Nonnative Interactions in Coupled Folding and Binding Processes of Intrinsically Disordered Proteins

Yongqi Huang[1,2,3], Zhirong Liu[1,2,3]*

1 State Key Laboratory for Structural Chemistry of Unstable and Stable Species, College of Chemistry and Molecular Engineering, Peking University, Beijing, China, **2** Center for Theoretical Biology, Peking University, Beijing, China, **3** Beijing National Laboratory for Molecular Sciences, Peking University, Beijing, China

Abstract

Proteins function by interacting with other molecules, where both native and nonnative interactions play important roles. Native interactions contribute to the stability and specificity of a complex, whereas nonnative interactions mainly perturb the binding kinetics. For intrinsically disordered proteins (IDPs), which do not adopt rigid structures when being free in solution, the role of nonnative interactions may be more prominent in binding processes due to their high flexibilities. In this work, we investigated the effect of nonnative hydrophobic interactions on the coupled folding and binding processes of IDPs and its interplay with chain flexibility by conducting molecular dynamics simulations. Our results showed that the free-energy profiles became rugged, and intermediate states occurred when nonnative hydrophobic interactions were introduced. The binding rate was initially accelerated and subsequently dramatically decreased as the strength of the nonnative hydrophobic interactions increased. Both thermodynamic and kinetic analysis showed that disordered systems were more readily affected by nonnative interactions than ordered systems. Furthermore, it was demonstrated that the kinetic advantage of IDPs ("fly-casting" mechanism) was enhanced by nonnative hydrophobic interactions. The relationship between chain flexibility and protein aggregation is also discussed.

Introduction

Deciphering how physical interactions affect protein behavior is fundamental to structural and functional biology. As a first approximation, interactions presented in the native state (native interactions) dominate in processes such as protein folding and binding, resulting in a funnel-like energy landscape with minimal frustration [1–3]. Under such conditions, the Gō-model [4,5] has been widely adopted to generate valuable insights into protein folding and binding [6–10]. In realistic systems, however, the existence of nonnative interactions is inevitable. The effects of nonnative interactions on protein folding have been demonstrated in many experimental studies [11–23]. Nonnative interactions perturb the denatured state ensemble and thus affect the equilibrium stability [11,12], lead to the accumulation of the on-pathway or off-pathway intermediate states [13–15], and most importantly, moderate the protein folding kinetics by perturbing the transition state [16,17,20,21]. Nonnative interactions also act as a major driving force in the rapid collapse of an unfolded protein during the early stages of the folding process, which is important in preventing proteins from aggregating [22]. In the phosphorylation-activation process of a signal protein (nitrogen regulatory protein C), the disruption of some native contacts was compensated by the transient formation of nonnative interactions [23]. For protein binding processes, nonnative interactions have been recognized to be important in the initial formation of the non-specific encounter complexes, where long-range electrostatic

interactions increase the diffusion process by the "steering effect", and then short-range hydrophobic interactions facilitate the formation of the final specific complexes by a two-dimensional search on the surface [24–27].

The effects of nonnative interactions on protein folding have been extensively studied by simulations and analytical theories [28–45]. An all-atom simulation suggested that ~20–25% of the energy in the transition state arose from nonnative contacts [28]. Nonnative hydrogen bonds in a simulation on the helix-coil transition were found to be most populated around the transition temperature [29]. Mutation can also change the population of nonnative contacts [30]. In general, the existence of nonnative interactions may influence protein stability and folding kinetics [31]. A lattice model simulation showed that nonnative interactions have little effect on protein stability, but would accelerate protein folding and thus give rise to the Φ-values that are negative or larger than unity [32]. Further detailed analytical and simulation studies found that the folding rate generally enhances initially as the nonnative interactions increase, but drops rapidly when the nonnative interactions are larger than a critical value [33–36]. The importance of nonnative interactions on folding kinetics was also validated by a direct comparison between simulation and experiment for the SH3 protein [37]. In the novel designed protein Top7 [46], it was revealed that the noncooperative folding kinetics is caused by both native topology and nonnative interactions [38,39]. For proteins with more complicated topologies, e.g., knotted proteins, it was suggested that

nonnative interactions play an essential role in the correct formation of the knots [40]. Despite the perturbation on folding kinetics, the protein folding mechanism is usually robust with respect to nonnative interactions [43,44]. Compared with the extensive studies on protein folding, theoretical investigations on protein binding are relatively rare [47–51]. The mechanisms of the electrostatic rate enhancement via lowering the transition state energy and the dimensionality-reducing effect by non-specific binding to DNA and cell membranes are well understood; however, the effects of short-range hydrophobic interactions on general protein–protein binding processes remain unresolved.

The influences of nonnative interactions may be more prominent in intrinsically disordered proteins (IDPs). IDPs are a special family of proteins that lack unique tertiary structures under physiological conditions, either along the entire chain or in particular regions [52–55]. IDPs were predicted to be enriched in both prokaryotic and eukaryotic genomes [56,57] and perform various functions, including transcription and translation regulation, cellular signal transduction, protein modifications, and molecular assemblies. In particular, IDPs have been shown to be associated with human diseases such as cancer, cardiovascular disease, amyloidosis and neurodegenerative diseases [58]. Although disordered when alone in solution, in many cases, IDPs undergo conformational transitions to folded states upon binding their biological targets to perform functions [59]. Gō-like models have been successfully applied to study the coupled folding and binding processes of IDPs [9,49,60–63]. This approach has contributed important insights into the characteristics of IDPs, e.g., the kinetic advantages in molecular recognition for IDPs through the "fly-casting" mechanism [64]. Considering the significant chain flexibility of IDPs, IDPs are expected to possess more nonnative interactions in the folding and binding processes than conventional ordered proteins. It would be important to investigate the different effects of nonnative interactions on IDPs and ordered proteins and whether taking into account the nonnative interactions would change the principles of IDPs elucidated by native-centric models.

In this paper, we conducted computer simulations to study how nonnative hydrophobic interactions affect the binding thermodynamics and kinetics of IDPs, and how these effects are related to the chain flexibility which distinguishes IDPs from the ordered proteins.

Results

IDPs are readily trapped into non-specific states

To investigate the effects of nonnative hydrophobic interactions on the coupled folding and binding of IDPs, we modified a coarse-grained Gō-like model of IDPs [62] to include a sequence-dependent hydrophobic-polar (HP) component which accounts for the nonnative hydrophobic interactions [37] (see *Materials and Methods*). We used our model to simulate the binding of the phosphorylated kinase-inducible domain (pKID) of the transcription factor cAMP response-element binding protein to the kinase inducible domain interacting domain (KIX) of the cAMP response-element binding protein. The pKID domain is a well characterized IDPs which folds upon binding to KIX [26,65,66]. In the model, a parameter α was introduced to scale the strength of the intra-molecular native interactions and thus tune the chain flexibility of pKID [62], while another parameter K_{HP} was used to describe the strength of the nonnative hydrophobic interactions.

The influences of nonnative hydrophobic interactions on the binding free energy of the pKID-KIX complex and their interplay with the chain flexibility are summarized in Figure 1. Here, we

used the fraction of native contacts between the two proteins (Q_b) as a reaction coordinate to depict the binding process. The nonnative hydrophobic interactions were found to mainly stabilize the partially bound states with moderate Q_b values, but had little influence on either the unbound or bound states. Over-strong nonnative hydrophobic interactions can even trap an intermediate state centered at $Q_b \sim 0.2$ (Figures 1A,B). Interestingly, our results clearly showed that the effect of nonnative hydrophobic interactions on a disordered system (with a small $\langle Q_f^{(free)} \rangle$) was more remarkable than that on an ordered system (with a large $\langle Q_f^{(free)} \rangle$). To quantitatively measure the influence on the equilibrium properties, we divided the conformation space into three states: the unbound state (U), intermediate state (I), and bound state (B) (Figure 1B). The population of the non-specific intermediate state increased as the strength of the nonnative hydrophobic interactions K_{HP} was increased, and the disordered system exhibited a more remarkable increase (Figure 2A). Although the nonnative hydrophobic interactions frustrated the binding free-energy landscapes, its influence on the stability of the complex was rather small because the free-energy difference between the bound and unbound states showed negligible change with K_{HP} (Figure 2B). When analyzing the transition temperature (T_m, defined as the temperature corresponding to the peak in the heat capacity curve [62]), the same conclusion was reached: T_m remained constant with increasing K_{HP} (Figure 2C).

