Supporting information

Sequential protein unfolding through a carbon nanotube pore

Zhonghe Xu[1], Shuang Zhang[1,2], Jeffrey K. Weber[3], Binquan Luan[3], Ruhong Zhou[3,4*], Jingyuan Li[1*]

[1] Key Laboratory for Biomedical Effects of Nanomaterials and Nanosafety, Institute of High Energy Physics, Chinese Academy of Sciences, Beijing 100049, P. R. China

[2] Department of Physics, Zhejiang University, Hangzhou 310027, P. R. China

[3] IBM Thomas J. Watson Research Center, Yorktown Heights, New York 10598, United States

[4] Department of Chemistry, Columbia University, New York, New York 10027, United State

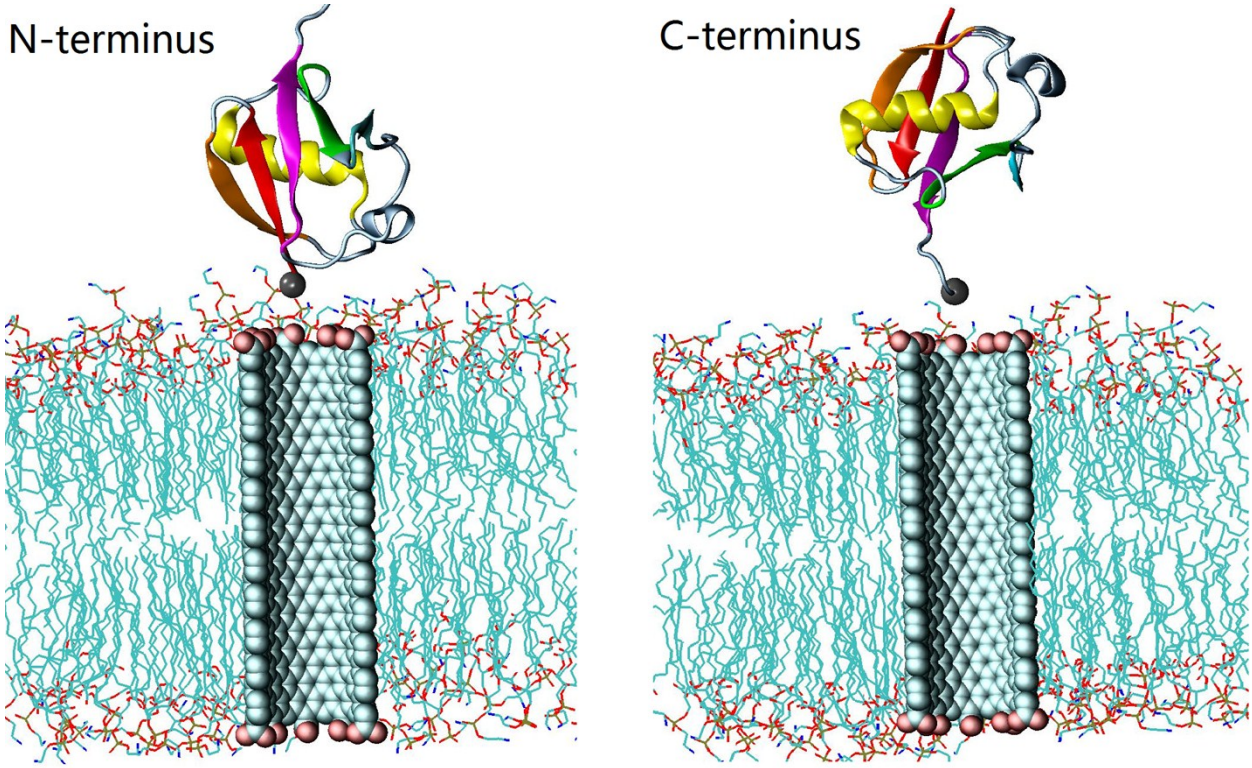

Figure S1: The system setup of protein unfolding through CNT-based nanopore. Water and ions are not shown for clarity. The protein was initially positioned such that the chosen pulling terminus (i.e. the N-terminus, left; and C-terminus, right) was situated 8 Å above the CNT entrance. After a 5-ns equilibration, the protein terminus (rendered as a sphere) was then threaded into the nanopore via steered molecular dynamics.

Table S1: Protein fragment designations corresponding to structure motifs detected during nanopore-mediated unfolding processes.

Fragments	Residues range	Fragments	Residues range
C1	1-50	N1	10-72
C2	1-45	N2	22-72
C3	1-35	N3	40-72
C4	1-17		

