Single Molecule Force Spectroscopy Reveals the Mechanical Design Governing the Efficient Translocation of the Bacterial Toxin Protein RTX

Han Wang[1#], Xiaoqing Gao[2#] and Hongbin Li[1]*

Supporting Information

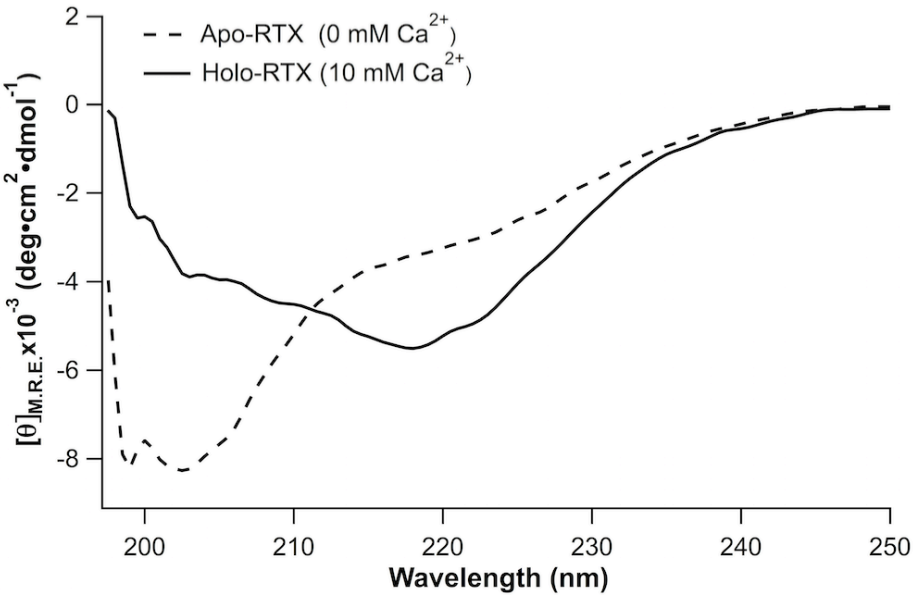

Figure S1. Far-UV CD spectra of wt-RTX in the absence of Ca^{2+} and in the presence 10 mM Ca^{2+}.

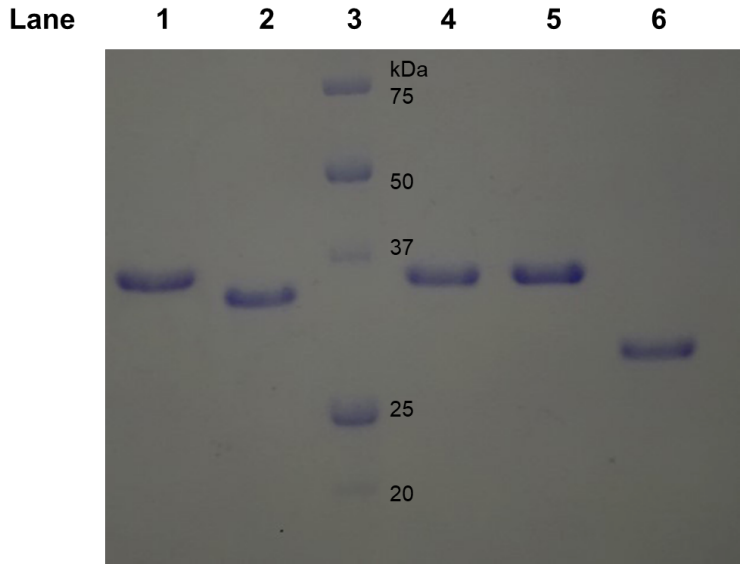

Figure S2. Coomassie blue stained SDS-PAGE photograph of RTX and RTX mutants. Lane 1, NuG2-wt RTX -NuG2 (34 kDa); Lane 2, NuG2-RTX Δhelix-NuG2 (32 kDa); Lane 3, standard protein marker; Lane 4, NuG2-RTX W1645G Y1646G-NuG2 (34 kDa); Lane 5, NuG2-RTX L5-NuG2 (34 kDa); Lane 6, NuG2-RTX A1615C (28 kDa).