Flexibility promotes dynamic and extensive non-specific interactions

To further characterize the nonnative hydrophobic interactions along the binding process, we examined the number of nonnative contacts (N_{HP}) of pKID in the free state, binding intermediate state, and the bound state. In the free state, the radius of gyration (R_g) showed that pKID underwent a minor compression in the presence of nonnative hydrophobic interactions (Figure 3A). Compared with the total number of intramolecular native contacts in the bound state (i.e., 25), the number of nonnative contacts was rather small (Figure 3B). For the most disordered system, $\langle N_{HP}^{(free)} \rangle$ was only ~ 1.0 even when $K_{HP} = 1.5$, and as expected, the ordered system showed no nonnative contacts. The small number of nonnative contacts observed for pKID in the free state explains the insensitivity of the equilibrium binding free-energy with respect to the strength of nonnative hydrophobic interactions (Figure 2B) because, by definition, the native bound state was affected only by native interactions. This observation was also consistent with results on protein folding which showed that the number of nonnative contacts of globular proteins is usually significantly smaller than the number of native contacts in the folded state [29,39]. Therefore, intramolecular native contacts were not affected in the free state (Figure 3C). In contrast, in the binding process the intermediate state showed a considerable number of nonnative contacts, particularly for the disordered system (Figures 4A–C). Under the same strength of the nonnative hydrophobic interactions, the nonnative contact number of the intermediate state showed a remarkable decrease with decreasing chain flexibility (Figure 4D). Compared with the intermediate state, the bound state possessed fewer nonnative hydrophobic contacts (Figure 4E). This feature strongly indicates that chain flexibility promotes dynamic and extensive non-specific interactions during the binding process and will have a significant effect on the binding kinetics. When correlating the nonnative contact number of the intermediate state to the strength of the nonnative interactions, a sharp increase in the nonnative contact number appeared (Figure 4F), which leads to misbinding states (Figure 2A).
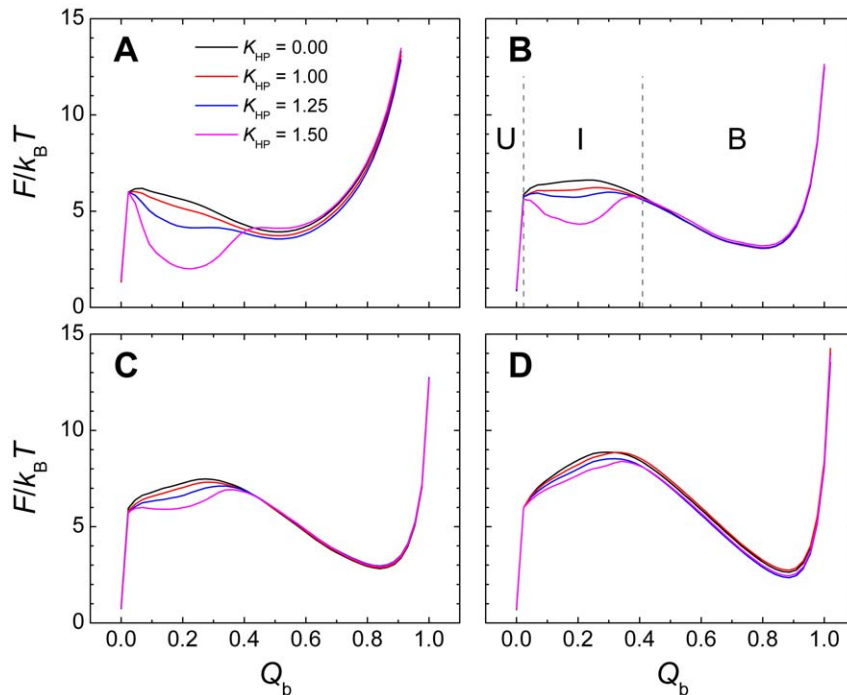
Figure 1. Free-energy profiles of the binding process. Free-energy profiles were calculated using the fraction of native intermolecular contacts (Q_b) as a reaction coordinate for systems with different degrees of chain flexibility. (A–D) $\langle Q_f^{(\text{free})} \rangle = 0.29$, 0.46, 0.65, and 0.85 by tuning the intramolecular interaction parameter α from 0.1, 1.0, 1.5 to 3.0. The two vertical dashed lines in panel (B) indicate the definition of the unbound state (U), intermediate state (I), and bound state (B). The strength of the nonnative hydrophobic interactions (K_{HP}) ranges from 0.00 to 1.50.
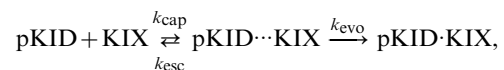doi:10.1371/journal.pone.0015375.g001

Binding rate is accelerated by nonnative hydrophobic interactions

It was established from the Gō-models that IDPs possess faster binding rates than ordered proteins via the "fly-casting" mechanism which would facilitate molecular recognition [60,62,64]. Considering the findings in protein folding studies that nonnative interactions can accelerate or decelerate the folding rates [32–37,39] and the above observation that nonnative contacts are more prevalent in IDPs than in ordered proteins, it would be intriguing to explore whether the kinetic advantage of IDPs will be smeared by the nonnative interactions. We simulated the binding kinetics of the pKID-KIX complex under various strengths of nonnative interactions, and the results are summarized in Figure 5. In a similar manner to protein folding, nonnative interactions have a nonlinear effect on the binding kinetics, i.e., the binding rate initially increased and then slowed sharply as the strength of the nonnative hydrophobic interactions increased, and this observation was independent of chain flexibility (Figures 5A,B). The existence of the critical strength of nonnative interactions was also observed in the work of Turjanski et al. [60]. In the binding-rate increase region, the disordered system possessed a greater binding rate than the ordered system, showing a kinetic advantage in the binding process even in the presence of nonnative hydrophobic interactions. As revealed from the free-energy analysis that the disordered system was more readily affected by nonnative hydrophobic interactions (Figure 1), the binding rate of the disordered system showed greater amplitude of change (Figure 5B). For the disordered system, the largest binding rate was 1.46 times as large as that when nonnative hydrophobic interactions were not included ($K_{\text{HP}} = 0.0$). However, for the ordered system, the largest binding rate was only 1.08 times as large as that at $K_{\text{HP}} = 0.0$. This indicates that the nonnative

hydrophobic interactions further amplify the kinetic advantages of IDPs in the binding process. The most striking finding was that the strength of the nonnative hydrophobic interactions corresponding to the maximum binding rate ($K_{\text{HP}}^{\text{max-rate}}$) showed a strong dependence on the chain flexibility (Figures 5A, C). The disordered system exhibited a smaller $K_{\text{HP}}^{\text{max-rate}}$ than the ordered system. The reduction in the binding rate under strong nonnative hydrophobic interactions was caused by non-specific kinetic traps along the binding trajectory. Figure 5D exemplifies a binding trajectory of a system with $\langle Q_f^{(\text{free})} \rangle = 0.46$ under $K_{\text{HP}} = 1.50$. Non-specific intermediate states are clearly shown. The chain flexibility dependence of $K_{\text{HP}}^{\text{max-rate}}$ indicates that, during the binding process, IDPs are more ready to form non-specific binding intermediates and even kinetically trapped misbinding states.

