 nN. After a molecule was picked up, we performed cyclic stretch-relax measurements on it. In those measurements of chimeric proteins, the extension was limited so I27 domains were not unfolded while the AFM tip was slightly lifted above the sample substrate (5 ~ 10 nm) to prevent adsorption of other molecules.

Force peak analysis of D34

During the AFM measurements on D34 monomers, we have limited control over pulling geometry that is somewhat random depending on the random attachment of the protein to AFM tip or to the substrate (when clean glass substrate is used). To obtain the force / extension histograms, the force-extension traces of D34 monomers with longer than 70 nm (~6 ARs unfolded length) are analyzed. First, the force peaks in the force-extension curves were fitted to the Worm-like-chain (WLC) model (9) (Fig. 6A, Fig. S1, Fig. S3 and Fig. S4). To determine the unfolding and refolding forces of ankyrin repeats of D34, the maximum force generated upon unfolding and refolding events relative to the force baselines was identified. (Fig. S1*B*). However, in this approach we may miss some unfolding / refolding events, if they occur close to the force noise level. Therefore, it is possible that our average forces (especially the refolding forces) could be somewhat overestimated rather than underestimated.

AFM measurements on D34-I27 constructs

We engineered a D34 monomer and a number of constructs of D34 with I27 handles. Unfortunately, the D34-I27 hybrid proteins did not produce consistent AFM force-extension recordings. In some recordings, complete D34 unfolding preceded the unfolding of I27 domains, but in a significant percentage of recordings we observed only partial unfolding of D34 even though the whole D34 insert was picked up by the AFM tip, as judged by the number of unfolding force peaks generated by flanking I27 domains. In addition, these partial unfolding recordings varied from one measurement to another (Fig. S4*B*). These difficulties are presumably caused by the SBD fragment that folds back on the ankyrin stack, which possibly causes some entanglement between the ankyrin stack and flanking I27 domains. Because of these difficulties, we focused our AFM measurements on D34 monomers (see main text). However, this approach limits our control over the pulling geometry that becomes somewhat random depending on the random attachment of the D34 monomeric protein to the substrate and an AFM tip. In addition, these traces of D34 monomers do not represent the mechanics of intact molecules but were obtained after initial stretching-relaxing cycles aimed at isolating a single molecule for force spectroscopy measurements.

SMD simulations using a coarse-grained model

The consensus AR protein NIC6 contains an N-terminal, a C-terminal, and 6 internal repeats, and is composed of 253 amino acids. The D34 protein is composed of 12 AR repeats and 408 amino acids. Initial geometries of D34 and NI6C were built based on PDB structure 1N11 and 2QYJ, respectively. The missing 5 residues between ARs to the SBD in D34 structure were inserted and minimized using all-atom CHARMM force field (10). The crystal structures were coarse-grained to C_α representation and the proteins were simulated using the structure based Go-like force field (11). The spring constants in the SMD were set to be the same as in AFM experiments. The simulations followed the same protocol as described in previous work (12). Langevin friction coefficient $\gamma = 1$ amu/ps is used in the SMD.

SUPPORTING REFERENCES

1. Wetzel, S. K., G. Settanni, M. Kenig, H. K. Binz, and A. Plückthun. 2008. Folding and Unfolding Mechanism of Highly Stable Full-Consensus Ankyrin Repeat Proteins. Journal of Molecular Biology 376:241-257.
2. Steward, A., J. L. Toca-Herrera, and J. Clarke. 2002. Versatile cloning system for construction of multimeric proteins for use in atomic force microscopy. Protein Science 11:2179-2183.
3. Lee, G., K. Abdi, Y. Jiang, P. Michaely, V. Bennett, and P. E. Marszalek. 2006. Nanospring behavior of ankyrin repeats. Nature 440:246-249.