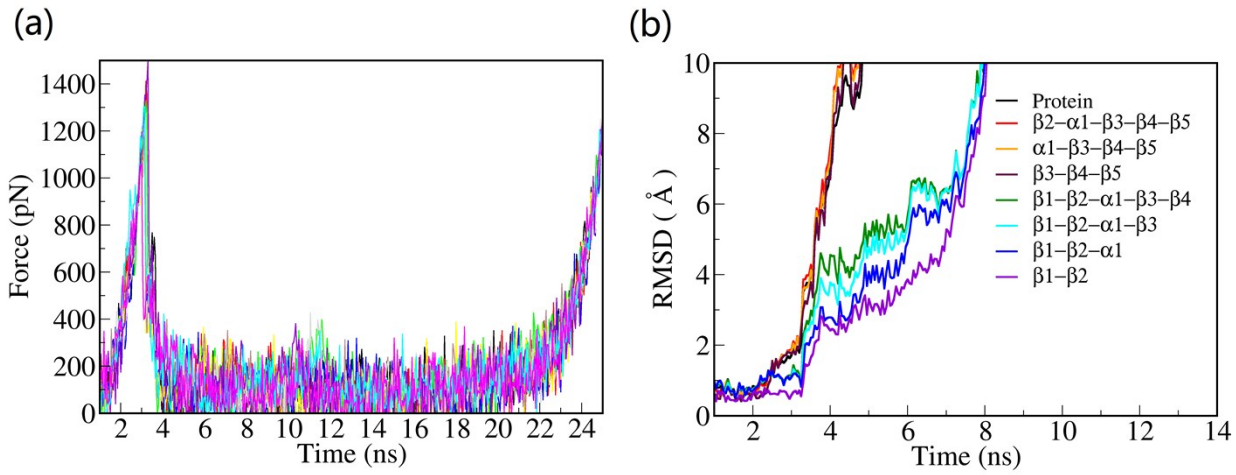

Figure S2: Biterminal unfolding at a pulling rate of 1nm/ns. Left: force profiles (ten trajectories); right: RMSDs of protein and its combinations of secondary structures in a representative trajectory. There is only one clear force profile peak appears in the early stages of stretching ($t \approx 3$ns), corresponding to the initial unfolding of protein. All RMSDs of the combinations of secondary structures sharply increases at the same time, suggesting a dramatic disassembly event.

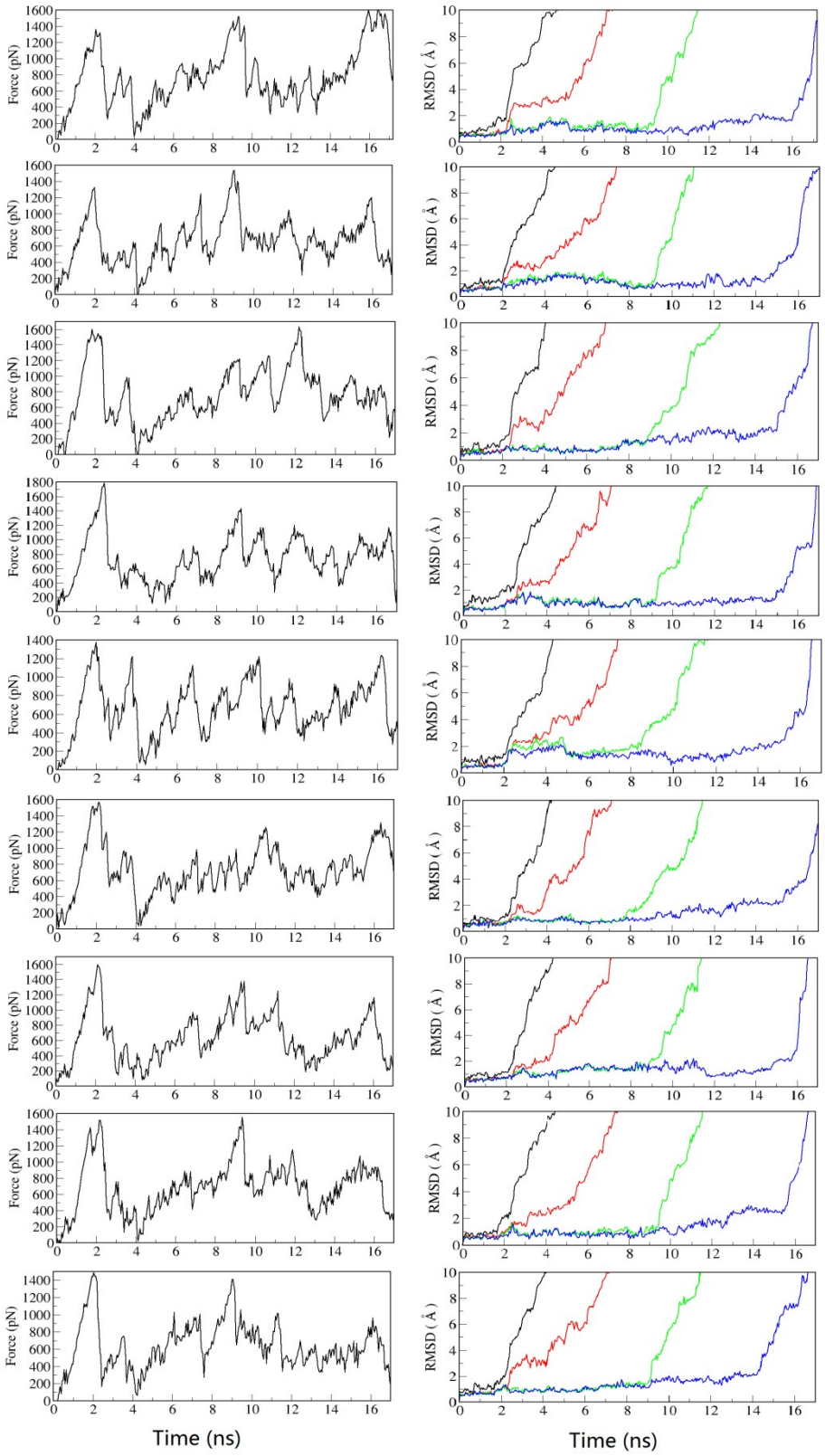

Figure S3: N-terminal unfolding at a pulling rate of 1 nm/ns. Left column: force profiles; right column: RMSDs of ubiquitin (black) and its N1 (red), N2 (green) and N3 (blue) motifs. The sequential detachment of secondary structure element can be found throughout ten independent simulations, and the remaining part always maintains native like characteristics (i.e. unfoldon motif). In addition, the correlation between the detachments of secondary structure elements and the distinct force peaks can also be found in all simulations.