The encounter complex is stabilized

The kinetic advantage of IDPs originates from the chain flexibility facilitating the encounter complex to evolve into the final binding complex rather than escape to the unbound state [62]. To investigate the influence of nonnative hydrophobic interactions on such a mechanism, we made an analysis by dissecting the binding process into a capture process and a further evolution process [62]:

$$\text{pKID} + \text{KIX} \underset{k_{\text{esc}}}{\overset{k_{\text{cap}}}{\rightleftarrows}} \text{pKID}\cdots\text{KIX} \overset{k_{\text{evo}}}{\longrightarrow} \text{pKID}\cdot\text{KIX},$$

where pKID+KIX is the unbound state, pKID·KIX is the native bound state, while pKID···KIX is the loosely bound encounter complex state formed by the capture event. We considered an encounter complex occurred when the system evolved from an unbound state to a state with $Q_b > 0$ (usually had one intermolecular native contact). The effect of nonnative hydrophobic
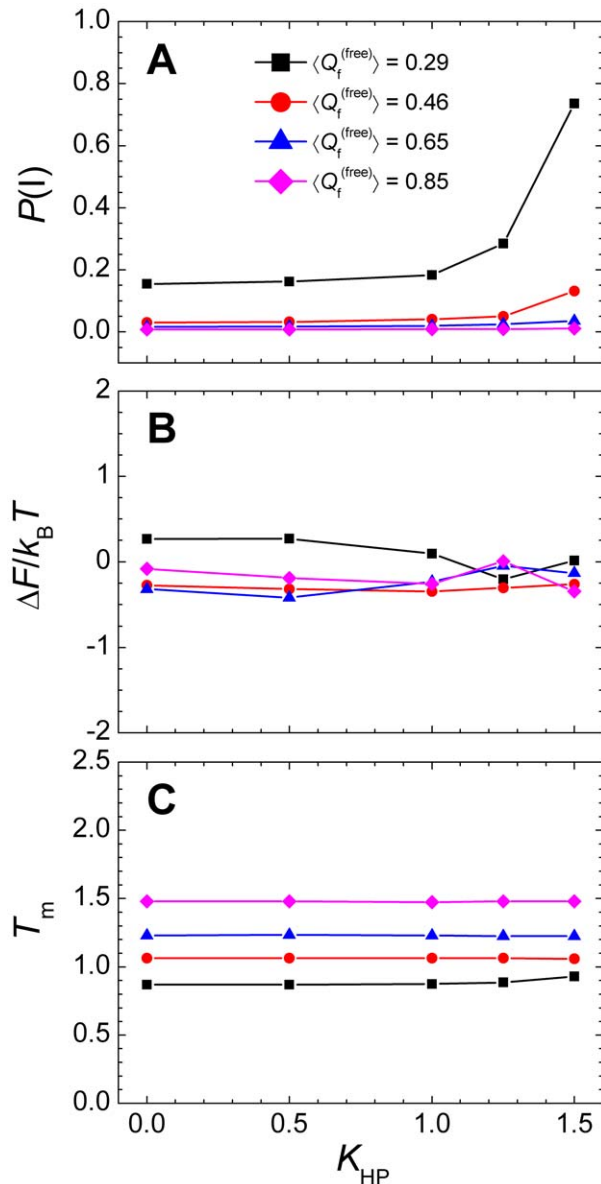
Figure 2. Thermodynamic properties of systems with different degrees of chain flexibility. (A) Correlation between the population of the intermediate state $P(I)$ and the strength of the nonnative hydrophobic interactions, K_{HP}. (B) Effect of the nonnative hydrophobic interactions on the stability of the complex which is measured by the free-energy difference between the bound state and the unbound state. (C) Effect of the nonnative hydrophobic interactions on the transition temperature T_m.
doi:10.1371/journal.pone.0015375.g002

Figure 3. Properties of the free pKID domain in the presence of nonnative hydrophobic interactions. (A) radius of gyration (R_g), (B) average number of nonnative contacts ($\langle N_{HP}^{(free)} \rangle$), and (C) the average fraction of native contacts ($\langle Q_f^{(free)} \rangle$).
doi:10.1371/journal.pone.0015375.g003

interactions on the capture rate k_{cap} was rather small for all systems and the disordered system possessed a slower capture rate (Figure 6A). Unlike the capture process, the evolving and the escape rates from the encounter state showed significant responses to the presence of the nonnative hydrophobic interactions (Figures 6B,C). Both Figures 6B and 6C show a two-stage profile, i.e., an initial plateau stage followed by a sharp decrease stage. The behaviors of the evolving and the escape rates as a function of the nonnative hydrophobic interaction strength were synchronic, but with opposite amplitude, namely, the disordered system ($\langle Q_f^{(free)} \rangle = 0.29$) showed the greatest decrease of the evolving
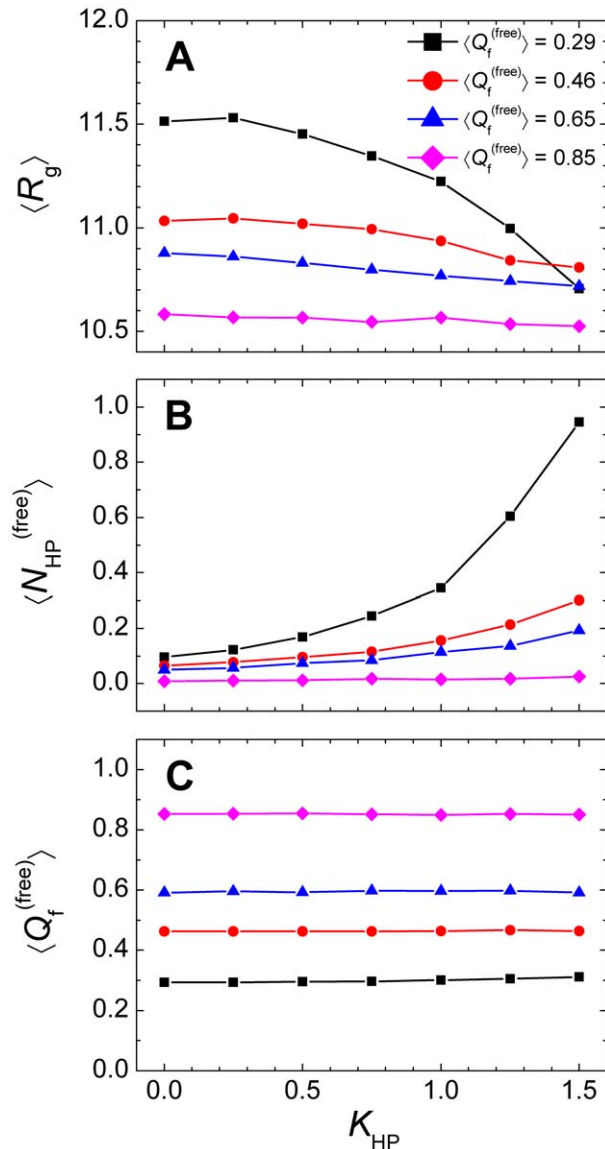
rate and the smallest decrease of the escape rate, whereas the ordered system ($\langle Q_f^{(free)} \rangle = 0.85$) showed the smallest decrease of the evolving rate and the greatest decrease of the escape rate. Compared with Figure 5C, we also noticed that the points corresponding to the sharp decreases of the evolving and the escape rates were smaller than those corresponding to the sharp decreases of the overall binding rate.