4. Schmid, E. L., T. A. Keller, Z. Dienes, and H. Vogel. 1997. Reversible oriented surface immobilization of functional proteins on oxide surfaces. Analytical Chemistry 69:1979-1985.
5. Schmitt, L., M. Ludwig, H. E. Gaub, and R. Tampe. 2000. A metal-chelating microscopy tip as a new toolbox for single-molecule experiments by atomic force microscopy. Biophysical Journal 78:3275-3285.
6. Oberhauser, A. F., P. E. Marszalek, H. P. Erickson, and J. M. Fernandez. 1998. The molecular elasticity of the extracellular matrix protein tenascin. Nature 393:181-185.
7. Rabbi, M., and P. E. Marszalek. 2007. Measuring Protein Mechanics by Atomic Force Microscopy. Cold Spring Harb Protoc 2007:pdb.prot4901.
8. Florin, E. L., M. Rief, H. Lehmann, M. Ludwig, C. Dornmair, V. T. Moy, and H. E. Gaub. 1995. Sensing Specific Molecular-Interactions with the Atomic-Force Microscope. Biosensors & Bioelectronics 10:895-901.
9. Bustamante, C., J. F. Marko, E. D. Siggia, and S. Smith. 1994. ENTROPIC ELASTICITY OF LAMBDA-PHAGE DNA. Science 265:1599-1600.
10. Mackerell Jr, A. D., M. Feig, and I. Brooks, Charles L. . 2004. Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. Journal of Computational Chemistry 25:1400-1415.
11. Clementi, C., H. Nymeyer, and J. N. Onuchic. 2000. Topological and energetic factors: what determines the structural details of the transition state ensemble and "en-route" intermediates for protein folding? an investigation for small globular proteins. Journal of Molecular Biology 298:937-953.
12. Lee, W., X. Zeng, H.-X. Zhou, V. Bennett, W. Yang, and P. E. Marszalek. 2010. Full Reconstruction of a Vectorial Protein Folding Pathway by Atomic Force Microscopy and Molecular Dynamics Simulations. Journal of Biological Chemistry 285:38167-38172.
13. Li, L., S. Wetzel, A. Plückthun, and J. M. Fernandez. 2006. Stepwise Unfolding of Ankyrin Repeats in a Single Protein Revealed by Atomic Force Microscopy. Biophys. J. 90:L30-32.

