Structure
Article

Prediction of Protein-Protein Association Rates from a Transition-State Theory

Ramzi Alsallaq[1] and Huan-Xiang Zhou[1,*]
[1] Department of Physics and Institute of Molecular Biophysics and School of Computational Science, Florida State University, Tallahassee, FL 32306, USA
*Correspondence: zhou@sb.fsu.edu
DOI 10.1016/j.str.2007.01.005

SUMMARY

We recently developed a theory for the rates of protein-protein association. The theory is based on the concept of a transition state, which separates the bound state, with numerous short-range interactions but restricted translational and rotational freedom, and the unbound state, with, at most, a small number of interactions but expanded configurational freedom. When not accompanied by large-scale conformational changes, protein-protein association becomes diffusion limited. The association rate is then predicted as $k_a = k_a^0 \exp(-\Delta G_{el}^{\ddagger}/k_B T)$, where $\Delta G_{el}^{\ddagger}$ is the electrostatic interaction free energy in the transition state, k_a^0 is the rate in the absence of electrostatic interactions, and $k_B T$ is thermal energy. Here, this transition-state theory is used to predict the association rates of four protein complexes. The predictions for the wild-type complexes and 23 mutants are found to agree closely with experimental data over wide ranges of ionic strength.

INTRODUCTION

Rapid association between proteins is crucial in a wide array of biological processes such as the utilization of and self-defense against toxins (Wallis et al., 1995; Jucovic and Hartley, 1996; Schreiber and Fersht, 1996; Terlau et al., 1996; Radic et al., 1997), receptor activation by growth hormones and cytokines (Shen et al., 1996; Wells, 1996; Darling et al., 2002), and regulation of actin polymerization (Marchand et al., 2001; Hemsath et al., 2005). In contrast to the study of protein folding kinetics, in which theoretical developments (Bryngelson and Wolynes, 1987; Dill and Chan, 1997; Munoz et al., 1997; Onuchic et al., 1997; Baker, 2000) and molecular dynamics simulations (Daggett and Levitt, 1993; Duan and Kollman, 1998; Garcia and Sanbonmatsu, 2001; Simmerling et al., 2002; Snow et al., 2002; Mayor et al., 2003; Huang and Zhou, 2006) have worked in concert to elucidate mechanisms, modeling efforts of protein association kinetics have mostly focused on Brownian dynamics simulations (Northrup and Erickson, 1992; Zhou, 1993; Gabdoulline and Wade, 1997, 2001; Elcock et al., 1999; Zou and Skeel, 2003; Lin and Beratan, 2005; Haddadian and Gross, 2006; Spaar et al., 2006). Recently, we developed a theory for the rates of protein-protein association (Alsallaq and Zhou, 2007). Here, we extensively test this theory against experimental data for 4 protein-protein complexes and 23 of their mutants over wide ranges of ionic strength.

The theory is based on the concept of a transition state for protein-protein association. The notion of a transition state has been invoked in previous studies of protein-protein association (Gabdoulline and Wade, 1997, 2001; Vijayakumar et al., 1998; Elcock et al., 1999; Frisch et al., 2001; Zhou, 2001, 2003; Miyashita et al., 2004; Schlosshauer and Baker, 2004; Spaar et al., 2006). In our recent theory, this concept is fully developed. As illustrated in Figure 1A, the bound state of two associating proteins is defined by a deep well of the interaction energy. The transition state is located at the outer boundary of the bound state and marks the end point of the sharp increase in interaction energy. It also marks the starting point of increased freedom in relative translation and rotation between the two proteins.

The rate of association can be approximated by $k_a k_+/(k_+ + k_-)$, where k_a is the diffusion-limited rate for reaching the transition state, k_+ is the rate of forming the stereospecific complex from the transition state through conformational rearrangement, and k_- is the rate of dissociation from the transition state (Zhou et al., 1997). If no large-scale conformational changes are involved, then one expects $k_+ >> k_-$, and the rate of association approaches k_a. For an intermolecular force to affect k_a, it must be present in the diffusion process that leads to the transition state and thus must be long ranged. Electrostatic interactions between associating proteins provide the dominant long-range force and can significantly increase the magnitude of k_a. In theoretical studies, it has been shown that the electrostatic rate enhancement can be accurately predicted by a transition-state theory formula (Zhou, 1996, 1997):

$$k_a = k_a^0 \exp\left(-\Delta G_{el}^{\ddagger}/k_B T\right), \quad (1)$$

where k_a^0 is the basal rate, i.e., the rate in the absence of electrostatic interactions, $\Delta G_{el}^{\ddagger}$ is the electrostatic interaction free energy in the transition state, and $k_B T$ is thermal energy. Note that though Equation 1 formally resembles a transition-state theory, its derivation was actually not