The above results showed that the effect of the nonnative hydrophobic interactions on the binding process was primarily exerted on the evolution and escape stages. Unlike electrostatic interactions, which are long-range and accelerate the binding rate by the steering effect [24], nonnative hydrophobic interactions are short-range and their effects on the capture process is negligible (Figure 6A). However, in the encounter state, nonnative hydrophobic interactions contribute energetically to the stability of the encounter complex and so a reduction in the evolving and
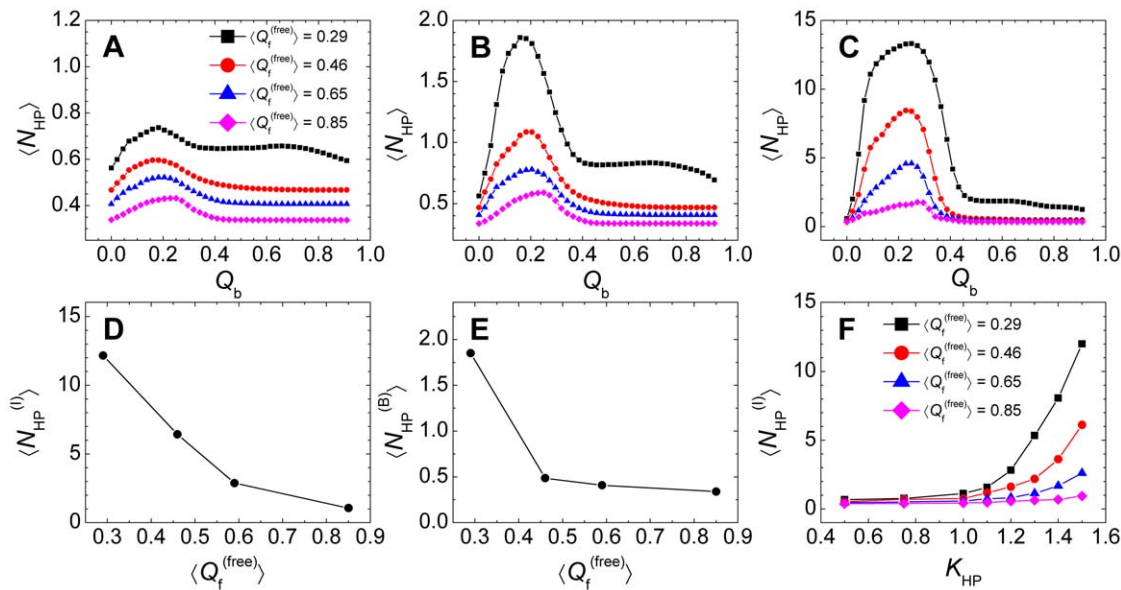
Figure 4. Characterization of the number of nonnative contacts. (A–C) The average number of nonnative contacts $\langle N_{HP} \rangle$ along the binding process when the strength of nonnative hydrophobic interactions was increased: (A–C) $K_{HP} = 0.50$, 1.00, and 1.50. (D) Correlation between the average number of nonnative contacts at the intermediate state $\langle N_{HP}^{(I)} \rangle$ and $\langle Q_f^{(free)} \rangle$. K_{HP} was set 1.50. (E) Correlation between the average number of nonnative contacts in the bound state $\langle N_{HP}^{(B)} \rangle$ and $\langle Q_f^{(free)} \rangle$. K_{HP} was set 1.50. (F) Correlation between $\langle N_{HP}^{(I)} \rangle$ and K_{HP}. The definitions of the intermediate state and bound state are presented in Figure 1.
doi:10.1371/journal.pone.0015375.g004

escape rates were observed (Figures 6B,C). Evolution and escape are two opposite processes in a binding event. Although increasing the nonnative hydrophobic interactions reduces the escape rate, it also reduces the evolving rate. Therefore, there is a balance between the escape and evolving rates to lead to the maximum binding rate.
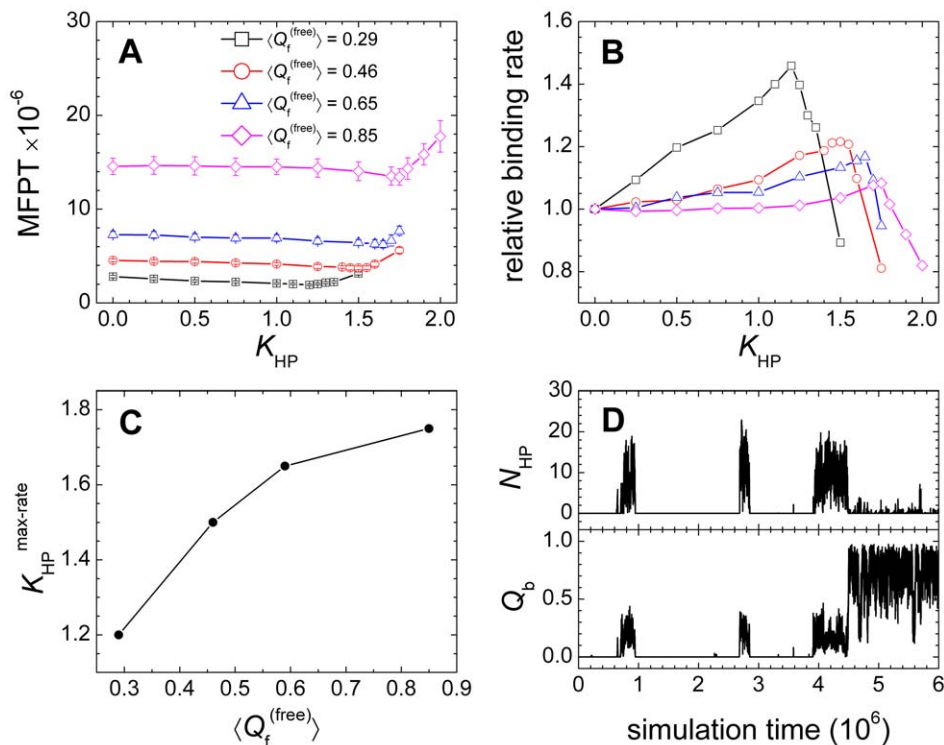

Figure 5. Effect of the nonnative hydrophobic interactions on the binding kinetics. (A) Mean first passage time (MFPT) of the binding process as a function of K_{HP}. (B) Correlation between the relative binding rate and K_{HP}. The relative binding rate was computed as $\mathrm{MFPT}(K_{HP} = 0.0)/\mathrm{MFPT}(K_{HP})$. (C) Correlation between the K_{HP} corresponding to the maximum binding rate, $K_{HP}^{max\text{-}rate}$, and $\langle Q_f^{(free)} \rangle$. (D) A typical binding trajectory for the system with $\langle Q_f^{(free)} \rangle = 0.46$ under $K_{HP} = 1.50$.
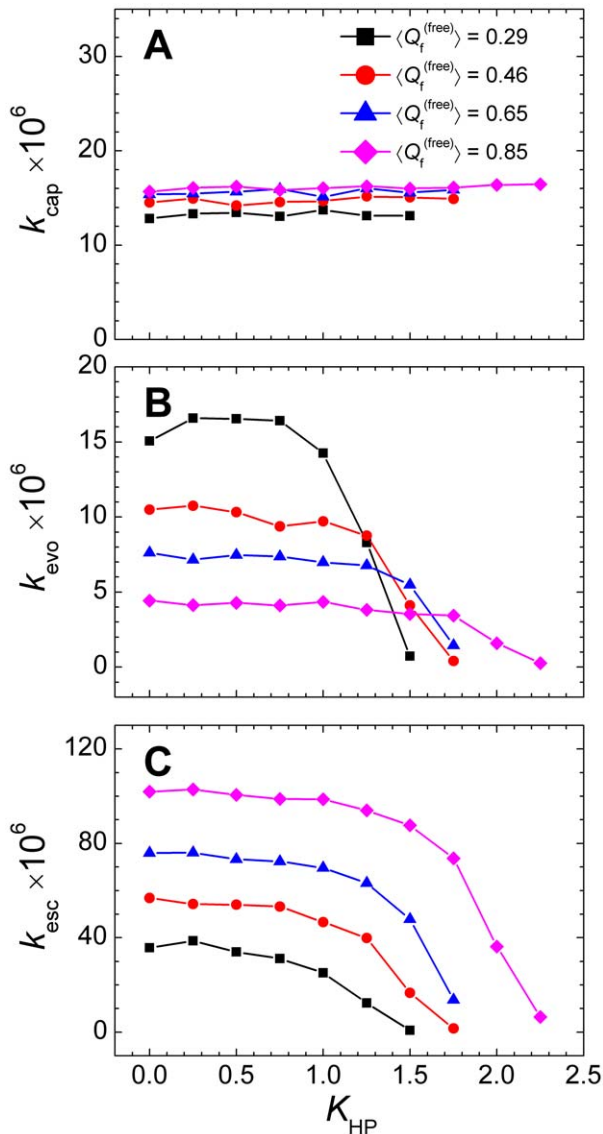doi:10.1371/journal.pone.0015375.g005

Figure 6. Kinetics analysis of the binding process. Effect of the nonnative hydrophobic interactions on (A) capture rate (k_{cap}), (B) evolving rate (k_{evo}), and (C) escape rate (k_{esc}) for systems with various chain flexibilities.
doi:10.1371/journal.pone.0015375.g006

Influence of nonnative contact distance on nonnative hydrophobic interaction

An adequate value for the nonnative hydrophobic contact distance ($\sigma_{h\varphi}$) is not well solved in the coarse-grained models. $\sigma_{h\varphi} = 5.0$ Å was adopted in Ref. [37], which is the same as the value used above. A slightly larger value, $\sigma_{h\varphi} = 5.5$ Å, was employed in Ref. [60]. Recently, a delicate scheme with adjustable $\sigma_{h\varphi}$ values between 5.0 and 7.0 Å was applied to the designed protein Top7 [39]. By analyzing the C_α distance distribution of intermolecular native contact pairs (Figure 7), we found that a value of $\sigma_{h\varphi} = 5.0$ Å is positioned at the lower bound of the distribution. The average native contact distance between hydrophobic residues was 7.5 Å, which is about 1.8 Å smaller than that of other contacts. The effect of nonnative hydrophobic interactions on the binding process is dependent on the contact distance $\sigma_{h\varphi}$ because lengthening $\sigma_{h\varphi}$ will enhance nonnative hydrophobic interactions under the same interaction strength K_{HP}. To confirm whether the $\sigma_{h\varphi}$ value will

alter the above findings that IDPs are likely to form nonnative contacts and their kinetic advantage is enhanced by nonnative hydrophobic interactions, we performed simulations with a larger $\sigma_{h\varphi}$, i.e., 7.5 Å. A similar nonlinear effect on the binding kinetics was observed (Figures 8A,B).

