Electronic Supplementary Information (ESI)

Xiakun Chu[1], and Jin Wang[1,2*]

[1] Department of Chemistry
State University of New York at Stony Brook, Stony Brook, NY 11794, USA

[2] State Key Laboratory of Electroanalytical Chemistry
Changchun Institute of Applied Chemistry, Changchun, Jilin 130022, China

[*] jin.wang.1@stonybrook.edu

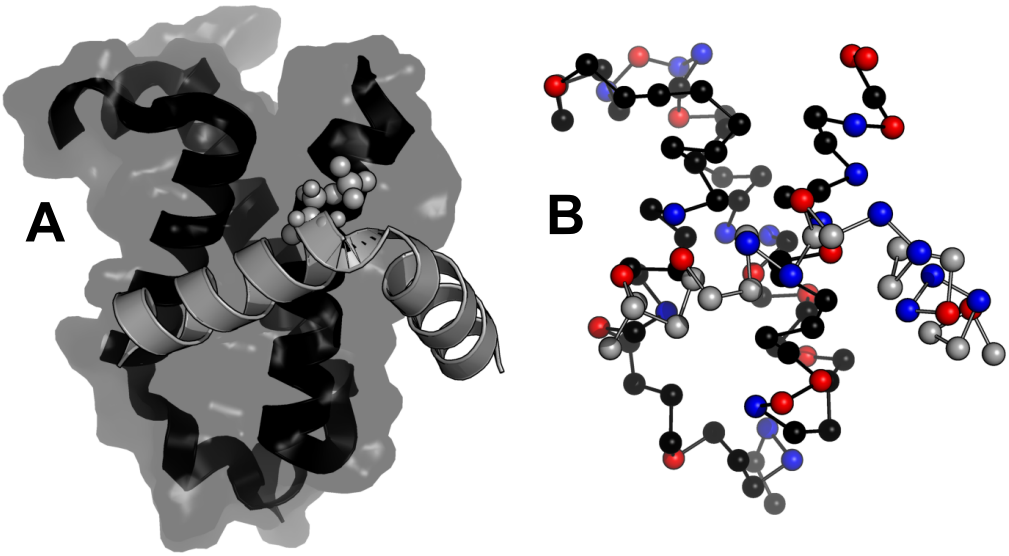

Figure S1: Structural illustrations of bound pKID-KIX complex (PDB:1KDX, first frame). (A) Cartoon representation. KIX and phosphorylated Ser-133 of pKID are shown with transparent surface and ball-and-stick model, respectively. (B) C_α-SBM. The negatively and positively charged residues are colored red and blue, respectively.

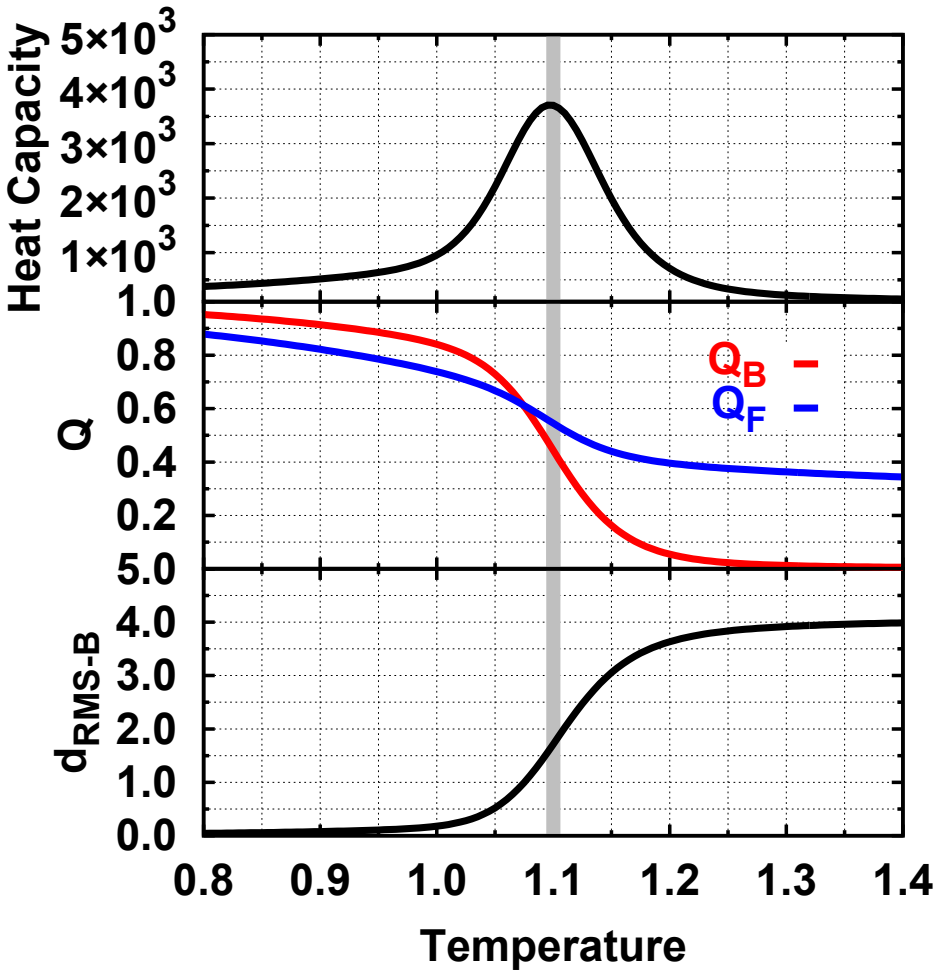

Figure S2: Binding thermodynamics of plain SBM in terms of (top) heat capacity, (middle) melting curves of Q_B and Q_F, and (bottom) melting curve of d_{RMS-B}. The thermodynamics is obtained by WHAM through collections of all the trajectories from REMD simulations. Q_B and Q_F are respectively fractions of native contacts for binding and folding of pKID, d_{RMS-B} is the root-mean-square deviation of distances of native binding contact pairs to those in native PDB structure. The grey lines in each panel indicate the binding temperature, which is derived from the peak of heat capacity curve. d_{RMS-B} is in units of *nm* and temperature is in energy units.