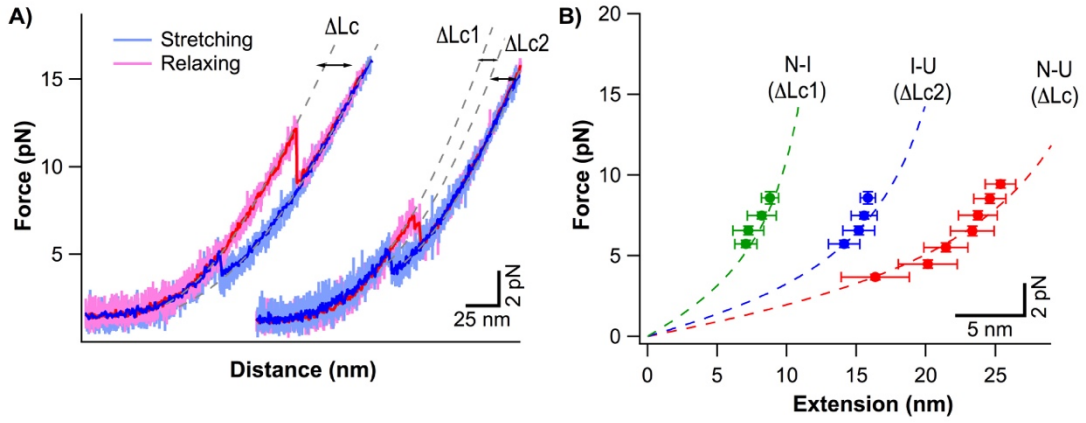

Figure S3. Optical tweezer experiments showed that the deletion of α helix at the C-terminus does not affect the unfolding and refolding pathways of RTX protein. A) Representative curves show successful unfolding and folding of RTX Δhelix protein via bifurcated pathways at a pulling speed of 50 nm s^{-1}. Grey dashed curves are fake worm-like chain fittings used to primarily help estimate the ΔLc of the unfolding and folding events. B) Force-extension relationships of RTX Δhelix protein in different unfolding pathways (N→I, I→U and N→U). Dotted lines are WLC fits to the experimental data, which yield the persistence length of 0.8 nm and ΔLc of 16.4 nm, 29.1 nm and 45.5 nm, respectively.

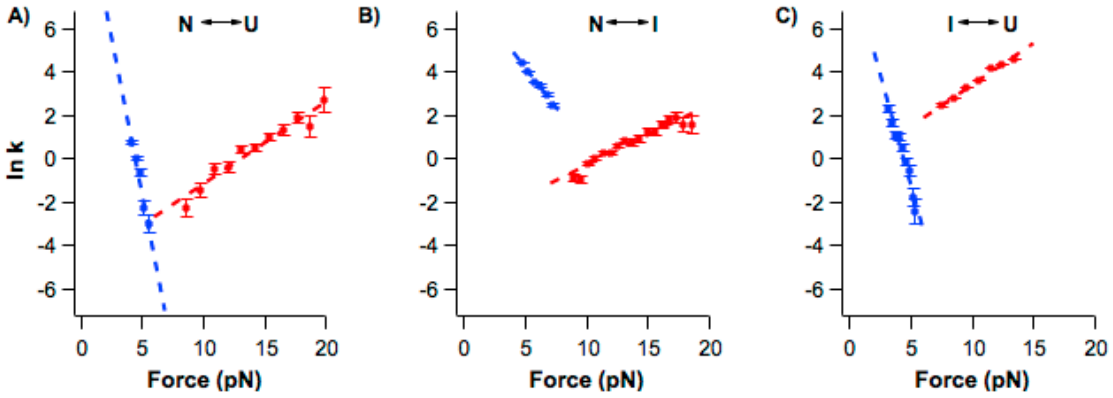

Figure S4. Signatures of the mechanical unfolding (red) and refolding (blue) of RTX protein. A,C,E) Unfolding and refolding force histograms for the pathways of (A) N ↔U, (C) N ↔I and (E) I ↔U at pulling speed of 50 nm s^{-1}, averaged unfolding/folding forces are shown in Table S1. B,D,F) Force-dependency of the unfolding and refolding rate constants of RTX for the pathways of N ↔U (A), N ↔I (B) and I ↔U (C). Unfolding is colored in red and refolding in blue. The rate constants were calculated using the Oesterhelt approach. Dotted lines correspond to the fits of Bell-Evans model to the experimetnal data.