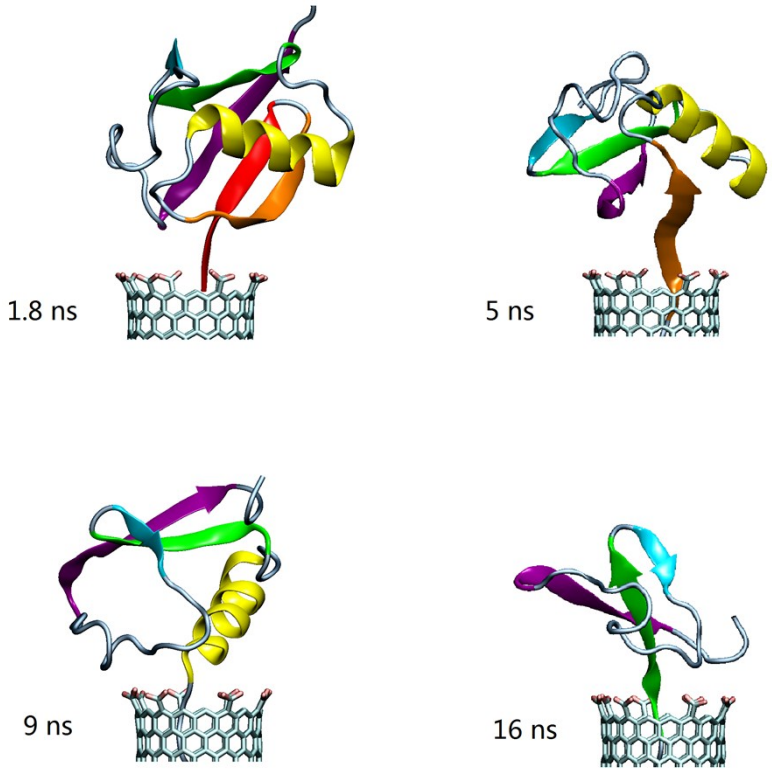

Figure S4: Snapshots of protein during N-terminal unfolding (1 nm/ns, same trajectory as figure 1). The protein underwent considerable rotation to allow the secondary structure elements to dissociate and thread into nanopore.

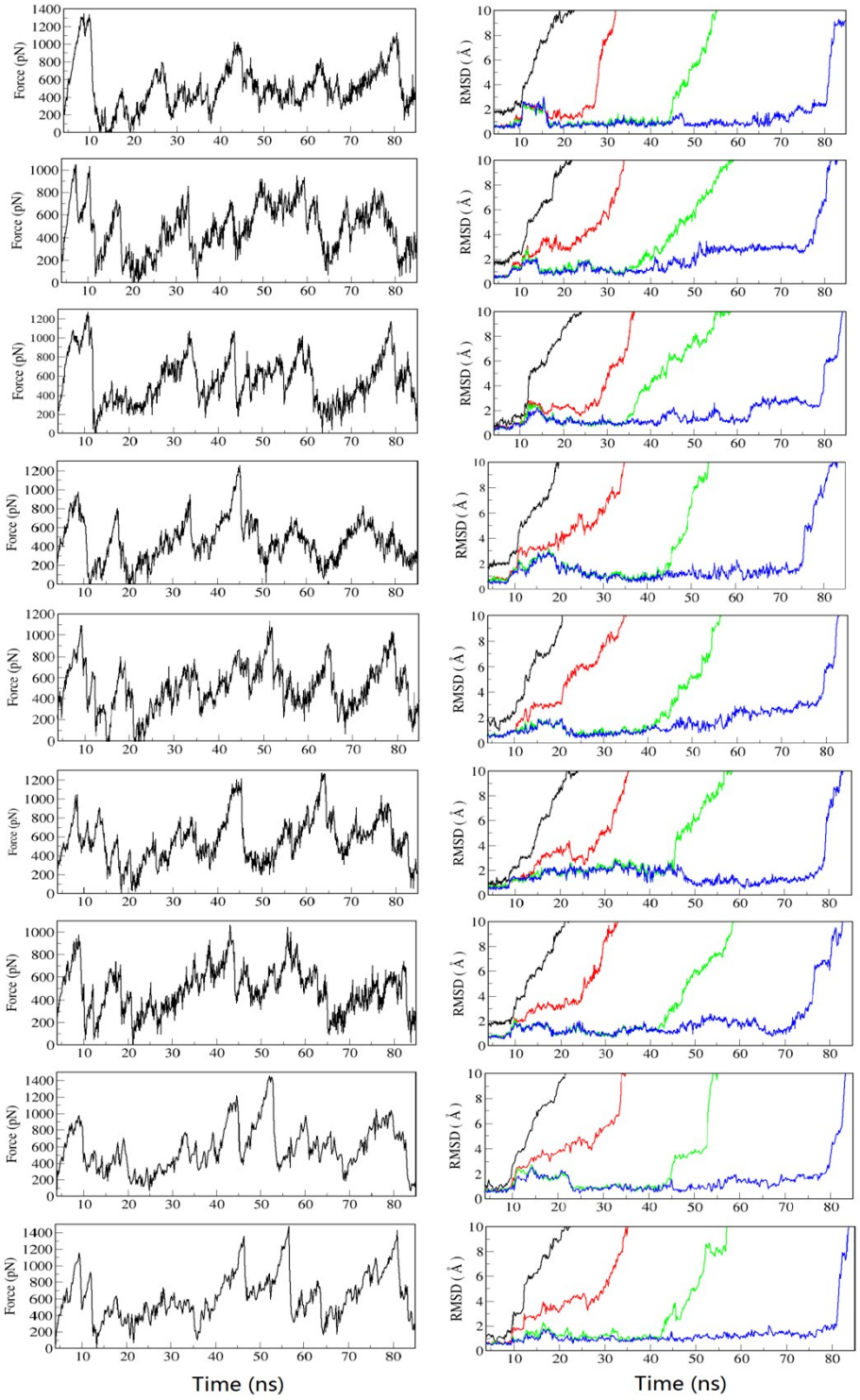

Figure S5: N-terminal unfolding at a pulling rate of 0.2 nm/ns. Left column: force profiles; right column: RMSDs of ubiquitin (black) and its N1 (red), N2 (green) and N3 (blue) motifs. The sequential detachment of secondary structure element can be found throughout ten independent simulations, and the remaining part always maintains native-like characteristics (i.e. unfoldon motif). In addition, the correlation between the detachments of secondary structure elements and the distinct force peaks can also be found in all simulations.