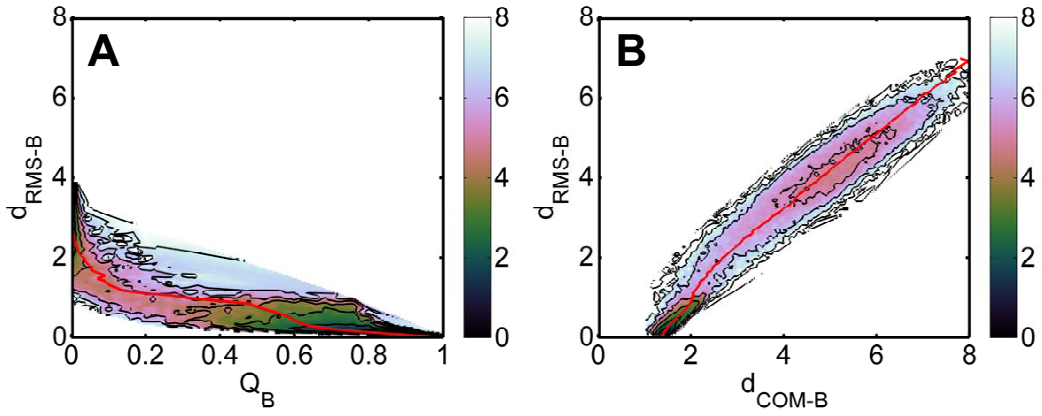

Figure S3: 2D binding free energy landscapes of plain SBM at binding temperature. d_{COM-B} is the distance of center of mass between pKID and KIX. The red lines show the mean values of Q_B and d_{COM-B} along d_{RMS-B}, i. e. $< Q_B(d_{RMS-B}) >$ and $< d_{COM-B}(d_{RMS-B}) >$. From (A), we can see when $d_{RMS-B} > 4.0$, $Q_B = 0$. d_{COM-B} and d_{RMS-B} are in units of *nm* and free energy is in units of kT_B, where T_B is the binding temperature.

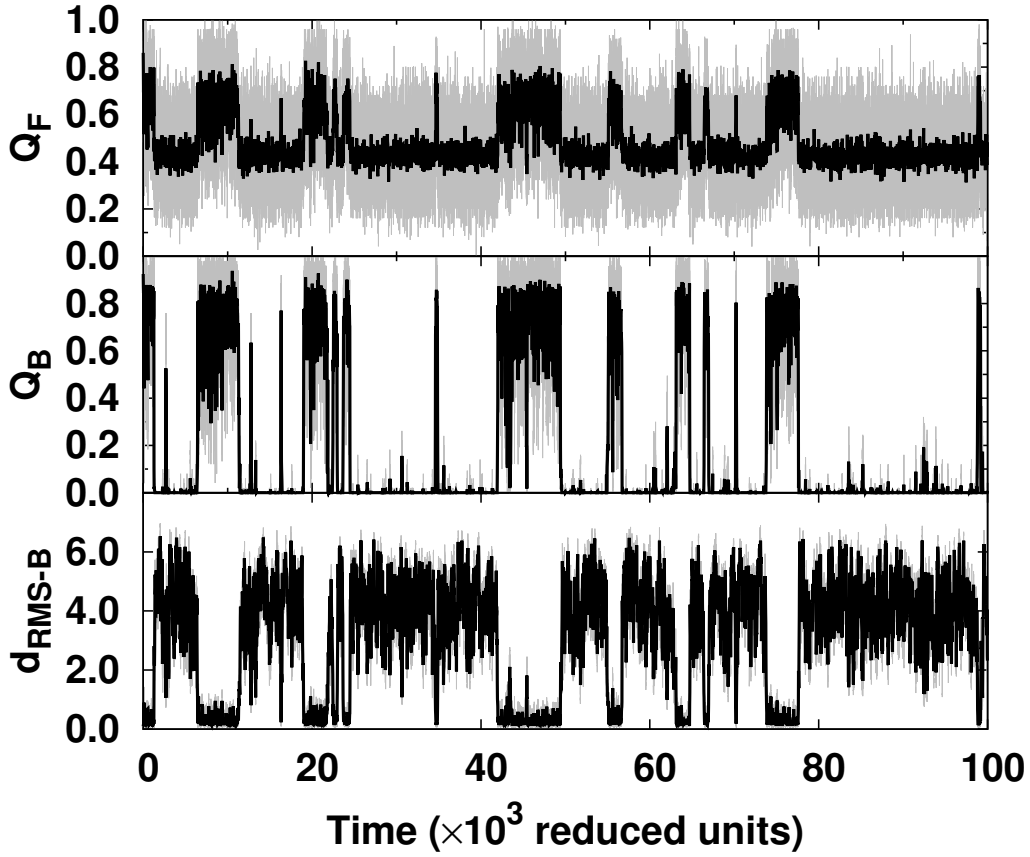

Figure S4: Constant temperature trajectories of plain SBM at binding temperature. Black lines are after block analysis from the raw data, which are shown as grey lines. d_{RMS-B} is in units of *nm*.

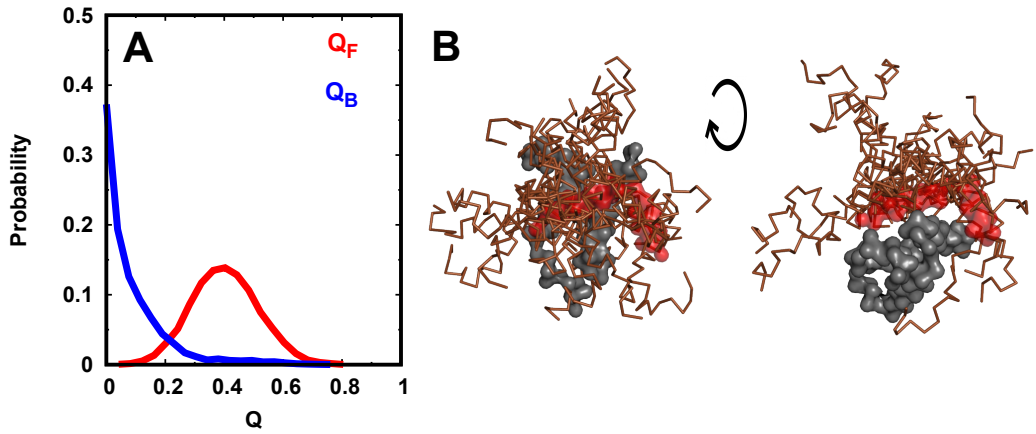

Figure S5: Loosely compact transition states. (A) The probability distributions of fraction of native contacts for folding and binding of pKID (Q_F and Q_B) at transition state ensemble, which is identified by transition path analysis along d_{RMS-B} in Fig. 1 in the main text. (B) Snapshots are illustrated for transition state ensemble. 20 structures of pKID are colored orange, while only one KIX is shown with grey surface and one pKID in PDB structure is shown with transparent red surface for clarity.