Increasing the contact distance $\sigma_{h\varphi}$ to 7.5 Å enhanced the effect of nonnative hydrophobic interactions on the binding process (Figure 8B); meanwhile the $K_{HP}^{\text{max-rate}}$ value was reduced to a value smaller than or comparable to the strength of the native interactions (Figure 8C). To quantitatively describe the effect of $\sigma_{h\varphi}$ on the binding rates, a linear fit in the regions where the binding rates were increasing in Figures 5B and 8B were performed, and the slope of the fit represented the sensitivity of the binding rates (Figure 8D). Under greater $\sigma_{h\varphi}$ values, not only the binding rates of the disordered system is significantly affected, but also the binding rate of the ordered system is affected. However, Figure 8D shows that the binding rate of the disordered system was increased faster than that of the ordered system under different $\sigma_{h\varphi}$ values.

Discussion

Nonnative interactions are proved experimentally to play an important role in the encounter process and increase the binding rate by reducing the phase space by non-specific binding followed by a two-dimensional diffusion [24–26]. This mechanism is supported by our results that the nonnative hydrophobic interactions stabilize the encounter complexes (Figures 6B,C) and increase the binding rates (Figure 5A). More importantly, our results showed that the relative binding rate of the disordered system increased greater than that of the ordered system, thereby indicating that weak nonnative hydrophobic interactions further amplify the kinetic advantages of IDPs (Figure 5B) [62]. Electrostatic interactions have been shown to increase the "fly-casting" effect of IDPs [67]. From the viewpoint of interaction distance, electrostatic interactions are long-range and mainly affect the diffusion process, whereas nonnative hydrophobic interactions are short-range and mainly act on the two-dimensional searching process. For IDPs with short hydrophobic binding motifs flanked by charged regions, these two factors may combine to produce kinetic advantages over ordered proteins.

Our simulations provided important insights into the binding kinetics in the presence of nonnative hydrophobic interactions for systems with different chain flexibilities. Chain flexibility represents an important advantage to IDPs because such flexibility will facilitate the binding of IDPs to several targets [68,69] with greater binding rates [62,64]. However, nonnative hydrophobic interactions can also hamper specific binding of IDPs to target proteins because of kinetic traps. In our simulations, a disordered chain was found to be more readily trapped into misbinding states (Figures 1,5,8), possibly because the disordered chain lacks structural constraints to prevent misbinding [70]. Structural inspection of snapshots from simulations gave non-specific binding states consistent with those in Ref. [60]. Although hydrophobic residues are sparsely found in IDPs, the interface of complexes formed by IDPs often have more hydrophobic–hydrophobic contacts than protein complexes formed by ordered proteins [71]. Consequently, IDPs possess a misbinding potential. Compared with the specific binding state, misbinding states may be less stable and as shown by Vavouri et al., the off-target binding is tightly regulated through the control of the concentration of IDPs [72].

Although simple, the model adopted in this work allowed us to isolate the effect of chain flexibility and provided insights into the relationship between chain flexibility and protein misbinding/ aggregation. These findings support the importance of structural
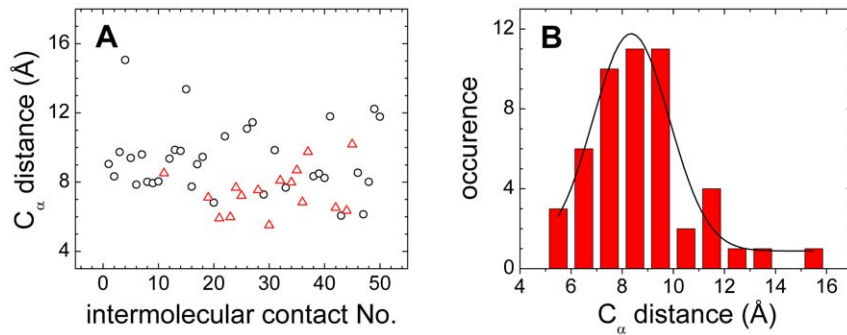
Figure 7. C$_\alpha$ distance distribution of native contacts in the pKID-KIX complex. (A) C$_\alpha$ distance distribution of native contacts formed by two hydrophobic residues (triangles) and others (circles). (B) C$_\alpha$ distance distribution of all native contacts.
doi:10.1371/journal.pone.0015375.g007

constraints in preventing aggregation [70,73] and structural flexibility in binding promiscuity [72], and also provide new clues to regulate protein-protein interactions through controlling the flexibility of proteins or binding motifs.

In conclusion, in this work, we introduced nonnative hydrophobic interactions into the Gō-like model to investigate how they affect the binding process. Using the KIX-pKID complex as a model system, we continuously tuned the chain from a disordered to an ordered form to characterize the interplay between chain flexibility and nonnative interaction effects. The results showed that the free-energy profiles became rugged and the intermediate states occurred when nonnative hydrophobic interactions were introduced. The binding rate initially accelerated and then dramatically decreased as the nonnative hydrophobic interaction

strength was increased. Both thermodynamic and kinetic analysis showed that the disordered system was more readily trapped into non-specific misbinding states than an ordered system. This supports the idea that IDPs are prone to form promiscuous interactions and aggregate. Furthermore, our results showed that weak nonnative hydrophobic interactions amplify the kinetic advantages of IDPs in specific binding processes.

Materials and Methods

Protein structure and quantities describing the coupled folding-binding process

The protein complex used in this study is formed by the phosphorylated kinase-inducible domain (pKID) of the transcrip-
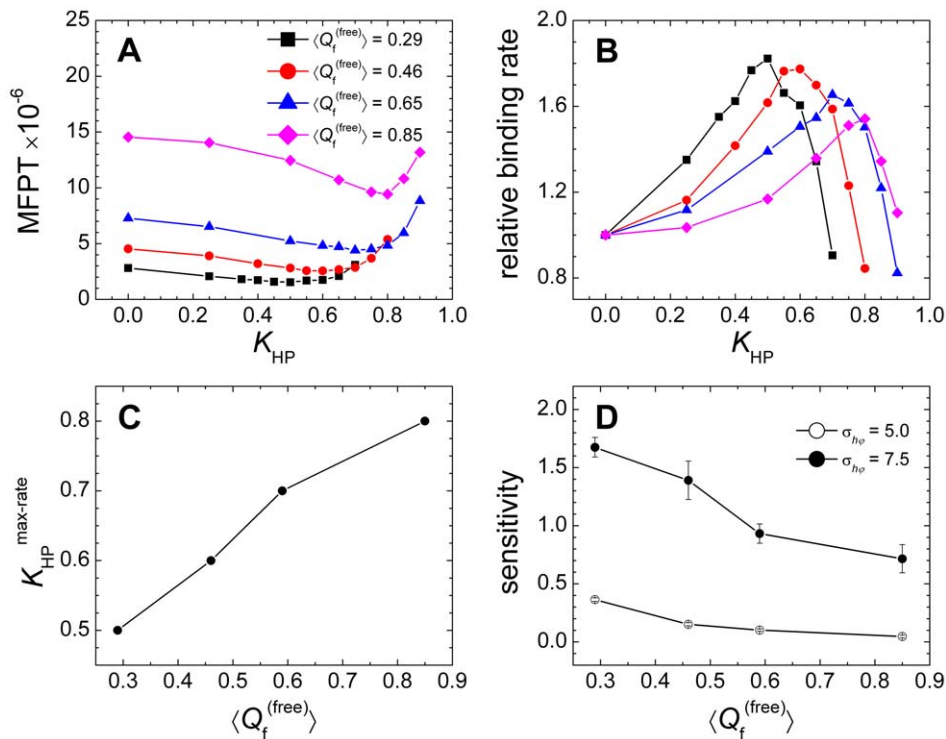