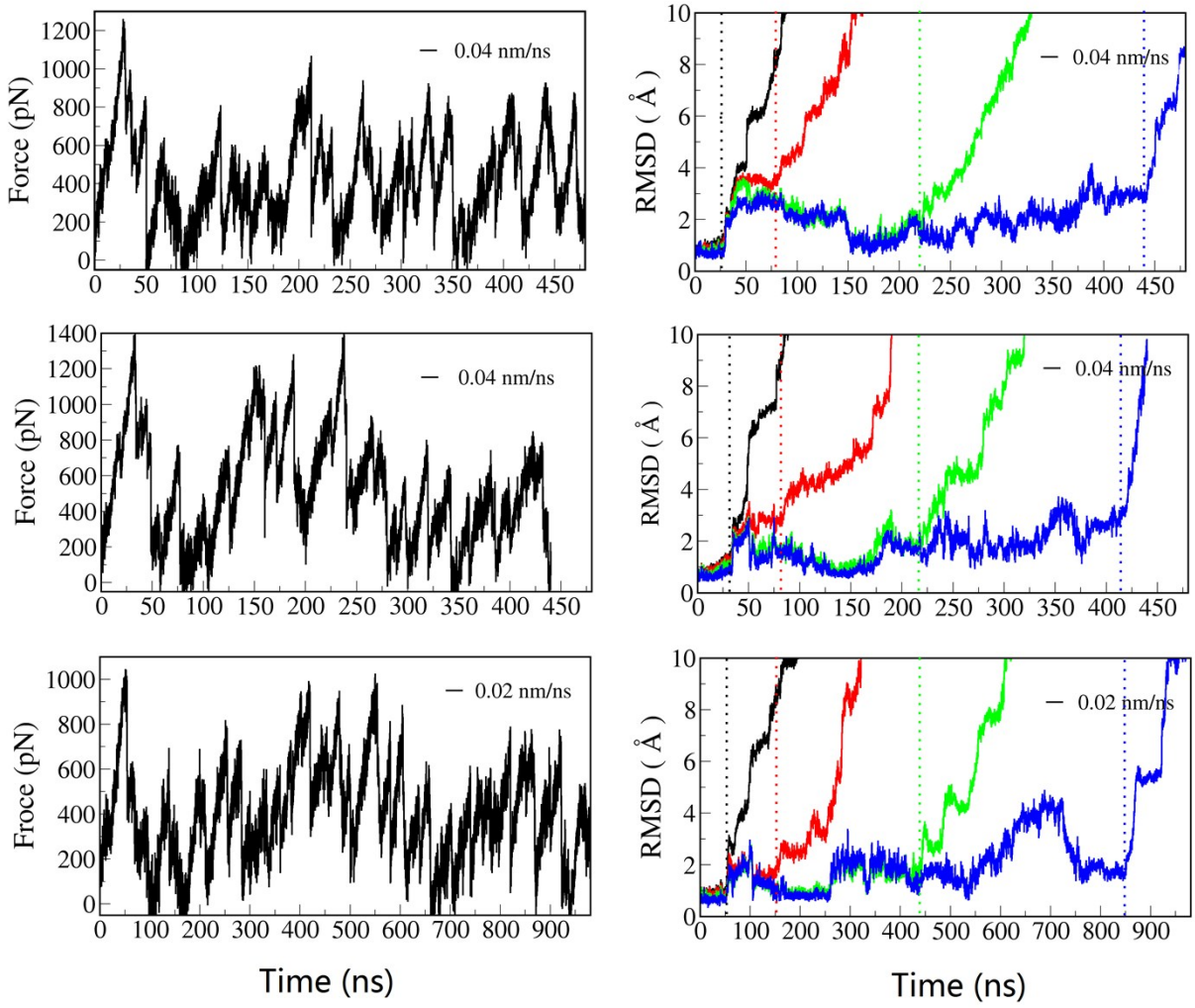

Figure S6 N-terminal unfolding at a pulling rate of 0.04 and 0.02 nm/ns. Left column: force profiles; right column: RMSDs of ubiquitin (black) and its N1 (red), N2 (green) and N3 (blue) motifs. All the features of nanopore-mediated protein unfolding, i.e. sequential detachment of secondary structure elements, stable unfoldon motifs and force-structure correlation are retained in the unfolding at much slower pulling rate.