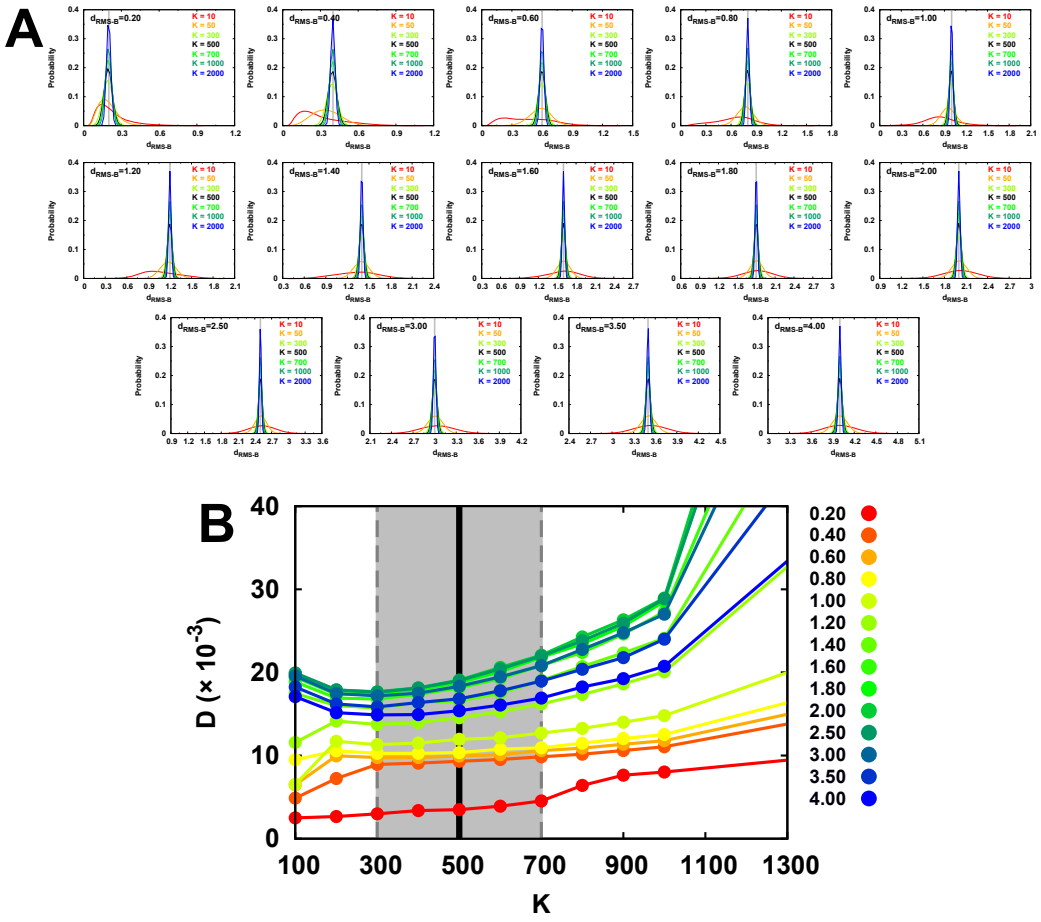

Figure S6: (A) The probability distributions with different harmonic restraining strengths K at different d_{RMS-B}. The grey line in each panel indicates the restraining value of d_{RMS-B}. All distributions show one single peak centered at the user-specified restraining d_{RMS-B} when $K > 300$, fullfilling the quasiharmonic approximation. (B) The diffusion coefficient $D(d_{RMS-B})$ along different harmonic restraining strength K at binding temperature. K is in units of $energy/nm^2$ and $D(d_{RMS-B})$ is in units of $nm^2/time$. The shadow grey region shows that D is robust to the changes in K and the black line with $K = 500$ is therefore used for extracting D.

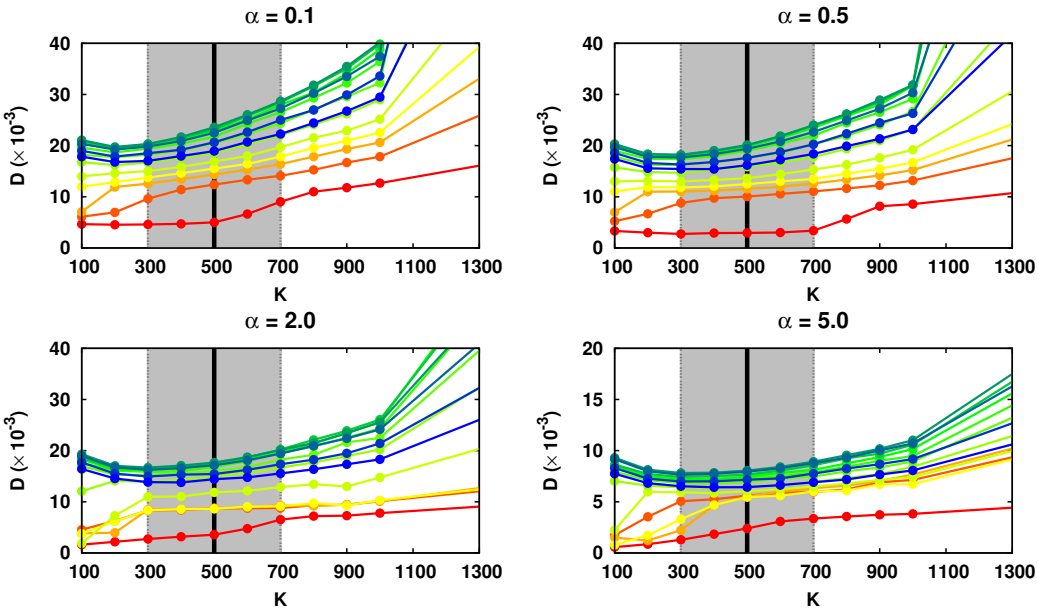

Figure S7: The diffusion coefficient $D(d_{RMS-B})$ along different harmonic restraining strength K for different conformational disorders of pKID with different α at binding temperature when $\alpha = 1.0$. The color schemes are same with that in Figure S6B.