Figure 8. Effect of the nonnative contact distance on the binding kinetics. (A) Mean first passage time (MFPT) of the binding process as a function of K_{HP}. (B) Correlation between the relative binding rate and K_{HP}. (C) Correlation between the K_{HP} corresponding to the maximum binding rate $K_{HP}^{max-rate}$ and $\langle Q_f^{(free)} \rangle$. Contact distance $\sigma_{h\varphi} = 7.5$ Å was used in (A–C). (D) The sensitivity of the binding rate with respect to nonnative hydrophobic interactions for $\sigma_{h\varphi} = 5.0$ and 7.5 Å.
doi:10.1371/journal.pone.0015375.g008

tion factor cAMP response-element binding protein (CREB) and the kinase-inducible domain interacting domain (KIX) of the CREB binding protein [26,65,66]. The pKID domain (Asp119–Pro146) is disordered in the free form and folds into two α-helices upon binding to the structured KIX domain [26]. The native contact set was built based on the CSU software [74]. The fraction of native intramolecular (folding) contacts, Q_f, was used to monitor the folding process, and the fraction of native intermolecular (binding) contacts, Q_b, was used to monitor the binding process. The average fraction of intramolecular native contacts of pKID in its free form, $\langle Q_f^{(free)} \rangle$, was used to quantify the degree of disorder (chain flexibility) of the model.

Modified native-centric Gō-like model with nonnative hydrophobic interactions

In this work, we modified a native-centric continuum Gō-model with coarse-grained C_α chain representation [7,62,75] to include nonnative hydrophobic interactions [37]. In the model system (pKID-KIX complex), the KIX domain was the ordered target and kept frozen during the simulations, whereas the pKID domain was free and tuned from a disordered to an ordered form by increasing the intramolecular interaction strength [62]. Thus, the total potential energy including nonnative hydrophobic interactions is proposed as:

$$V_{total} = V_{stretching}^{pKID} + V_{bending}^{pKID} + V_{torsion}^{pKID} + V_{non-bonded}^{pKID} + V_{non-bonded}^{pKID-KIX}$$

where

$$V_{stretching}^{pKID} = \sum_{bonds} K_r (r - r_0)^2$$

$$V_{bending}^{pKID} = \alpha \left\{ \sum_{angles} {}_\theta K (\theta - \theta_0)^2 \right\}$$

$$V_{torsion}^{pKID} = \alpha \left\{ \sum_{dihedrals} \left\{ K_\phi^{(1)} [1 - \cos(\phi - \phi_0)] + K_\phi^{(3)} [1 - \cos 3(\phi - \phi_0)] \right\} \right\}$$

Compared with Ref. [62], α is also introduced into the bending and torsion terms to better control the chain flexibility. Non-bonded interactions were divided into native interactions, excluded volume repulsions, and nonnative hydrophobic interactions:

$$V_{non-bonded}^{pKID} = \alpha \left\{ \sum_{i<j-3,native}^{pKID} \varepsilon \left[5 \left(\frac{r'_{ij}}{r_{ij}} \right)^{12} - 6 \left(\frac{r'_{ij}}{r_{ij}} \right)^{10} \right] + \sum_{i<j-3,repulsion}^{pKID} \varepsilon \left(\frac{r_{rep}}{r_{ij}} \right)^{12} \right\}$$
$$- K_{HP} \sum_{i,j \in HP}^{pKID} \kappa_i \kappa_j \exp \left[-(r_{ij} - \sigma_{h\varphi})^2 / 2 \right]$$

$$V_{non-bonded}^{pKID-KIX} = \left\{ \sum_{i,j,native}^{pKID-KIX} \varepsilon \left[5 \left(\frac{r'_{ij}}{r_{ij}} \right)^{12} - 6 \left(\frac{r'_{ij}}{r_{ij}} \right)^{10} \right] + \sum_{i,j,repulsion}^{pKID-KIX} \varepsilon \left(\frac{r_{rep}}{r_{ij}} \right)^{12} \right\}$$
$$- K_{HP} \sum_{i,j \in HP}^{pKID-KIX} \kappa_i \kappa_j \exp \left[-(r_{ij} - \sigma_{h\varphi})^2 / 2 \right]$$

r, θ, ϕ, and r_{ij} were the virtual bond length, bond angle, torsion angle, and non-bonded spatial distance defined by the C_α atoms, respectively; r_0, θ_0, ϕ_0, and r'_{ij} were the corresponding native values available from the PDB structure (PDB code for the pKID-KIX complex is 1KDX [66]). Non-bonded interactions were only considered when two C_α atoms i and j were separated sequentially by at least three residues within one chain (the pKID domain) or

when they came from different chains. For native interactions, a 12-10 Lennard-Jones (LJ) form potential was used; whereas for nonnative interactions, r_{rep} parameterized the excluded volume repulsion between residue pairs that do not belong to the given native contact set. The nonnative hydrophobic interaction term adopted here has been used to account for the nonnative hydrophobic interactions present in protein folding processes [37]. Alanine, valine, leucine, isoleucine, phenylalanine, methionine, tryptophan, and tyrosine were considered as hydrophobic amino acids. We did not distinguish between hydrophobic residue types, so κ_i was the same for all hydrophobic residues and set at 1.0. The overall strength of the hydrophobic interactions was controlled by K_{HP}. We adopted $\sigma_{h\varphi} = 5.0$ Å as in Ref. [37] to control the hydrophobic interactions, and also tested the influence of the $\sigma_{h\varphi}$ value (see *Influence of nonnative contact distance on nonnative hydrophobic interaction*). The summation in the nonnative hydrophobic interaction term excludes hydrophobic pairs in the native contact because they have already been included in the native interactions. Other parameters were set $r_{rep} = 4.0$ Å, $K_r = 100$ ε, $K_\theta = 20$ ε, $K_\phi^{(1)} = \varepsilon$, $K_\phi^{(3)} = 0.5$ ε. The interaction strength was controlled by the parameter ε, which was fixed at 1.0 in this study. The parameter α scaled the intramolecular interactions within the pKID domain and tuned the degree of chain disorder (flexibility): $\langle Q_f^{(free)} \rangle$ was calculated to be 0.29, 0.46, 0.59 and 0.85 for α = 0.1, 1.0, 1.5, and 3.0, respectively.

Thermodynamics and kinetics simulations

Simulations were performed by Langevin dynamics in an over-damped region, with a friction constant of $0.1 \ \tau^{-1}$ ($\tau = \sqrt{ma^2 / \varepsilon_0}$), where the length scale a was set to 4 Å, the mass m was set to 1.0, and the reference energy scale ε_0 was 1.0 as in Ref. [76,77]. The molecular dynamics time step was set to $0.005 \ \tau^{-1}$. Simulation temperatures were chosen to be the transition temperature of the binding process, i.e., the temperature where the system has equal probability in the bound state and the unbound state, when the nonnative hydrophobic interaction strength K_{HP} was set to zero. Other parameters were set as in Ref. [62].

A pKID chain and a KIX chain were put in a 200 Å cubic box with periodic boundary conditions. The KIX domain was kept frozen at the box center while the pKID was free to move. A high temperature unbinding simulation was performed to provide 400 randomly chosen unbound conformations. An unbound conformation was defined by the fraction of native contact $Q_b = 0$ and a mass-center distance between the two proteins $\Delta R > 45$ Å. Subsequently, 400 binding simulations were performed starting from 400 unbound structures. A bound state was considered to occur when the system reached the minimum of the free energy as in Ref. [62]. The encounter state was reached when the system evolved from an unbound state to a state with $Q_b > 0$ (usually have one native contact). Kinetic data were averaged from the resulting trajectories.

