Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.

Supplementary data

Nonnative Contact Effects in Protein Folding

Qiang Shao[1,2*], Weiliang Zhu[1,2]

[1]Drug Discovery and Design Center, CAS Key Laboratory of Receptor Research, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, 555 Zuchongzhi Road, Shanghai, 201203, China

[2]University of Chinese Academy of Sciences, Beijing, 100049, China

[*]To whom correspondence should be addressed. Qiang Shao, Tel: +86 21 50806600-1304, E-mail: qshao@simm.ac.cn.

Protein	Sequence
NTL9	MKVIFLKDVKGMGKKGEIKNVADGYANNFLFKQGLAIEA
NuG2b	DTYKLVIVLNGTTFTYTTEAVDAATAEKVFKQYANDAGVDGEWT YDAATKTFTVTE
CspA	SGKMTGIVKWFNADKGFGFITPDDGSKDVFVHFSAIQNDGYKSLDE GQKVSFTIESGAKGPAAGNVTSL

Table S1. Amino acid sequences of proteins in this work.

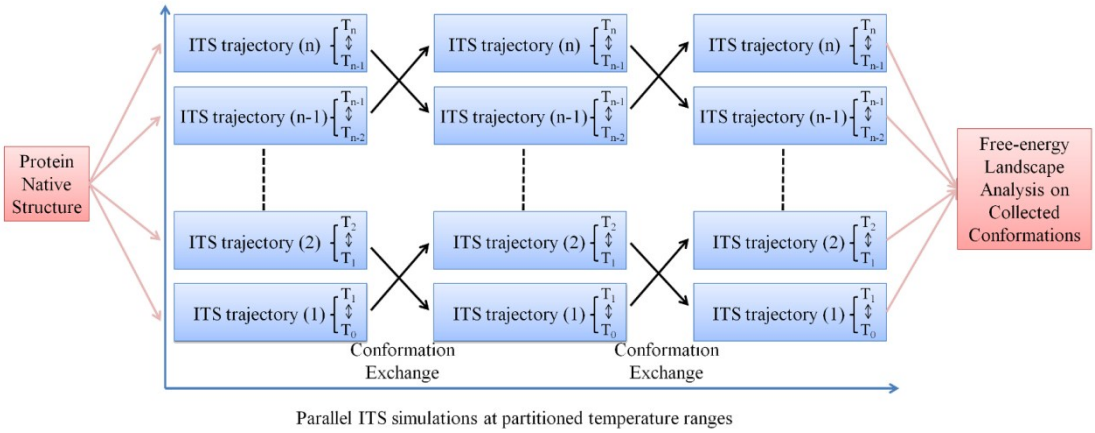

Figure S1. Schematic description of the P-ITS MD simulation for protein folding/unfolding.

Proteins	Replicas	Temp. (K)	$N_{i,k}$'s	Simulation Time (μs)
NTL9	3	280-320	80	4.3
		320-360	80	4.0
		360-430	50	4.2
NuG2b	2	280-370	100	3.6
		370-430	100	3.7
CspA	3	280-320	80	2.6
		320-360	80	2.6
		360-430	50	2.6

Table S2. Simulation parameters for P-ITS MD simulations under study.