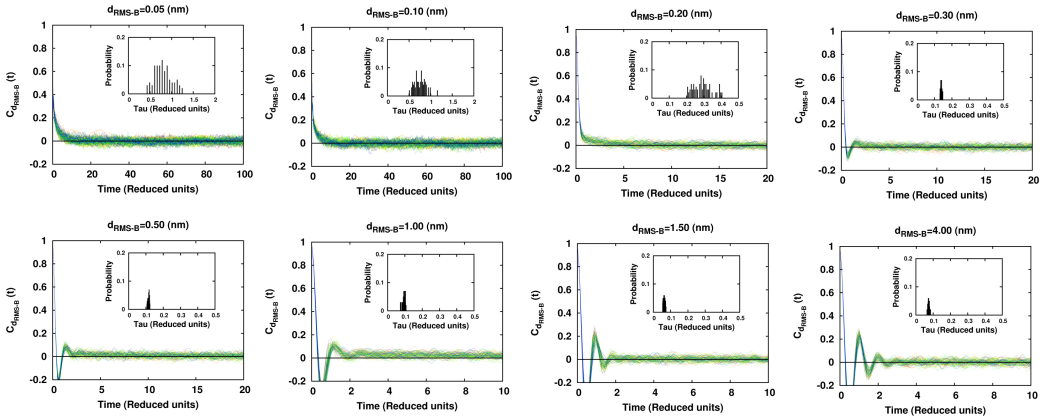

Figure S8: The autocorrelation function $C_{d_{RMS-B}}(t)$ and its relaxtion time $\tau_{corr}(d_{RMS-B})$ for each post-segmented trajectory (2×10^6 MD steps, equal to 1×10^3 long in reduced time units, divided from one 1×10^5 time long trajectory.). All the autocorrelation functions have very small relaxtion times, which are far below the length of the post-segmented trajectory, indicating that the length of the post-segmented trajectory is long enough to ensure a good estimation of relaxtion time τ_{corr}. The trajectory was performed with default parameter α=1.0 and harmonic restraining potential strength K=500.

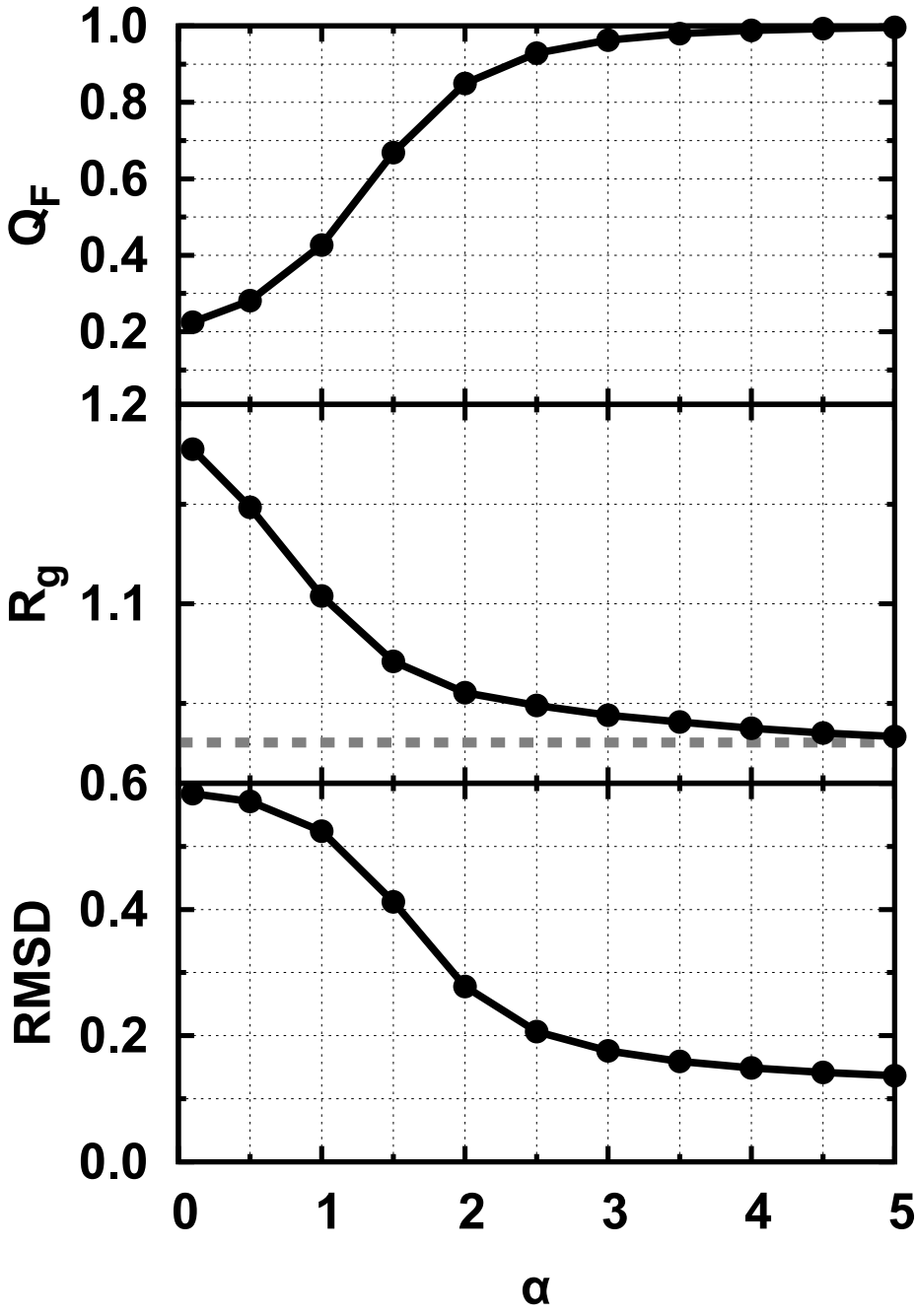

Figure S9: Conformational characteristics of isolated pKID at different strengths of native intra-chain contacts, measured by (top) fraction of native folding contacts Q_F, (middle) radius of gyration R_g and (bottom) root-mean-square deviation (RMSD) to native bound PDB structure. The dashed line in middel panel indicate the value at native bound PDB structure. R_g and RMSD are in units of *nm*.