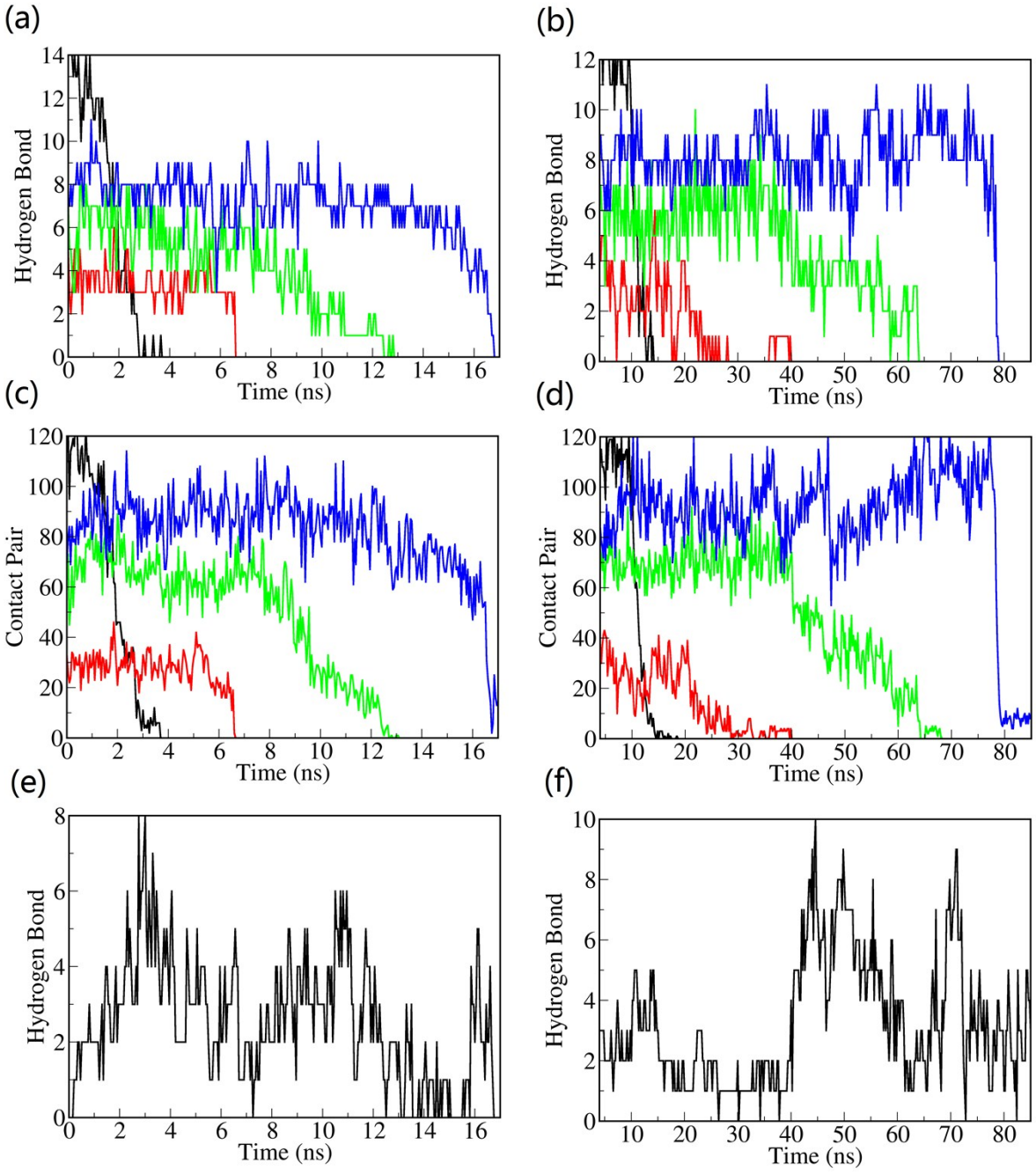

Figure S7: N-terminal unfolding at pulling rates of 1nm/ns (left) and 0.2nm/ns (right). Top: the number of hydrogen bonds between the secondary structural elements β1 (black), β2 (red), α1 (green) and β3 (blue) and the remainder of protein. Middle: the number of heavy atom contact pairs between β1 (black), β2 (red), α1 (green) and β3 (blue) and complementary protein atoms. Bottom: the number of hydrogen bonds between the protein and nanopore (i.e. the CNT end and the surrounding lipids).