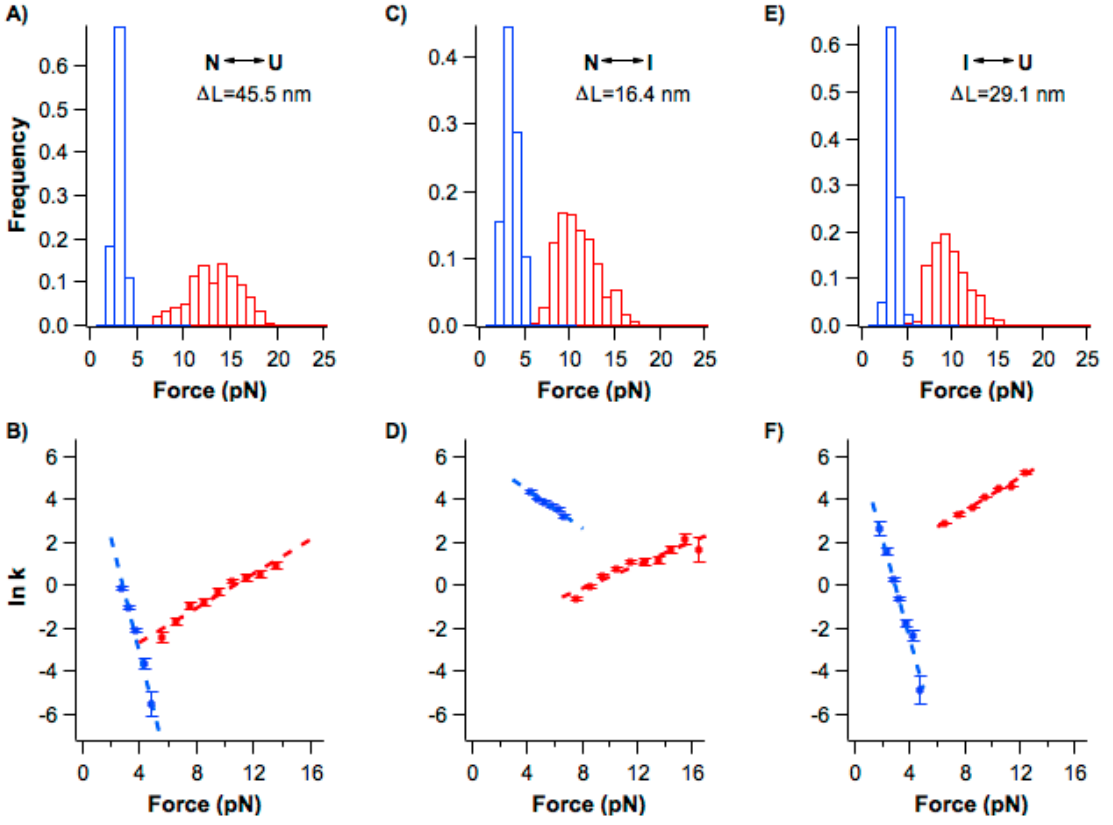

Figure S5. Signatures of the mechanical unfolding (red) and refolding (blue) of RTX Δhelix protein. A, C, E) Unfolding and refolding force histograms for the pathways of (A) N↔U, (C) N↔I and (E) I↔U at pulling speed of 50 nm s^{-1}, averaged unfolding/folding forces are shown in Table S1. B,D,F) Force-dependent unfolding and refolding rates calculated by Oesterhelt approach and dwell-time distribution analysis. Dotted lines correspond to the fitting of Bell-Evans model to the experimetnal data.