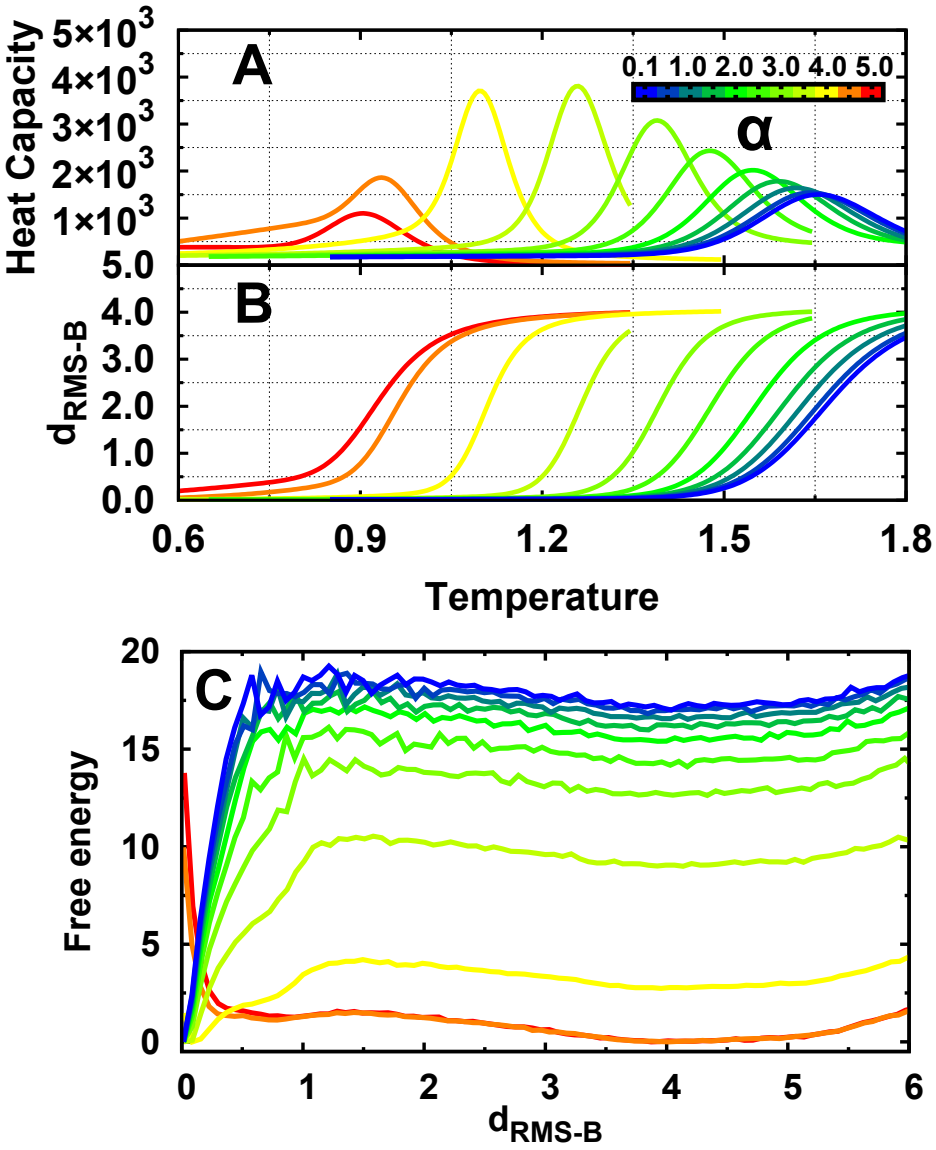

Figure S10: Thermodynamics of binding at different strengths of native intra-chain contacts of pKID. (A) Heat capacity and (B) melting curves of d_{RMS-B} as a function of temperature at different α. (C) The binding free energy projecting onto d_{RMS-B} at binding temperature when $\alpha = 1.0$. Free energy is in units of kT_B, where T_B is the binding temperature at $\alpha = 1.0$.

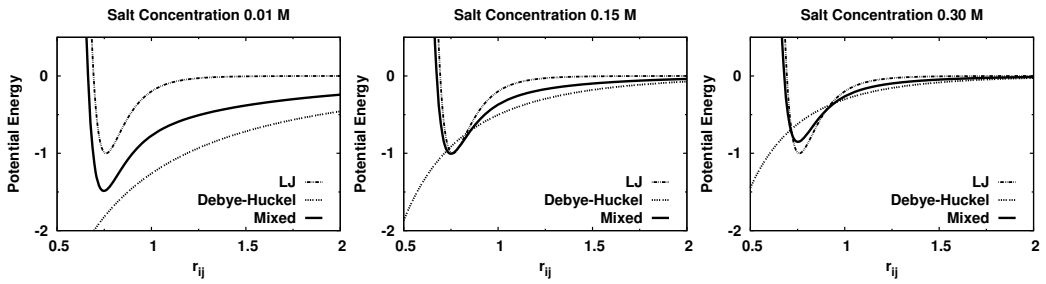

Figure S11: Parameterizations of the native salt bridges. The oppositely charged pairs forming native contacts have both non-local contacting (LJ) and electrostatic (Debye-H*ü*ckel) interactions, thus directly applying the default parameters in SBM would lead to abnormally strong interactions for those pairs (dashed and dotted lines). Therefore, an additional parameterization process is required. At low salt concentrations, which are validated in Debye-H*ü*ckel model typically in the range from 0.01 M to 0.3 M, the energetic contributions for two oppositely charged residues located at the mean distances of the native contacts in this system, which is 0.76 nm, vary from -1.79 to -0.60. The values are in same order of magnitude of the native LJ interaction 1.0, indicating a reasonable balance between LJ and electrostatic interaction with default parameters. Therefore we used the moderate salt concentration at 0.15 M as the reference: a same re-scaling prefactor are both applied to the LJ and electrostatic interactions of native contacting pairs involving two oppositely charged residues to ensure the energetic contribution in native distance is still 1.0 at the salt concentration of 0.15 M (maintaining the biasing to native structure). We then applied the same obtained pre-factor into other salt concentration situations. The new mixed potential (solid line) shows separate LJ and electrostatic interaction behaviors at near and far away native contact distance respectively, which are supposed to capture the "steering" effect of electrostatic interactions on protein binding and preserve the biasing to native structure simultaneously. In the figure, the potentials with the native distance at 0.76 nm and three salt concentration 0.01 M, 0.15 M, 0.3 M are shown as examples. r_{ij} is in units of nm and energy is in reduced units.