By dissecting a binding trajectory into an encounter step, an escape step, and an evolution step, we accumulated the transition number (N) and the averaged transition time (measured by the mean passage time, MPT) between any two states. The escape rate k_{esc} and the evolving rate k_{evo} were calculated as:

$$k_{esc} = [(MPT_{esc} \times N_{esc} + MPT_{evo} \times N_{evo}) \times N_{esc} / (N_{esc} + N_{evo})]^{-1}$$

$$k_{evo} = [(MPT_{esc} \times N_{esc} + MPT_{evo} \times N_{evo}) \times N_{evo} / (N_{esc} + N_{evo})]^{-1}$$

where $\mathrm{MPT_{esc}}$ and $\mathrm{MPT_{evo}}$ are the mean passage time from the encounter state to the unbound state and from the encounter state to the bound state, respectively; $\mathcal{N}_{esc}$ and $\mathcal{N}_{evo}$ are the corresponding numbers of transitions. The capture rate was calculated as $k_{cap} = \mathrm{MPT_{cap}}^{-1}$; $\mathrm{MPT_{cap}}$ is the mean passage time from the unbound state to the encounter state.

The bias potential and the histogram technique were used for conformational sampling [78,79]. The free energy was calculated as $F(Q_b)/k_B T = -\ln[P(Q_b)]$, where $P(Q_b)$ is the normalized population distribution as a function of Q_b.

References

1. Dill KA, Chan HS (1997) From Levinthal to pathways to funnels. Nat Struct Biol 4: 10–19.
2. Bryngelson JD, Onuchic JN, Socci ND, Wolynes PG (1995) Funnels, pathways, and the energy landscape of protein folding: a synthesis. Proteins 21: 167–195.
3. Tsai CJ, Kumar S, Ma BY, Nussinov R (1999) Folding funnels, binding funnels, and protein function. Protein Sci 8: 1181–1190.
4. Ueda Y, Taketomi H, Gō N (1978) Studies on protein folding, unfolding, and fluctuations by computer simulation. II. A. three-dimensional lattice model of lysozyme. Biopolymers 17: 1531–1548.
5. Gō N (1983) Theoretical studies of protein folding. Annu Rev Biophys Bioeng 12: 183–210.
6. Micheletti C, Banavar JR, Maritan A, Seno F (1999) Protein structures and optimal folding from a geometrical variational principle. Phys Rev Lett 82: 3372–3375.
7. Clementi C, Nymeyer H, Onuchic JN (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. J Mol Biol 298: 937–953.
8. Liu ZR, Chan HS (2005) Desolvation is a likely origin of robust enthalpic barriers to protein folding. J Mol Biol 349: 872–889.
9. Levy Y, Cho SS, Onuchic JN, Wolynes PG (2005) A survey of flexible protein binding mechanisms and their transition states using native topology based energy landscapes. J Mol Biol 346: 1121–1145.
10. Qi YF, Huang YQ, Liang HH, Liu ZR, Lai LH (2010) Folding simulations of a de novo designed protein with a βαβ fold. Biophys J 98: 321–329.
11. Cho J-H, Sato S, Raleigh DP (2004) Thermodynamics and kinetics of non-native interactions in protein folding: a single point mutant significantly stabilizes the N-terminal domain of L9 by modulating non-native interactions in the denatured state. J Mol Biol 338: 827–837.
12. Cho J-H, Raleigh DP (2005) Mutational analysis demonstrates that specific electrostatic interactions can play a key role in the denatured state ensemble of proteins. J Mol Biol 353: 174–185.
13. Capaldi AP, Kleanthous C, Radford SE (2002) Im7 folding mechanism: misfolding on a path to the native state. Nat Struct Biol 9: 209–216.
14. Otzen D (2005) Antagonism, non-native interactions and non-two-state folding in S6 revealed by double-mutant cycle analysis. Protein Eng Des Sel 18: 547–557.
15. Neudecker P, Zarrine-Afsar A, Choy W-Y, Muhandiram DR, Davidson AR, et al. (2006) Identification of a collapsed intermediate with non-native long-range interactions on the folding pathway of a pair of Fyn SH3 domain mutants by NMR relaxation dispersion spectroscopy. J Mol Biol 363: 958–976.
16. Morton VL, Friel CT, Allen LR, Paci E, Radford SE (2007) The effect of increasing the stability of non-native interactions on the folding landscape of the bacterial immunity protein Im9. J Mol Biol 371: 554–568.
17. Rothwarf DM, Scheraga HA (1996) Role of non-native aromatic and hydrophobic interactions in the folding of hen egg white lysozyme. Biochemistry 35: 13797–13807.
18. Salvatella X, Dobson CM, Fersht AR, Vendruscolo M (2005) Determination of the folding transition states of barnase by using Φ_I-value-restrained simulations validated by double mutant Φ_{IJ}-values. Proc Natl Acad Sci USA 102: 12389–12394.
19. Blanco FJ, Ortiz AR, Serrano L (1997) Role of a nonnative interaction in the folding of the protein G B1 domain as inferred from the conformational analysis of the α-helix fragment. Fold Des 2: 123–133.
20. Viguera AR, Vega C, Serrano L (2002) Unspecific hydrophobic stabilization of folding transition states. Proc Natl Acad Sci USA 99: 5349–5354.
21. Di Nardo AA, Korzhnev DM, Stogios PJ, Zarrine-Afsar A, Kay LE, et al. (2004) Dramatic acceleration of protein folding by stabilization of a nonnative backbone conformation. Proc Natl Acad Sci USA 101: 7954–7959.
22. Canet D, Lyon CE, Scheek RM, Robillard GT, Dobson CM, et al. (2003) Rapid formation of non-native contacts during the folding of HPr revealed by real-time photo-CIDNP NMR and stopped-flow fluorescence experiments. J Mol Biol 330: 397–407.
23. Gardino AK, Villali J, Kivenson A, Lei M, Liu CF, et al. (2009) Transient non-native hydrogen bonds promote activation of a signaling protein. Cell 139: 1109–1118.