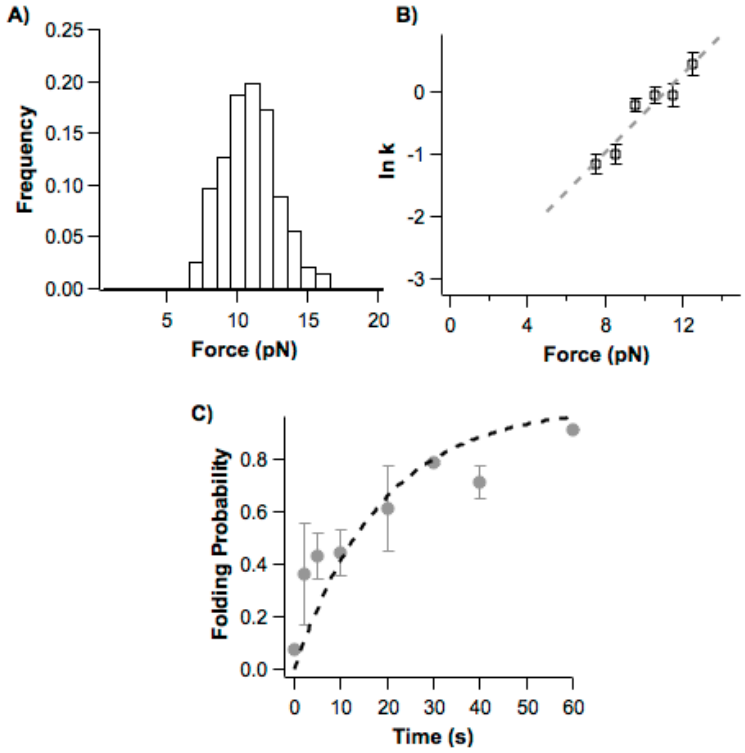

Figure S6. π-π interactions between W1645 and Y1646 is critical for the rapid folding of RTX. A) Unfolding force histograms of RTX W1645G Y1646G protein for the two-state pathway at pulling speed of 20 nm s^{-1}. B) Force-dependent unfolding rates calculated by Oesterhelt approach. Dotted lines correspond to the fitting of Bell-Evans model with α_0 of 0.026 s^{-1} and Δx_u of 1.3 nm. C) Relationship of folding probability vs folding time at zero force. Fitting the data to the first order kinetics measured a folding rate constant β_0 of 0.054 s^{-1}.

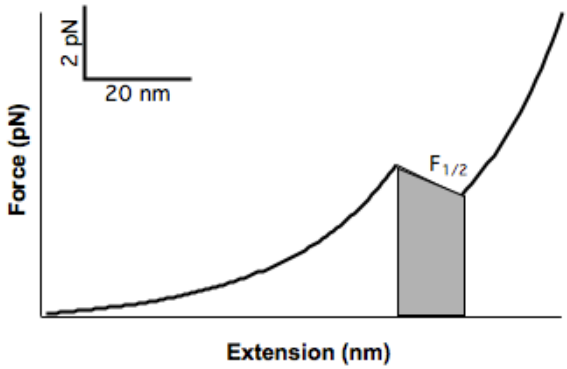

Figure S7. Work generation due to protein folding. During the folding of the RTX intermediate state, the RTX polypeptide chain shortens. The shortening of the polypeptide chain is force-dependent and follows the WLC model of polymer elasticity. $F_{1/2}$, at which the folding and unfolding rate constant equal to each other, is predicted to be ~5 pN for the folding intermediate state of RTX. At $F_{1/2}$, RTX shortens by ~13 nm, leading to the generation of work of ~15 k_BT (the shaded area).