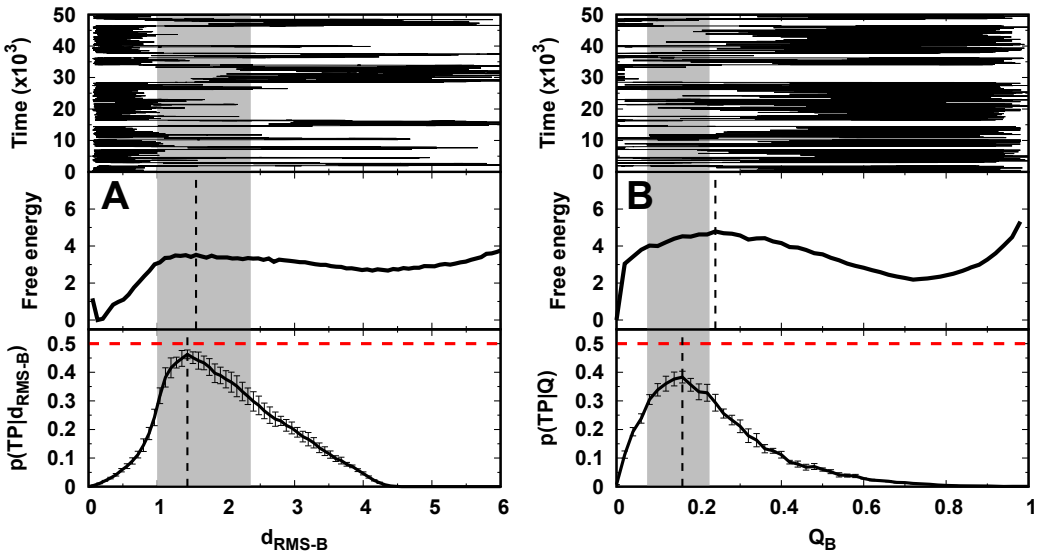

Figure S12: Salt-dependent (C_{salt}=0.15M) trajectories (top), 1D free energy landscapes (middle) and transition path analyses (bottom) at binding temperature along (A) d_{RMS-B} and (B) Q_B. The plot details are similar with Fig. 1 in the main text. The values at the peak of $p(TP|x)$ for d_{RMS-B} and Q_B are 0.46 ± 0.01 and 0.38 ± 0.02, respectively. It implies that d_{RMS-B} is better than Q_B to characterize the binding transition states. The transition state locations identified by the barrier of free energy landscapes and the peak of $p(TP|x)$ has more discrepancy than that is done by plain SBM simulations (Fig. 1 in the main text). In details, for d_{RMS-B}, the transition state d_{RMS-B} are 1.56 and 1.44 from free energy landscape and $p(TP|x)$, respectively; while for Q_B, the transition state Q_B are 0.24 and 0.16 from free energy landscape and $p(TP|x)$ respectively, The free energy landscapes are quantified from REMD simulations, while $p(TP|x)$s are calculated from additional constant temperature simulations, which have accumulated $\sim$ 1000 (un)binding transitions. The error bars are calculated by analyzing different trajectories. Time is in the reduced units, d_{RMS-B} is in units of *nm* and free energy is in units of $kT_B^{0.15M}$, where $T_B^{0.15M}$ is the binding temperature at salt concentration C_{salt}=0.15M.

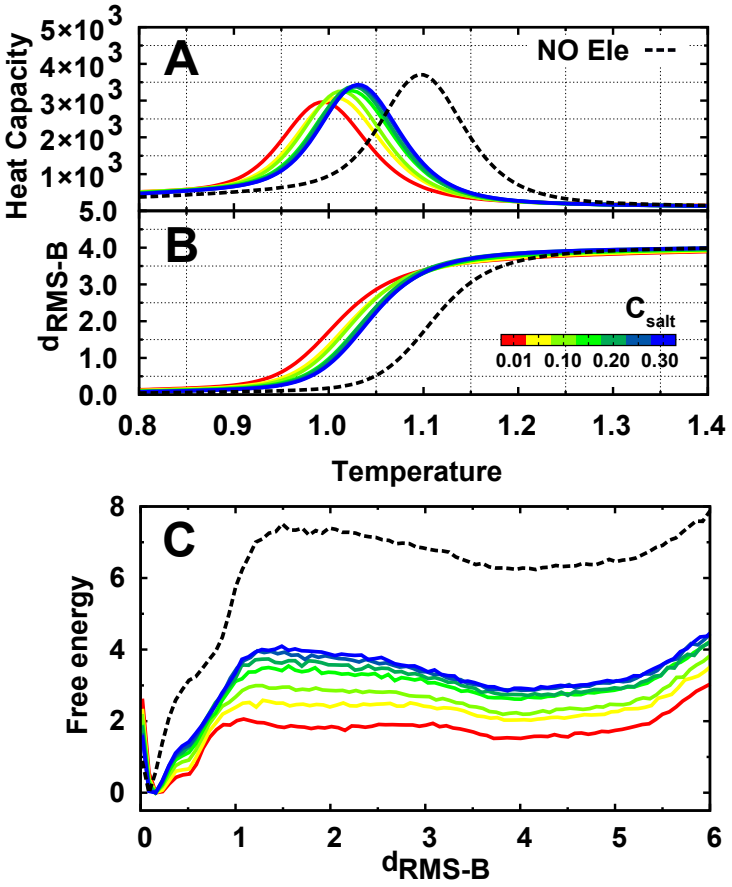

Figure S13: Thermodynamics of binding at different salt concentrations. (A) Heat capacity and (B) melting curves of d_{RMS-B} as a function of temperature at different salt concentrations. (C) The binding free energy projecting onto d_{RMS-B} at binding temperature of $C_{salt} = 0.15M$. Free energy is in units of $kT_B^{0.15M}$, where $T_B^{0.15M}$ is the binding temperature at salt concentration $C_{salt} = 0.15M$.