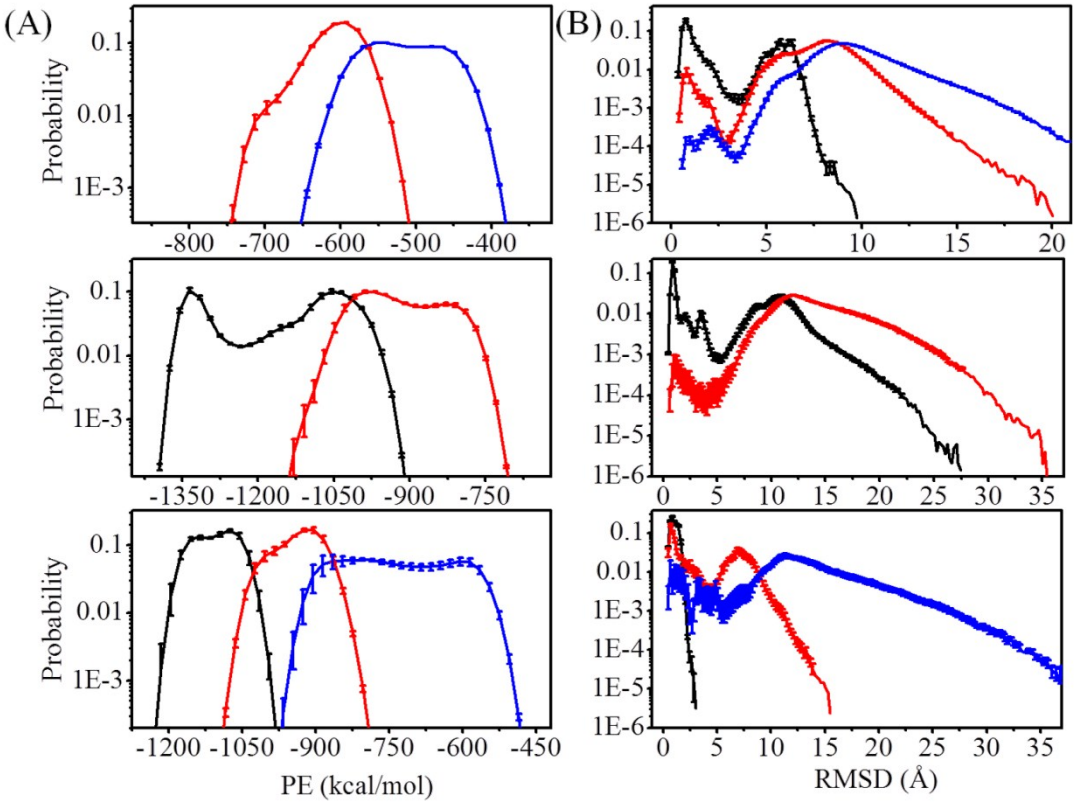

Figure S2. The distribution of potential energies and RMSDs sampled in individual replicas of P-ITS MD simulations for the three proteins under study (top: NTL9, middle: NuG2b, bottom: CspA).

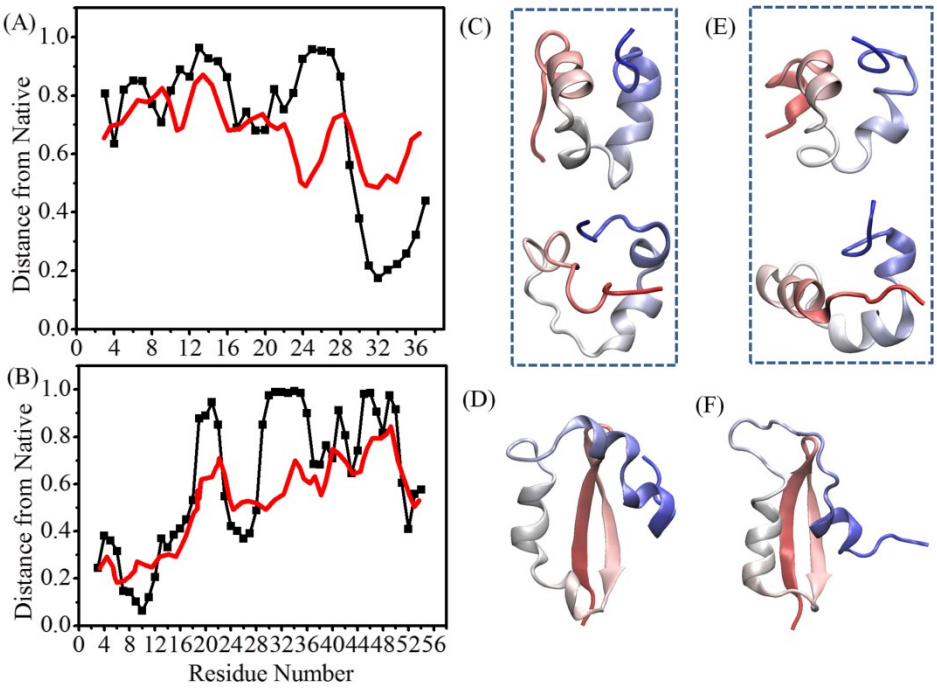

Figure S3. Average distance from the native structure in the unfolded state at 300 K (with a smaller value indicating more native-like for a residue) for (A) NTL9 and (B) NuG2b (Black line: results from the present P-ITS simulations, red line: results from previous extensive MD simulations by Lindorff-Larsen *et al.*). (C-D) Representative structures of the unfolded states of NTL9 and NuG2b from the present P-ITS simulations and (E-F) the comparison to those from the extensive MD simulations by Lindorff-Larsen *et al*.