Table S1. Averaged unfolding/folding forces and number of events for wt RTX and RTX Δhelix

		F_u (pN)	F_f (pN)	α_0 (s^{-1})	Δx_u (nm)	β_0 (s^{-1})	Δx_f (nm)
wt RTX	N-U	14.2 ± 2.7 (n=160)	3.7 ± 0.5 (n=821)	$6.9x10^{-3}$	1.6	$2.8x10^{5}$	11.7
	N-I	12.8 ± 2.5 (n=890)	4.4 ± 0.7 (n=214)	$4.6x10^{-2}$	1.1	$2.8x10^{3}$	3.1
	I-U	11.7 ± 2.2 (n=890)	3.8 ± 0.6 (n=214)	$6.8x10^{-1}$	1.6	$8.2x10^{3}$	8.4
RTXΔhelix	N-U	12.7 ± 2.7 (n=313)	2.4 ± 0.5 (n=903)	$1.5x10^{-2}$	1.6	$3.6x10^{3}$	11.7
	N-I	10.3 ± 2.3 (n=685)	2.9 ± 0.8 (n=113)	$1.0x10^{-1}$	1.1	$2.7x10^{3}$	3.1
	I-U	9.1 ± 2.1 (n=685)	2.8 ± 0.5 (n=113)	1.4	1.6	$3.7x10^{2}$	8.4

Supplementary Methods

Protein engineering

The plasmid encoding RTX with the restriction sites (5' *Bam*HI and 3' *Bgl*II and *Kpn*I) was a generous gift from Dr. Shanshan Lv from Beijing University of Chemical Technology. Following a well-established iterative molecular biology strategy[1] we constructed the gene of NuG2-wt RTX -NuG2. The NuG2-wt RTX -NuG2 gene was then subcloned into the expressing vector pQE80L-CC, which was constructed as described[2,3], to build pQE80L/Cys-NuG2-RTX-NuG2-Cys. The full sequence of the engineered Cys-NuG2-RTX-NuG2-Cys is as follows:
MRGSHHHHHHGSCKRS*MDTYKLVIVLNGTTFTYTTEAVDAATAEKVFKQYANDN GVDGEWTYDDATKTFTVTE*RS**GSARDDVLIGDAGANVLNGLAGNDVLSGGA GDDVLLGDEGSDLLSGDAGNDDLFGGQGDDTYLFGVGYGHDTIYESGG GHDTIRINAGADQLWFARQGNDLEIRILGTDDALTVHDWYRDADHRVEII HAANQAVDQAGIEKLVEAMAQYPDP**RS*MDTYKLVIVLNGTTFTYTTEAVDAAT AEKVFKQYANDNGVDGEWTYDDATKTFTVTE*RSGGTKC

where the sequence in bold corresponds to the RTX protein and the sequence in italic corresponds to NuG2 domains. Cysteine residues at N- and C-termini are underlined.

The gene encoding RTX Δhelix was constructed using polymerase chain reaction. The mutant RTX W1645G Y1646G and RTX A1615C were obtained using megaprimer approach of site-directed mutagenesis. In the mutant RTX-L5, 5 additional amino residues GGGLG was inserted between residue 1635-1636 located in one of the flexible loop of RTX using a previously described method[4]. All of the constructed genes

were confirmed by DNA sequencing.

Protein overexpression was carried in *Escherichia coli* strain DH5α at 37 °C in 250 mL 2.5% LB media supplemented with 100 mg L^{-1} antibiotics ampicillin. When the optical density at 600 nm reached between 0.8-1.0, protein overexpression was induced with 1 mM isopropyl-β-D-1-thiogalactopyranoside (IPTG, Thermo Fisher Scientific, Waltham, MA) and the expression continued for 4 hours at 37 °C. The bacteria cell pellets were harvested by centrifugation at 5,000 rpm at 4 °C for 10 minutes and resuspended in 10 mL phosphate buffered saline (PBS, 10 mM, pH 7.4) buffer. Cell lysis was carried out by adding 10 μL protease inhibitor cocktail (SIGMA-ALDRICH, St. Louis, MO), 50 μL 100 mg mL^{-1} lysozyme from egg white (SIGMA-ALDRICH, St. Louis, MO), 1 mL 10% (w/v) Triton X-100 (VWR, Tualatin, OR), 50 μL 1 mg mL^{-1} DNase I (SIGMA-ALDRICH, St. Louis, MO) and 50 μL 1 mg mL^{-1} RNase A (Bio Basic Canada Inc, Markham, ON) at 4 °C for 40 minutes. The supernatant containing the protein was isolated by centrifugation at 12,000 rpm at 4 °C for 1 hour, and polyprotein is purified by Co^{2+} affinity column with TALON His-tag purification kit (TaKaRa Bio USA Inc, Mountain View, CA). The polyprotein was eluted and stored in elution buffer (10 mM PBS, 300 mM NaCl, 250 mM imidazole). The purified apo-form RTX polyprotein (free of Ca^{2+}) was at a concentration of ~1.0 mg mL^{-1} and stored at -20 °C. Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) confirmed the expression and purity of the expressed proteins (Fig. S1).