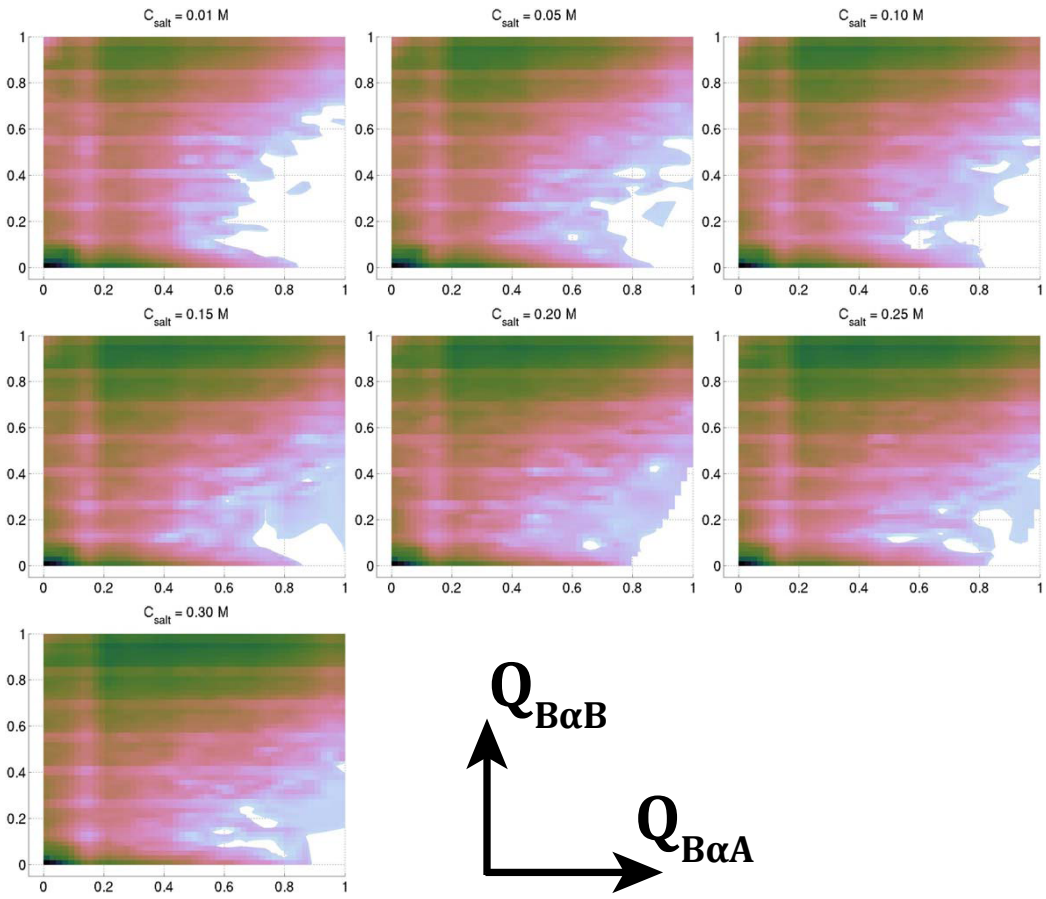

Figure S14: Free energy landscapes of binding projecting onto fraction of native binding contacts between the two helices αA and αB with KIX at different salt concentrations.

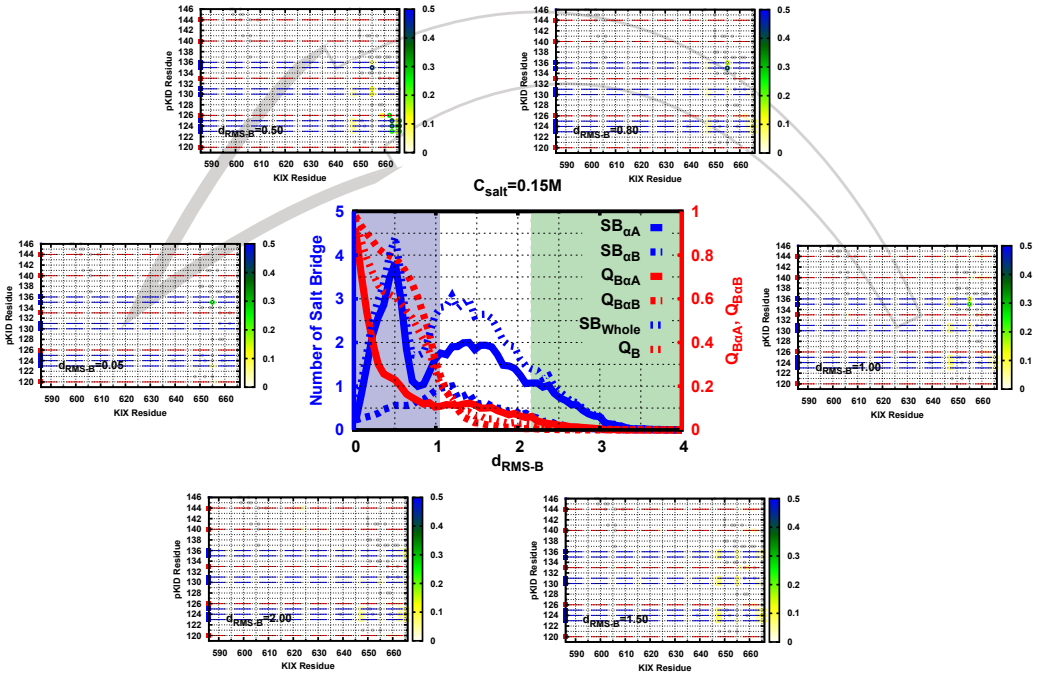

Figure S15: Binding evolution of the pKID and the two helices in pKID at moderate salt concentration ($C_{salt} = 0.15M$). Center panel shows the native contacts, non-native electrostatic contacts (SB) of pKID and the two helices in pKID binding to KIX go along d_{RMS-B}. The non-native electrostatic contact probabilities are plotted surrounding from d_{RMS-B}=2.00 to 0.05. The positive and negative residues in pKID are colored blue and red, respectively. The grey points are native contacts.