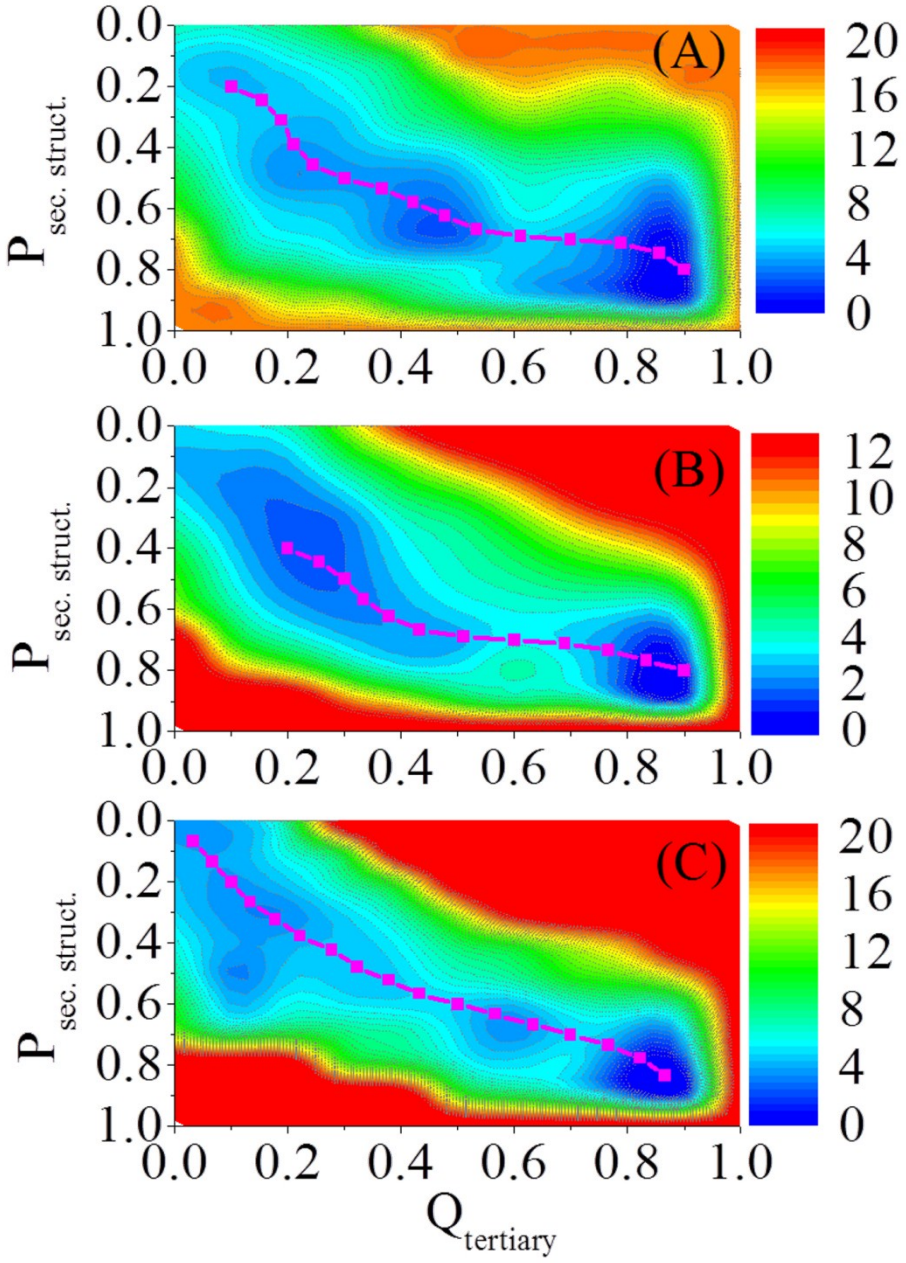

Figure S4. Two-dimensional free-energy landscapes spanning along the fraction of native contacts among the individual hairpins/helices ($Q_{tertiary}$) and the fraction of secondary structure contents ($P_{sec.\ struct.}$) for (A) NTL9, (B) NuG2b, and (C) CspA. The contours are spaced at intervals of 0.5 k_BT.

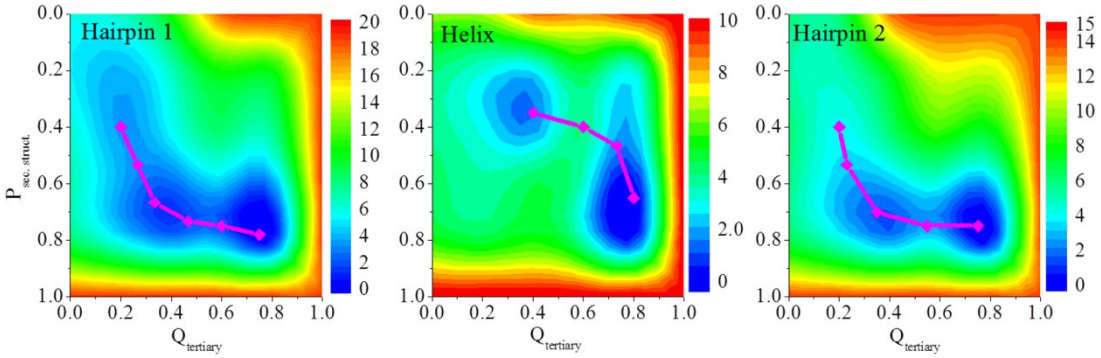

Figure S5. Two-dimensional free-energy landscapes spanning along $Q_{tertiary}$ and $P_{sec.\ struct.}$ for individual hairpins and helix of NTL9. The contours are spaced at intervals of 0.5 k_BT.