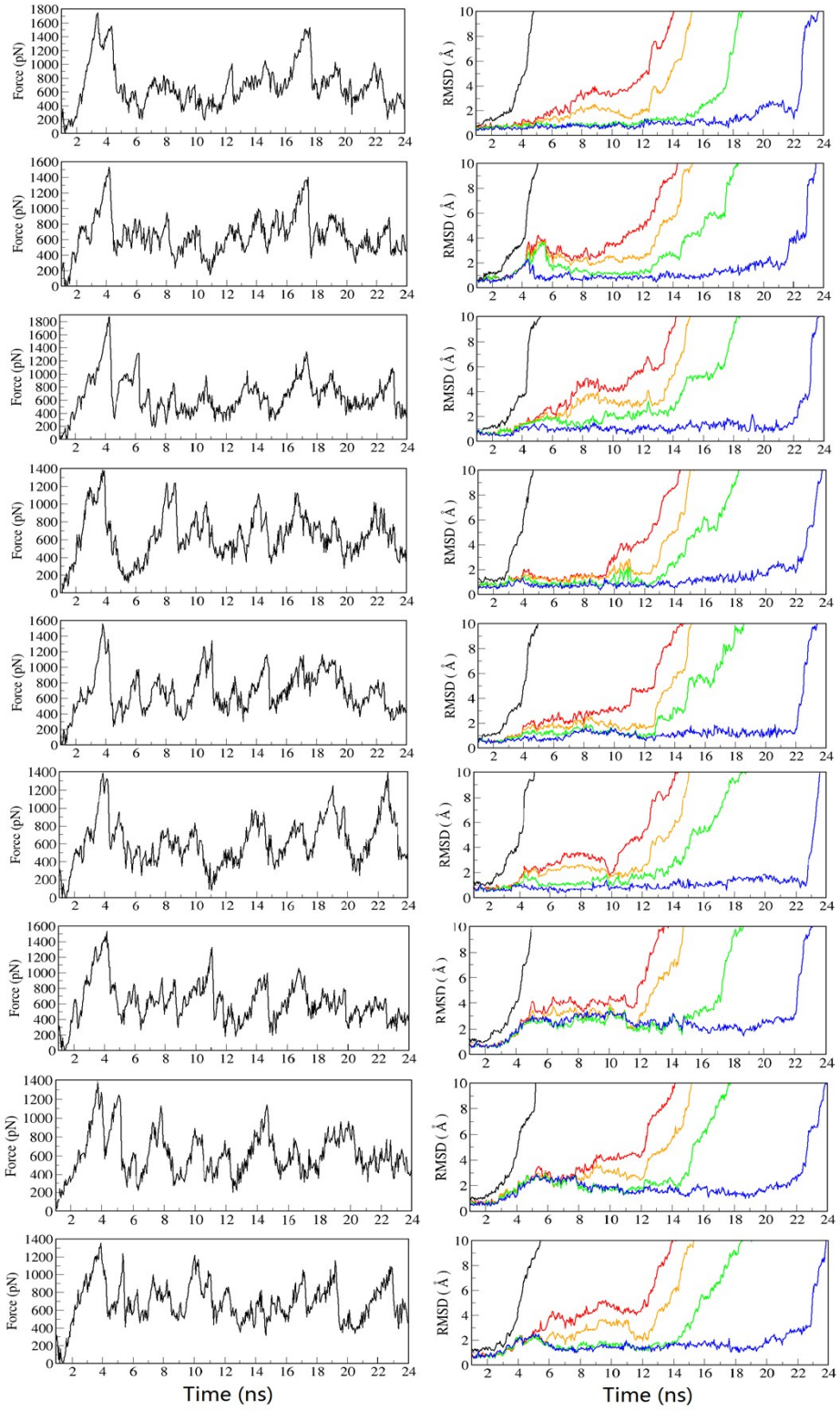

Figure S8: C-terminal unfolding at a pulling rate of 1 nm/ns. Left column: force profiles; right column: RMSDs of ubiquitin (black) and its C1 (red), C2 (orange), C3 (green) and C4 (blue) motifs. The sequential detachment of secondary structure element can be found throughout ten independent simulations, and the remaining part always maintains native-like characteristics (i.e. unfoldon motif). In addition, the correlation between the detachments of secondary structure elements and the distinct force peaks can also be found in all simulations.

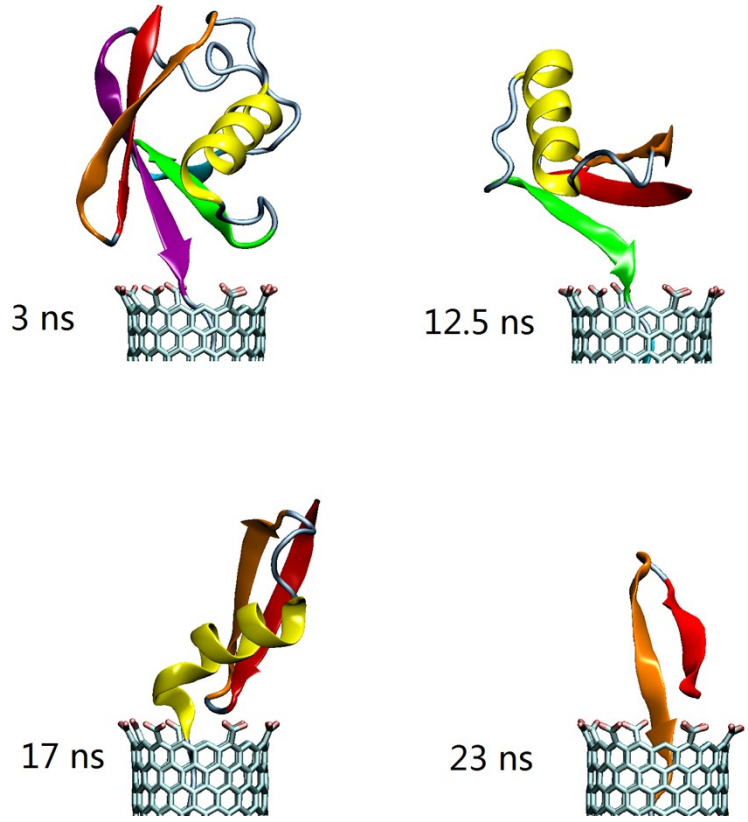

Figure S9: Snapshots of protein during C-terminal unfolding (1 nm/ns, same trajectory as figure 3). The protein underwent considerable rotation to allow the secondary structure elements to dissociate and thread into nanopore.