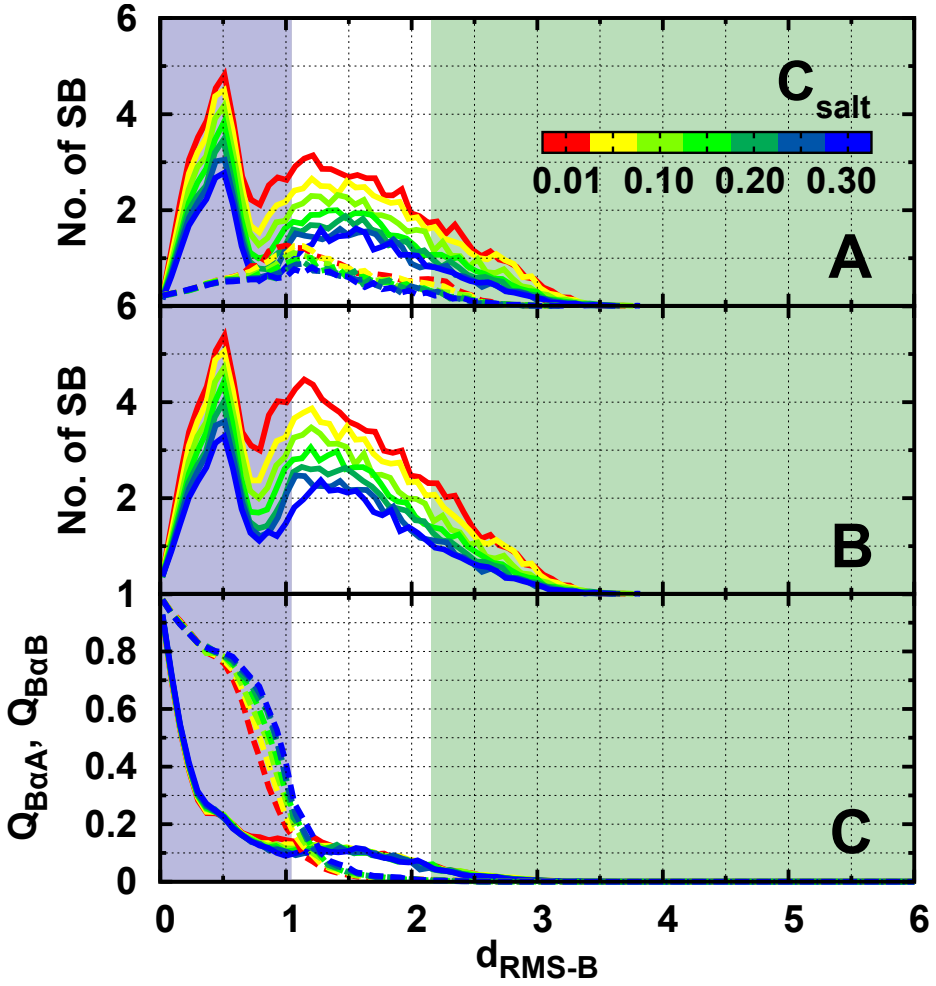

Figure S16: Binding evolution of the two helices in pKID along d_{RMS-B}. Number of non-native electrostatic contacts (SB) are formed by the interface between (A) helix αA, αB and KIX (solid and dashed lines), and (B) the whole pKID and KIX along d_{RMS-B}. (C) the evolution of fraction of native binding contacts of the two helices of pKID with KIX along d_{RMS-B}. The color of shadows and lines are with the same scheme used in Fig. 5 in the main text.

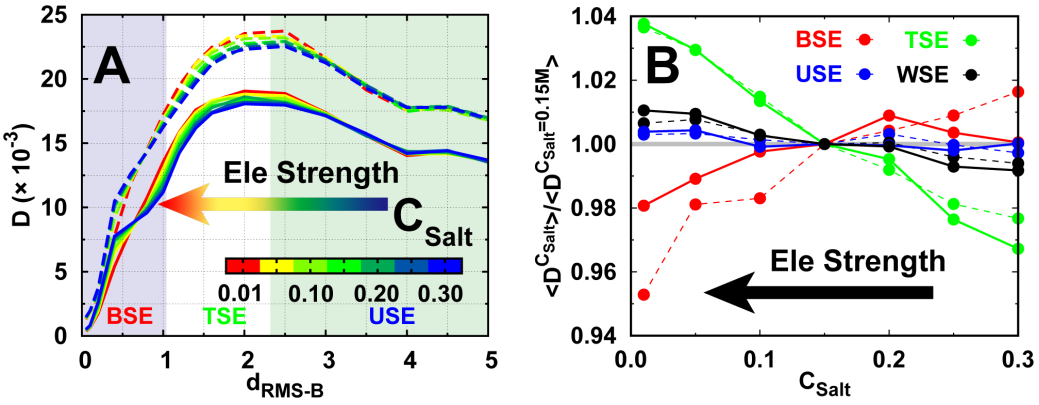

Figure S17: Diffusion coefficients at different strengths of electrostatic interactions. The solid and dashed lines correspond to different degrees of conformational disorder with α=1.0 and 0.1, respectively.

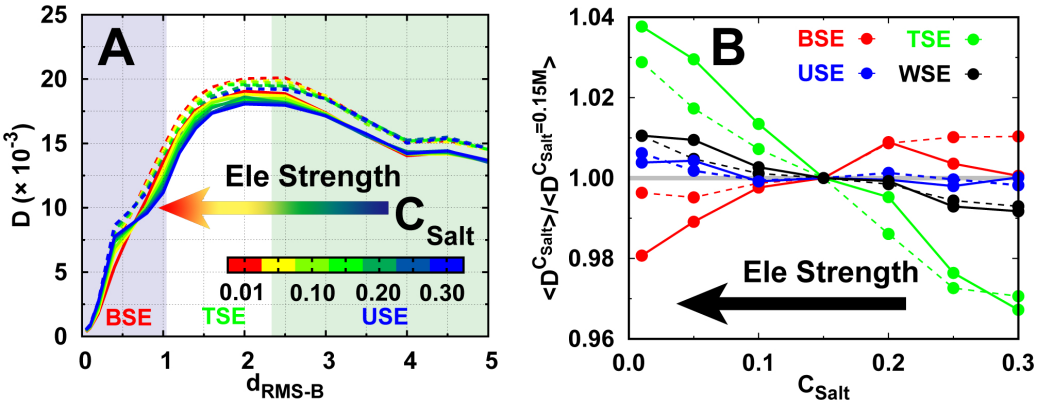

Figure S18: Diffusion coefficients at different strengths of electrostatic interactions. The solid and dashed lines respectively correspond to different temperatures at 1.02, which is binding temperature at salt concentration C_{Salt}=0.15M, and 1.10, which is the binding temperature without electrostatic interactions.