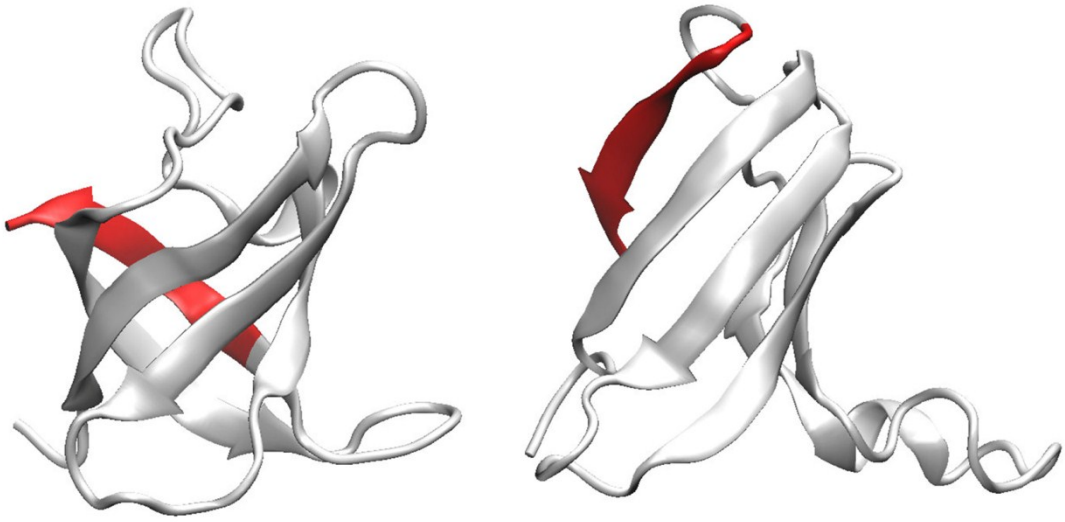

Figure S6. Comparison of the representative conformations of the unfolded (U) and folded (F) states of CspA. The C-terminal β-strand that is misfolded and packed to the first β-strand of hairpin 1 in the unfolded state is colored in red.

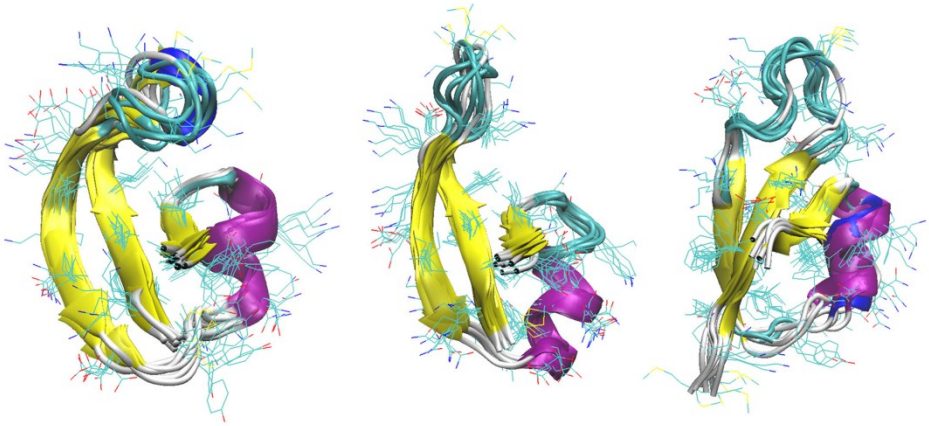

Figure S7. Structural ensembles of the nonnative β-sheet structures of NTL9. Colors code the secondary structural elements as defined in VMD software (purple: α-helix; blue: 3_{10}-helix; yellow: β-sheet; cyan: turn; silver: coiled).