SUPPORTING FIGURES

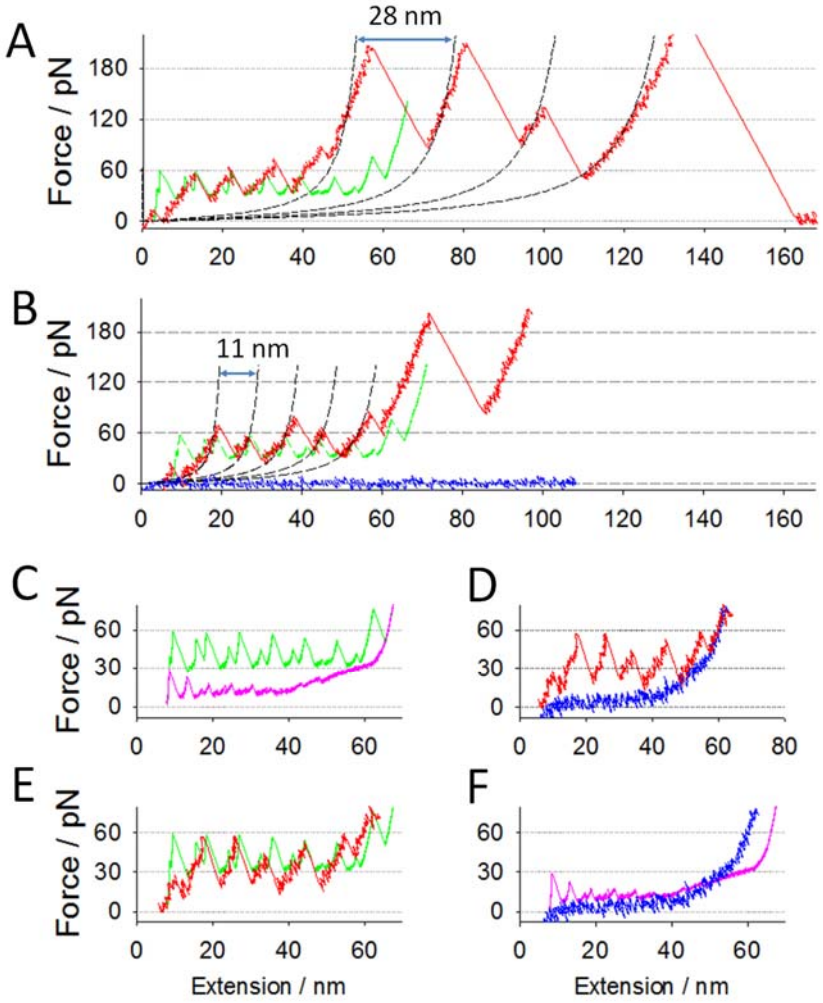


Fig. S1. Unfolding force-extension traces of $(I27)_3$-NI6C construct and the comparisons of the AFM and SMD results of NI6C.

(A) The fully stretched unfolding force-extension curve of $(I27)_3$-NI6C (red). I27 force peaks are fitted to a family of WLC curves (black dash lines) with ΔL_c = 28 nm and p = 0.36 nm. The simulated unfolding trace of NI6C-SMD2 (green) is superimposed.

(B) Another unfolding force-extension curve of $(I27)_3$-NI6C (red) and force baseline (blue). NI6C force peaks are fitted to a family of WLC curves (black dash lines) with ΔL = 11 nm and p = 0.6 nm. The simulated unfolding trace of NI6C-SMD2 (green) is superimposed. Interestingly, similar unfolding force peaks of ~60 pN were captured in the AFM study of NI4C which differs from NI6C only by the lesser number of internal consensus ARs (13). Because in that study, NI4C was not flanked by any other proteins, we assumed that the stretching forces were not applied to the terminal residues but rather to terminal regions. This could possibly force the unfolding of NI4C to proceed from the N-terminus to the C-terminus generating increased unfolding force peaks as compared to $(I27)_3$-NI6C-$(I27)_3$ study (12).

(C) The simulated unfolding (green) and refolding (pink) traces of NI6C-SMD2 (the same curves shown in Fig. 2E).

(D) The AFM unfolding (green) and refolding (pink) traces of $(I27)_3$-NI6C (the same curves shown in Fig. 2F).

(E) A comparison of the unfolding force-extension traces by SMD (green) and AFM (red).

(F) A comparison of the refolding force-extension traces by SMD (pink) and AFM (blue)

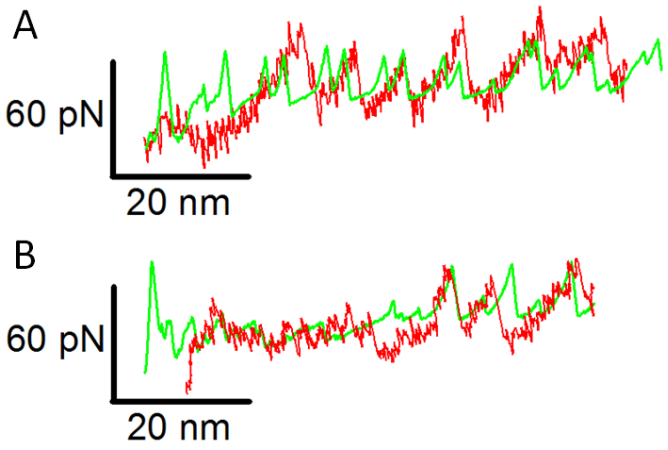


Fig. S2. Comparisons of the AFM and SMD unfolding results of D34.

(A) The unfolding force-extension curve of D34 (red, the same unfolding trace shown in Fig. 6A) and the simulated unfolding trace of D34-SMD1 (green, the same unfolding trace shown in Fig. 3A).

(B) Another unfolding force-extension curve of D34 (red, the same unfolding trace shown in Fig. 6C) and the simulated unfolding trace of D34-SMD3 (green, the same unfolding trace shown in Fig. 3C).