24. Zhou H-X (2005) How do biomolecular systems speed up and regulate rates? Phys Biol 2: R1–R25.
25. Tang C, Iwahara J, Clore GM (2006) Visualization of transient encounter complexes in protein-protein association. Nature 444: 383–386.
26. Sugase K, Dyson HJ, Wright PE (2007) Mechanism of coupled folding and binding of an intrinsically disordered protein. Nature 447: 1021–1025.
27. Sánchez IE, Ferreiro DU, Dellarole M, de Prat-Gay G (2010) Experimental snapshots of a protein-DNA binding landscape. Proc Natl Acad Sci USA 107: 7751–7756.
28. Paci E, Vendruscolo M, Karplus M (2002) Validity of Gō models: comparison with a solvent-shielded empirical energy decomposition. Biophys J 83: 3032–3038.
29. Chen YT, Ding JD (2010) Roles of non-native hydrogen-bonding interaction in helix-coil transition of a single polypeptide as revealed by comparison between Gō-like and non-Gō models. Proteins 78: 2090–2100.
30. Paci E, Vendruscolo M (2005) Detection of non-native hydrophobic interactions in the denatured state of lysozyme by molecular dynamics simulations. J Phys: Condens Matter 17: S1617–S1626.
31. Wrabl JO, Shortle D (1996) Perturbations of the denatured state ensemble: modeling their effects on protein stability and folding kinetics. Protein Sci 5: 2343–2352.
32. Li L, Mirny LA, Shakhnovich EI (2000) Kinetics, thermodynamics and evolution of non-native interactions in a protein folding nucleus. Nat Struct Biol 7: 336–342.
33. Clementi C, Plotkin SS (2004) The effects of nonnative interactions on protein folding rates: theory and simulation. Protein Sci 13: 1750–1766.
34. Plotkin SS (2001) Speeding protein folding beyond the Gō model: how a little frustration sometimes helps. Proteins 45: 337–345.
35. Fan K, Wang J, Wang W (2002) Folding of lattice protein chains with modified Gō potential. Eur Phys J B 30: 381–391.
36. Cieplak M, Hoang TX (2002) The range of the contact interactions and the kinetics of the Go models of proteins. Int J Mod Phys C 13: 1231–1242.
37. Zarrine-Afsar A, Wallin S, Neculai AM, Neudecker P, Howell PL, et al. (2008) Theoretical and experimental demonstration of the importance of specific nonnative interactions in protein folding. Proc Natl Acad Sci USA 105: 9999–10004.
38. Zhang ZQ, Chan HS (2009) Native topology of the designed protein Top7 is not conducive to cooperative folding. Biophys J 96: L25–L27.
39. Zhang ZQ, Chan HS (2010) Competition between native topology and nonnative interactions in simple and complex folding kinetics of natural and designed proteins. Proc Natl Acad Sci USA 107: 2920–2925.
40. Wallin S, Zeldovich KB, Shakhnovich EI (2007) The folding mechanics of a knotted protein. J Mol Biol 368: 884–893.
41. Azia A, Levy Y (2009) Nonnative electrostatic interactions can modulate protein folding: molecular dynamics with a grain of salt. J Mol Biol 393: 527–542.
42. Chan HS, Dill KA (1994) Transition states and folding dynamics of proteins and heteropolymers. J Chem Phys 100: 9238–9257.
43. Karanicolas J, Brooks CL, III (2003) Improved Gō-like models demonstrate the robustness of protein folding mechanisms towards non-native interactions. J Mol Biol 334: 309–325.
44. Gin BC, Garrahan JP, Geissler PL (2009) The limited role of nonnative contacts in the folding pathways of a lattice protein. J Mol Biol 392: 1303–1314.
45. Gao M, Yao XQ, She ZS, Liu ZR, Zhu HQ (2010) Intermediate structure and slow hydration water dynamics in protein folding process. Acta Phys -Chim Sin 26: 1998–2006.
46. Kuhlman B, Dantas G, Ireton GC, Varani G, Stoddard BL, et al. (2003) Design of a novel globular protein fold with atomic-level accuracy. Science 302: 1364–1368.
47. Zhou H-X, Szabo A (2004) Enhancement of association rates by nonspecific binding to DNA and cell membranes. Phys Rev Lett 93: 178101.
48. Curcio R, Caflisch A, Paci E (2005) Change of the unbinding mechanism upon a mutation: a molecular dynamics study of an antibody-hapten complex. Protein Sci 14: 2499–2514.
49. Wang J, Lu Q, Lu HP (2006) Single-molecule dynamics reveals cooperative binding-folding in protein recognition. PLoS Comput Biol 2: e78.
50. Alsallaq R, Zhou H-X (2008) Electrostatic rate enhancement and transient complex of protein-protein association. Proteins 71: 320–335.

51. Givaty O, Levy Y (2009) Protein sliding along DNA: dynamics and structural characterization. J Mol Biol 385: 1087–1097.

52. Wright PE, Dyson HJ (1999) Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. J Mol Biol 293: 321–331.

53. Uversky VN, Dunker AK (2010) Understanding protein non-folding. Biochim Biophys Acta 1804: 1231–1264.

54. Turoverov KK, Kuznetsova IM, Uversky VN (2010) The protein kingdom extended: ordered and intrinsically disordered proteins, their folding, supramolecular complex formation, and aggregation. Prog Biophys Mol Biol; doi:10.1016/j.pbiomolbio.2010.01.003.

55. Huang YQ, Liu ZR (2010) Intrinsically disordered proteins: the new sequence-structure-function relations. Acta Phys -Chim Sin 26: 2061–2072.

56. Ward JJ, Sodhi JS, McGuffin LJ, Buxton BF, Jones DT (2004) Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. J Mol Biol 337: 635–645.

57. Oldfield CJ, Cheng Y, Cortese MS, Brown CJ, Uversky VN, et al. (2005) Comparing and combining predictors of mostly disordered proteins. Biochemistry 44: 1989–2000.

58. Midic U, Oldfield CJ, Dunker AK, Obradovic Z, Uversky VN (2009) Protein disorder in the human diseasome: unfoldomics of human genetic diseases. Bmc Genomics 10: S12.

59. Dyson HJ, Wright PE (2002) Coupling of folding and binding for unstructured proteins. Curr Opin Struct Biol 12: 54–60.

60. Turjanski AG, Gutkind JS, Best RB, Hummer G (2008) Binding-induced folding of a natively unstructured transcription factor. PLoS Comput Biol 4: e1000060.

61. Lu Q, Lu HP, Wang J (2007) Exploring the mechanism of flexible biomolecular recognition with single molecule dynamics. Phys Rev Lett 98: 128105.

62. Huang YQ, Liu ZR (2009) Kinetic advantage of intrinsically disordered proteins in coupled folding-binding process: a critical assessment of the "fly-casting" mechanism. J Mol Biol 393: 1143–1159.

63. Huang YQ, Liu ZR (2010) Smoothing molecular interactions: the "kinetic buffer" effect of intrinsically disordered proteins. Proteins;DOI: 10.1002/prot.22820.

64. Shoemaker BA, Portman JJ, Wolynes PG (2000) Speeding molecular recognition by using the folding funnel: the fly-casting mechanism. Proc Natl Acad Sci USA 97: 8868–8873.

65. Chrivia JC, Kwok RPS, Lamb N, Hagiwara M, Montminy MR, et al. (1993) Phosphorylated CREB binds specifically to the nuclear protein CBP. Nature 365: 855–859.

66. Radhakrishnan I, Pérez-Alvarado GC, Parker D, Dyson HJ, Montminy MR, et al. (1997) Solution structure of the KIX domain of CBP bound to the transactivation domain of CREB: a model for activator:coactivator interactions. Cell 91: 741–752.

67. Levy Y, Onuchic JN, Wolynes PG (2007) Fly-casting in protein-DNA binding: frustration between protein folding and electrostatics facilitates target recognition. J Am Chem Soc 129: 738–739.

68. Kriwacki RW, Hengst L, Tennant L, Reed SI, Wright PE (1996) Structural studies of p21Waf1/Cip1/Sdi1 in the free and Cdk2-bound state: conformational disorder mediates binding diversity. Proc Natl Acad Sci USA 93: 11504–11509.

69. Oldfield CJ, Meng JW, Yang JY, Yang MQ, Uversky VN, et al. (2008) Flexible nets: disorder and induced fit in the associations of p53 and 14-3-3 with their partners. Bmc Genomics 9: S1.

70. Pechmann S, Levy ED, Tartaglia GG, Vendruscolo M (2009) Physicochemical principles that regulate the competition between functional and dysfunctional association of proteins. Proc Natl Acad Sci USA 106: 10159–10164.

71. Mészáros B, Tompa P, Simon I, Dosztányi Z (2007) Molecular principles of the interactions of disordered proteins. J Mol Biol 372: 549–561.

72. Vavouri T, Semple JI, Garcia-Verdugo R, Lehner B (2009) Intrinsic protein disorder and interaction promiscuity are widely associated with dosage sensitivity. Cell 138: 198–208.

73. Goldschmidt L, Teng PK, Riek R, Eisenberg D (2010) Identifying the amylome, proteins capable of forming amyloid-like fibrils. Proc Natl Acad Sci USA 107: 3487–3492.

74. Sobolev V, Sorokine A, Prilusky J, Abola EE, Edelman M (1999) Automated analysis of interatomic contacts in proteins. Bioinformatics 15: 327–332.

75. Liu ZR, Chan HS (2005) Solvation and desolvation effects in protein folding: native flexibility, kinetic cooperativity and enthalpic barriers under isostability conditions. Phys Biol 2: S75–S85.

76. Kaya H, Chan HS (2003) Solvation effects and driving forces for protein thermodynamic and kinetic cooperativity: how adequate is native-centric topological modeling? J Mol Biol 326: 911–931.

77. Kaya H, Chan HS (2004) Solvation effects and driving forces for protein thermodynamic and kinetic cooperativity: how adequate is native-centric topological modeling? Corrigendum. J Mol Biol 337: 1069–1070.

78. Boczko EM, Brooks III CL (1993) Constant-temperature free energy surfaces for physical and chemical processes. J Phys Chem 97: 4509–4513.

79. Badasyan A, Liu ZR, Chan HS (2008) Probing possible downhill folding: native contact topology likely places a significant constraint on the folding cooperativity of proteins with ~40 residues. J Mol Biol 384: 512–530.