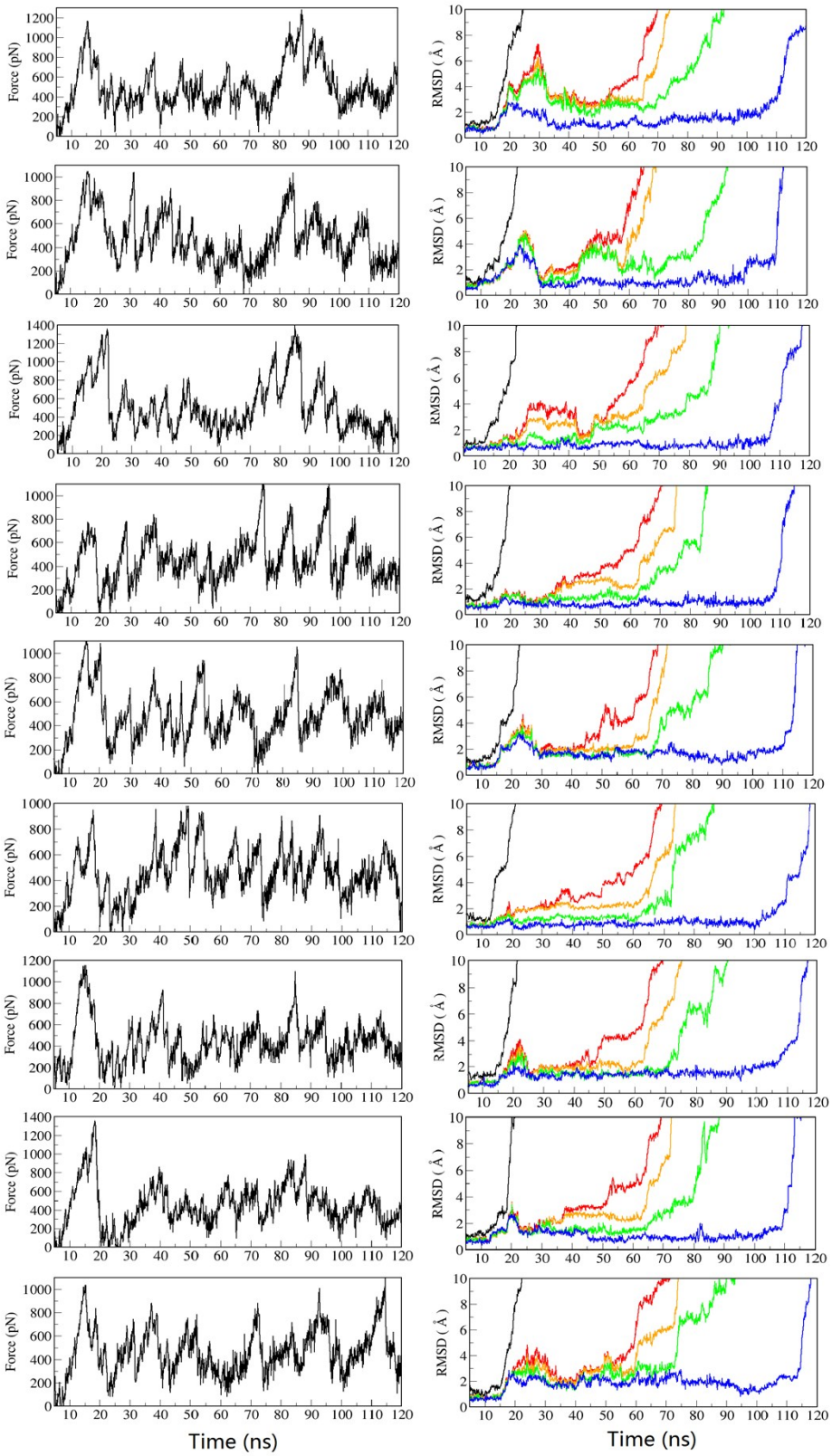

Figure S10: C-terminal unfolding at a pulling rate of 0.2 nm/ns. Left column: force profiles; right column: RMSDs of ubiquitin (black) and its C1 (red), C2 (orange), C3 (green) and C4 (blue) motifs. The sequential detachment of secondary structure element can be found throughout ten independent simulations, and the remaining part always maintains native-like characteristics (i.e. unfoldon motif). In addition, the correlation between the detachments of secondary structure elements and the distinct force peaks can also be found in all simulations.