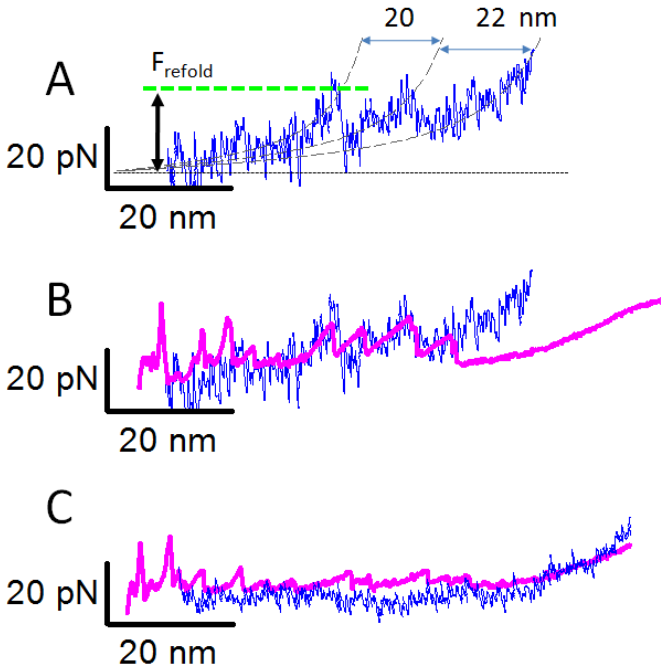


Fig. S3. Comparisons of the AFM and SMD refolding results of D34.

(A) The refolding force-extension curve of D34 (blue, the same refolding trace shown in Fig. 6*A*), fitted to WLC curves (black dash lines) with p = 0.77 nm. The maximum forces of the refolding force peaks, F_{refold}, are measured from the force baseline (black arrow and green dash line).

(B) The refolding trace in *A* (blue) and the simulated unfolding trace of D34-SMD2 (green, the same unfolding trace shown in Fig. 3*B*).

(C) Another refolding force-extension curve of D34 (red, the same refolding trace shown in Fig. 6*B*) and the simulated unfolding trace of D34-SMD1 (green, the same unfolding trace shown in Fig. 3*A*).

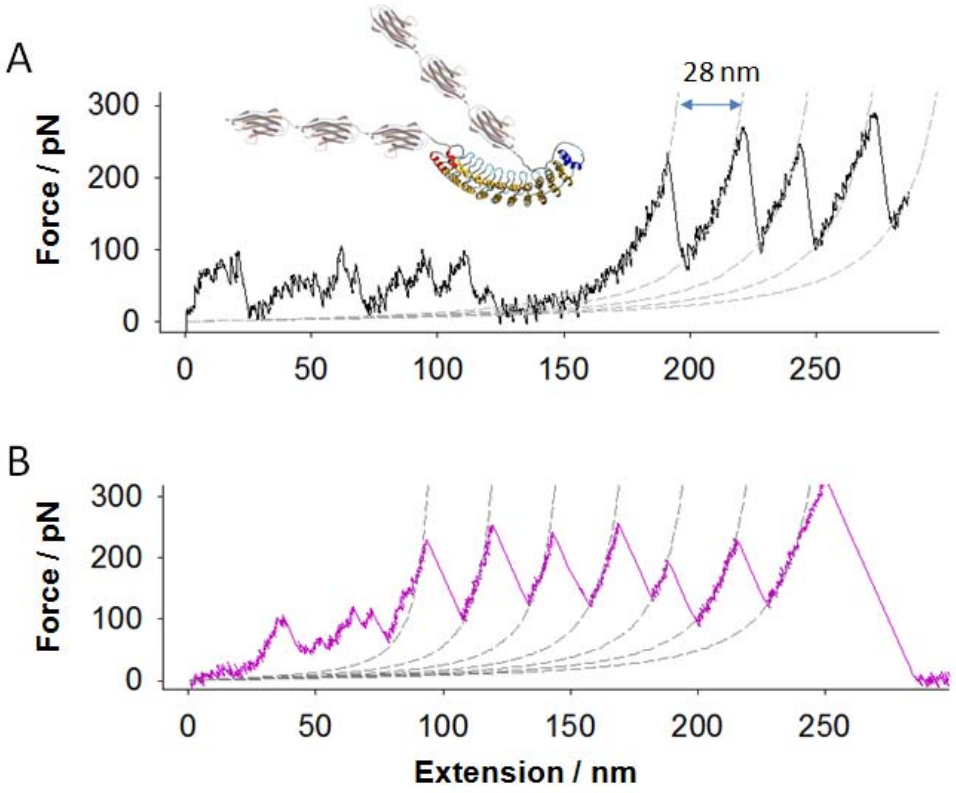
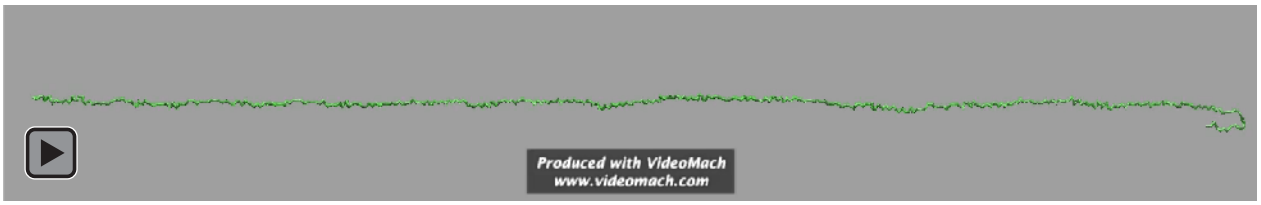