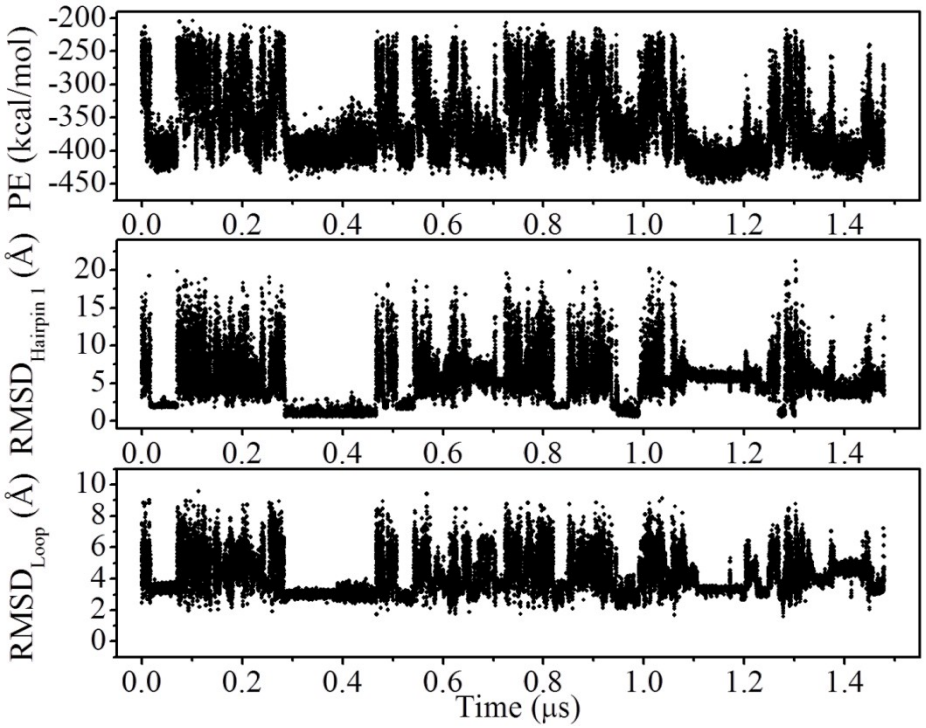

Figure S8. Top: time series of the potential energy (PE) in the trajectory of ITS MD simulations of the isolated structural segment including hairpin 1 and the loop of NTL9. Middle: time series of the RMSD of hairpin 1 and bottom: time series of the RMSD of the loop of NTL9.

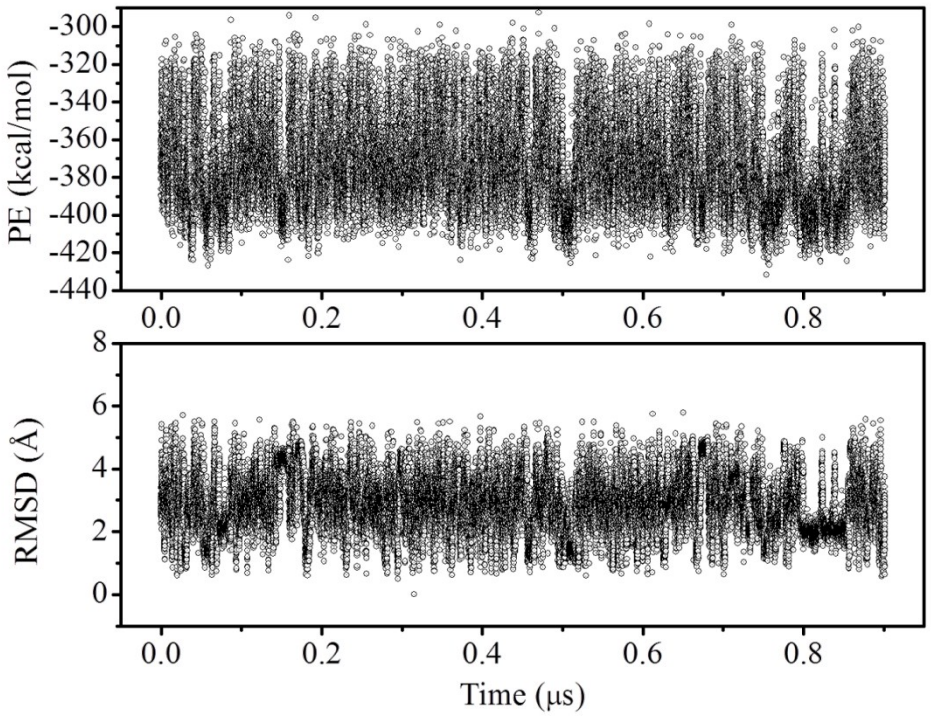

Figure S9. Time series of the potential energy (PE) and the RMSD in the trajectory of ITS MD simulations of the isolated helix of NTL9.