Preparation of DNA-protein chimera

Double-strand DNA (dsDNA) handles were prepared via the methods described previously[5]. Two dsDNA handles of 802 and 558 bp were generated via regular PCR amplification. The template pGEMEX-1 plasmid was a generous gift from Professor X. F. Zhang of Lehigh University. The modified primers were purchased from Integrated DNA Technologies (IDT Inc, San Jose, CA). The reverse primer was 5'-NH_2 modified, and the forward primers were 5'-biotin and 5'-digoxigenin modified, respectively. The sequences of the primers were as follows:

Forward primer for 802 bp dsDNA handles: 5'-Bio-CAA-AAA-ACC-CCT-CAA-GAC-CC

Reverse primer for 802 bp dsDNA handles: 5'-NH_2-CGA-CGA-TAA-ACG-TAA-GGA-CAT-C

Forward primer for 558 bp dsDNA handles: 5'-Dig-CAA-AAA-ACC-CCT-CAA-GAC-CC

Reverse primer for 558 bp dsDNA handles: 5'-NH_2-GCT-ACC-GTA-ATT-GAG-ACC-AC

After the PCR amplification, QIAquick PCR purification kit (QIAGEN, Germantown, MD) was used to purify the PCR products. Subsequently, dsDNA handles were allowed to react with 4-(N-Maleimidomethyl) cyclohexanecarboxylic acid N-hydroxysuccinimide ester (SMCC, SIGMA-ALDRICH, St. Louis, MO) overnight, which enabled the amine group at the end of the dsDNA handles replaced by maleimide group. 50 μM of the freshly expressed proteins were reduced with 1

mM tris(2-carboxyethyl) phosphine (TCEP, SIGMA-ALDRICH, St. Louis, MO) for 1 hour, and remaining TCEP was removed by Zeba desalting columns (7kMW, Thermo Fisher Scientific, Waltham, MA). Then the reduce proteins were diluted to ~3 μM by Tris buffer (20 mM Tris, 150 mM NaCl, pH 7.4), the same concentration with the mixed dsDNA handles with two lengths. 1 μL of the diluted protein was added into 1 μL of 3 μM of mixed dsDNA handles, and the thiol-maleimide reaction was kept at room temperature overnight. The formed dsDNA-protein chimera was diluted by Tris buffer (20 mM Tris, 150 mM NaCl, pH 7.4) to ~10 nM and ready for optical tweezers experiment.