Fig. S4. Unfolding force extension traces of the $(I27)_3$-D34-$(I27)_3$ construct (inset in *A*). The I27 force peaks are fitted with family WLC curves (gray dashed lines) of ΔL_c = 28 nm and p 0.36 nm.

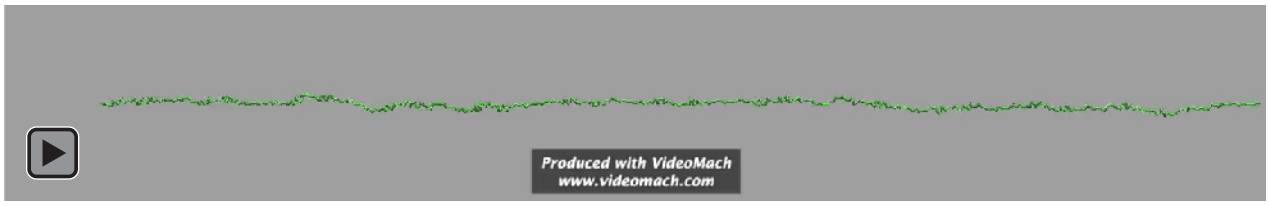
Movie S1. The unfolding process of D34-SMD1.

Movie S2. The refolding process of D34-rSMD1.

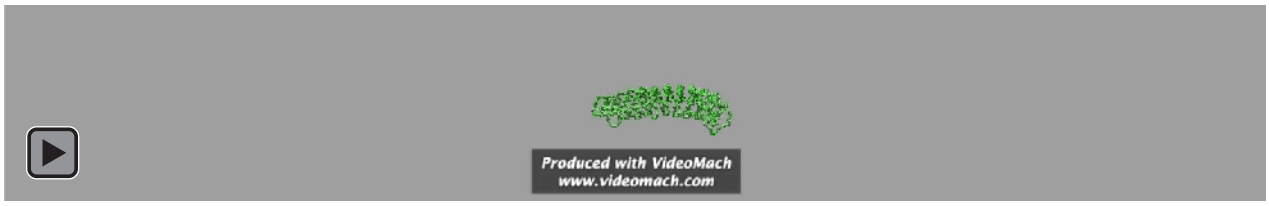


Movie S3. The unfolding process of D34-SMD2.

Movie S4. The refolding process of D34-SMD2.



Movie S5. The unfolding process of D34-SMD3.



Movie S6. The refolding process of D34-SMD3.