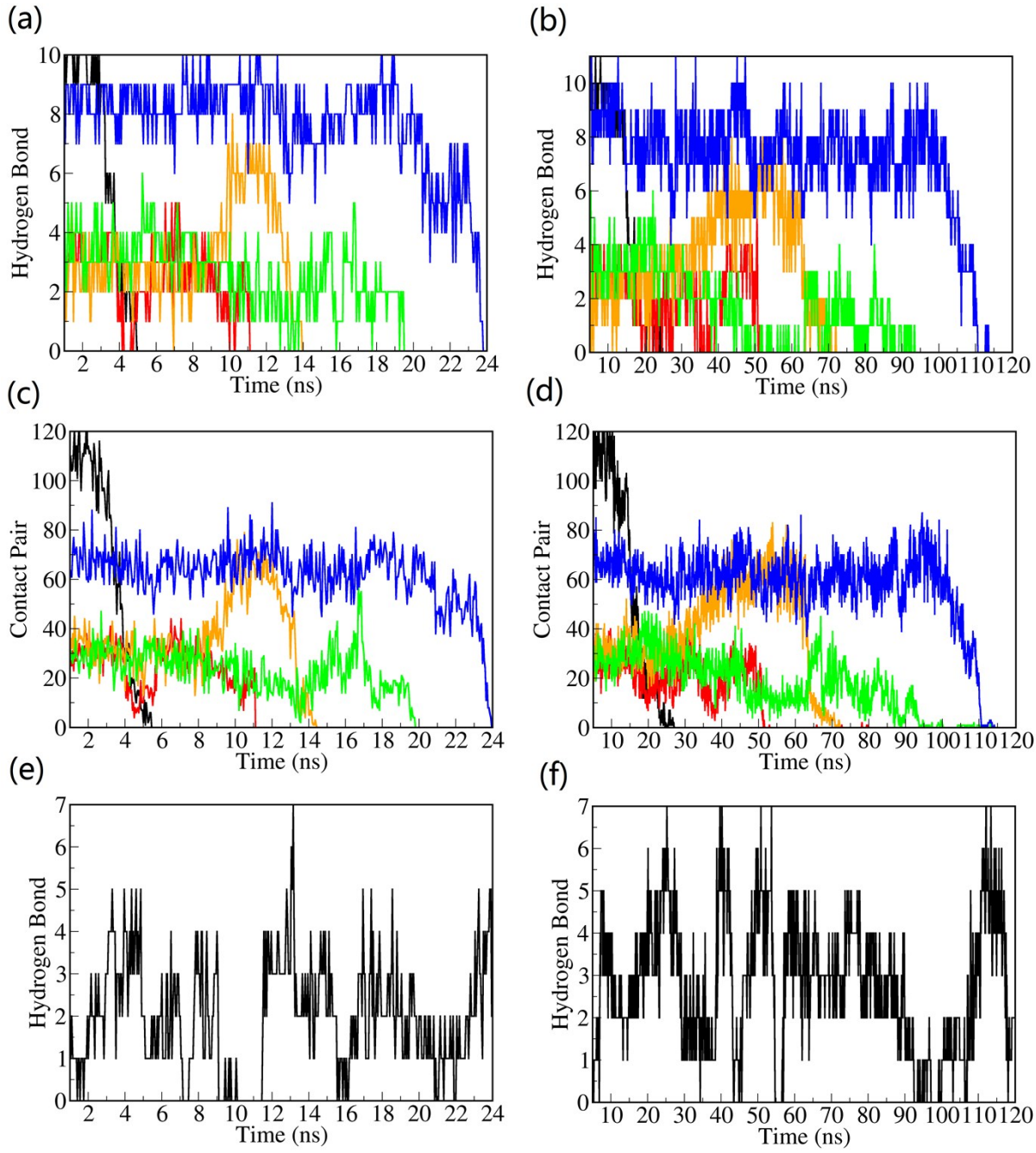

Figure S11: C-terminal unfolding at pulling rates of 1 nm/ns (left), and 0.2 nm/ns (right). Top: the number of hydrogen bonds between the secondary structural elements β5 (black), β4 (red), β3 (orange), α1 (green) and β2 (blue) and the remainder of protein. Middle: the number of heavy atom contacts between β5 (black), β4 (red), β3 (orange), α1 (green) and β2 (blue) and complementary protein atoms. Bottom: the number of hydrogen bonds between the protein and nanopore (i.e. the CNT end and the surrounding lipids) during C-terminal unfolding.

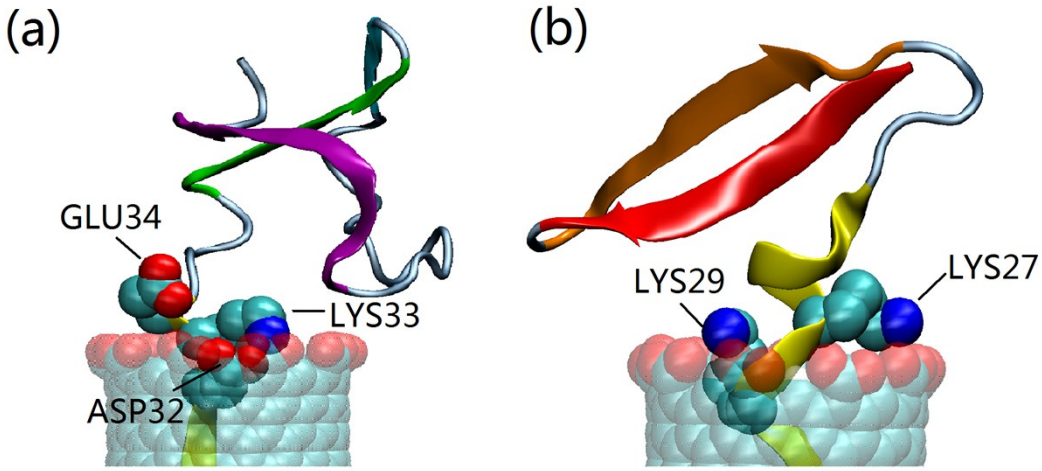

Figure S12: Representative configurations that illustrate the trapping of residues around nanopore entrance. a) ASP32, LYS33 and GLU34 trapped at the nanotube aperture (t=64 ns, N-terminal unfolding at a pulling rate of 0.2nm/ns). b) two positively charged residues -- LYS27 and LYS29 -- trapped at the nanopore entrance (t=82 ns, C-terminal unfolding at a pulling rate of 0.2nm/ns).

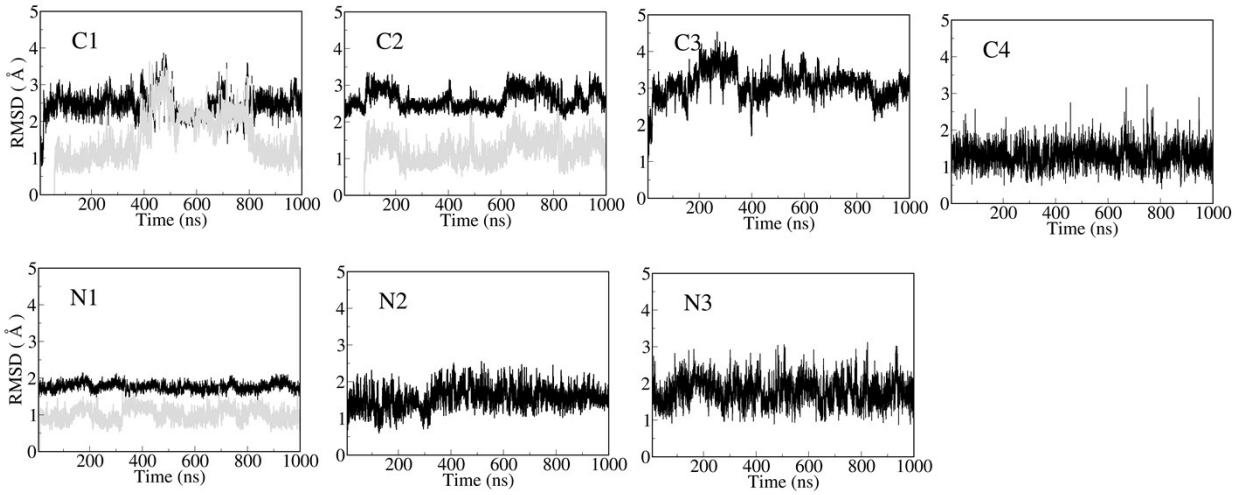

Figure S13: RMSDs of protein fragments corresponding to the motifs C1, C2, C3, and C4 (first row) and the motifs N1, N2, N3 (second row). RMSDs of protein fragments with respect to corresponding crystal structure fragments are presented in black, whereas RMSDs with respect to configurations resulting from initial structural rearrangements are shown in gray.