Single molecule optical tweezers experiment

The single molecule optical tweezers experiments were carried out on a Minitweezers setup[6]. During a typical single molecule OT experiment, 1 μL of streptavidin-coated polystyrene beads (1% w/v 1 μm, Spherotech Inc, Lake Forest, IL) was diluted by 3 mL Tris buffer (20 mM Tris, 150 mM NaCl, pH 7.4) in a 5 mL syringe and then injected into the liquid chamber. One streptavidin-coated bead was trapped by the laser beam and moved onto the tip of a glass micropipette. The bead was sucked tightly by the glass micropipette by applying a vacuum. Next, 1 μL of the 10 nM dsDNA-protein chimera was allowed to react with 4 μL of anti-digoxigenin-coated polystyrene beads (0.5% w/v, 2 μm, Spherotech Inc, Lake Forest, IL) for 30 minutes. The beads carrying the chimera on their surface were also diluted by 3 mL Tris buffer (20 mM Tris, 150 mM NaCl, pH 7.4) in a 5 mL syringe and injected into the liquid chamber. One anti-digoxigenin-coated bead was then captured by laser beam and allowed to touch the surface of the sucked streptavidin-coated bead. Once the interaction forms between biotin tethered at the end of dsDNA handle and streptavidin on the polystyrene bead, single dsDNA-protein chimera could be stretched and relaxed by moving the laser trap (Fig. 1B) to measure the force-distance curve. For experiments on holo-RTX, Tris buffer (20 mM Tris, 150 mM NaCl, pH 7.4) supplemented with 10 mM Ca^{2+} was used. Due to the fact that the stiffness of MiniTweezer is force-dependent, it is challenging to directly convert force-distance curve to force-extension curve. To measure the contour length increment ΔLc upon RTX unfolding, force-extension curves were determined from each pair of force-distance curves. Extension was determined by subtracting $distance_{stretching}$ from $distance_{relaxation}$. Fitting the measured force-extesnion relationship to the WLC model of polymer elasticiltiy allowed for the determination of ΔLc for each unfolding/refolding event.

Calculating the kinetics of unfolding/folding of proteins

We used the method proposed by Oesterhelt et al[7] to measure the force-dependent unfolding/folding rate constants of proteins from constant pulling velocity experiments. The force-dependency of the unfolding (or folding) rate constants was then fitted to the Bell-Evans model to extract the intrinsic unfolding and folding rate constants at zero force (α0 and β0, respectively), as well as the distantance from the native state to the transition state (Δx_u) and the distance between the unfolded and

transition state (Δx_f).

To measure the folding rate of the variant RTX W1645G Y1646G, a standard double-pulse protocol was applied. After the molecule was stretched to unfold, it was quickly relaxed to zero force and allowed to refold for a specific time window Δt. In the second pulse, the molecule was stretched again to check if the RTX variant has refolded within the time window of Δt. By varying Δt, the folding probability as a function of Δt can be determined. By fitting the data to the following equation $P(t)=1-\exp(\beta_0 \cdot t)$, the folding rate constant β_0 can be determined.

References

(1) Cao, Y.; Li, H. Polyprotein of GB1 is an ideal artificial elastomeric protein *Nat. Mater.* **2007**, *6*, 109.

(2) He, C.; Li, S.; Gao, X.; Xiao, A.; Hu, C.; Hu, X.; Hu, X.; Li, H. Direct observation of the fast and robust folding of a slipknotted protein by optical tweezers *Nanoscale* **2019**, *11*, 3945.

(3) Lei, H.; He, C.; Hu, C.; Li, J.; Hu, X.; Hu, X.; Li, H. Single-Molecule Force Spectroscopy Trajectories of a Single Protein and Its Polyproteins Are Equivalent: A Direct Experimental Validation Based on A Small Protein NuG2 *Angew. Chem. Int. Ed. Engl.* **2017**, *129*, 6213.

(4) Li, H.; Wang, H.-C.; Cao, Y.; Sharma, D.; Wang, M. Configurational entropy modulates the mechanical stability of protein GB1 *J. Mol. Biol.* **2008**, *379*, 871.

(5) Zhang, X.; Halvorsen, K.; Zhang, C.-Z.; Wong, W. P.; Springer, T. A. Mechanoenzymatic cleavage of the ultralarge vascular protein von Willebrand factor *Science* **2009**, *324*, 1330.

(6) Smith, S.; Riveti, C. http://tweezerslab.unipr.it/cgi-bin/mt/home.pl.

(7) Oberbarnscheidt, L.; Janissen, R.; Oesterhelt, F. Direct and model free calculation of force-dependent dissociation rates from force spectroscopic data *Biophys. J.* **2009**, *97*, L19.