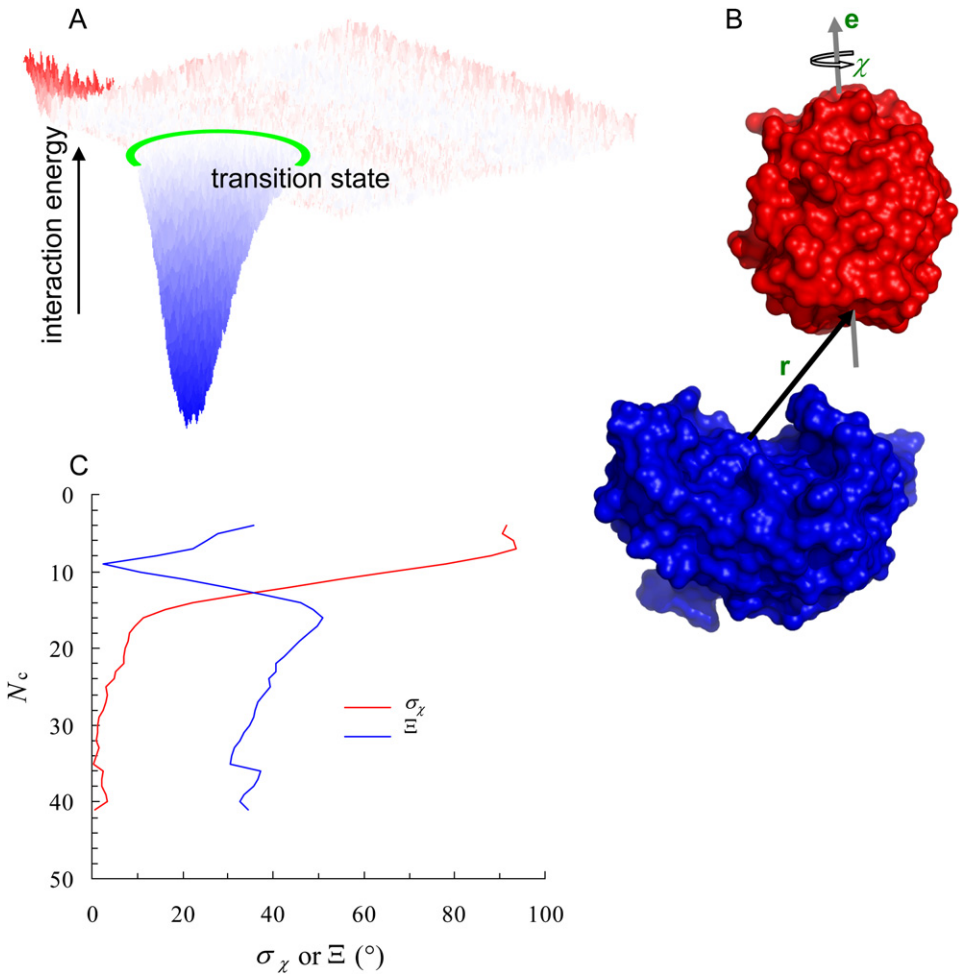


Figure 1. Illustration of the Energy Landscape for Protein-Protein Association

(A) The bound state is located in a deep well of the interaction energy. The transition state, indicated by a green ring, marks the end of the sharp increase in interaction energy and the start of the increase in translational and rotation freedom.

(B) The six degrees of relative translational and rotational freedom between two associating proteins. One protein, shown in blue, was fixed in space; the other, shown in red, was allowed to translate and rotate. The three translational degrees of freedom were represented by the displacement vector, **r**, between the centers of the binding sites on the two proteins. Of the three rotational degrees of freedom, two were a unit vector, **e**, attached to the moving protein, and the remaining one was the rotational angle, χ, around the unit vector. In the X-ray structure of the bound complex, the relative separation, r, is 0, and the rotation angle, χ, is 0°.

(C) Transition of the standard deviation of χ, σ_χ, from the bound state (with high contact levels) to the unbound state (with low contact levels). The start of the sharp increase in σ_χ marks the transition state, which was uniquely determined by a maximum of Ξ. Actual data for the IL4-IL4BP complex are shown for illustration (with $N_c^{\ddagger} = 16$).

trivial at all. It is an approximation that is valid only when the complex formed is stereospecific and the intermolecular force is long ranged, conditions that are fortuitously fulfilled for protein-protein association under the influence of electrostatic interactions. Implementation of Equation 1 entails generating configurations in the transition state and calculating the electrostatic interaction energies over these configurations.

Considerations based on orientational constraints for protein association have suggested that the basal rate lies in the range of 10^5–10^6 $M^{-1}s^{-1}$ (Northrup and Erickson, 1992; Zhou, 1997; Schlosshauer and Baker, 2004). Experimental results typically extrapolate to this range at high ionic strengths, which screen electrostatic interactions between proteins. Rates over this range then indicate electrostatic enhancement. As noted previously (Zhou, 2001, 2003), the presence of electrostatic rate enhancement is accompanied by a telltale sign, in the form of disparate ionic strength effects on the association and dissociation rates. For the association of proteins in which electrostatic rate enhancement is implicated, it is widely observed that the association rate shows large variations with ionic strength, while the dissociation rate (k_d) is modestly affected by ionic strength. The transition-state theory provides a simple, qualitative explanation for the disparate ionic strength effects. As the transition state is located at

the outer boundary of the bound state and, hence, is close to it, ionic strength is expected to screen electrostatic interactions in the two states to nearly the same extent. Hence, the effect of ionic strength on the binding affinity (i.e., k_a/k_d) will be almost solely confined to the association rate. In this sense, ionic strength in protein association is analogous to a mutation with a Φ value of 1 in protein folding. Like a native contact that is present in the transition state for protein folding, electrostatic interactions are almost fully formed in the transition state for protein association.

In this paper, we go beyond qualitative explanation and aim for a quantitative account of effects of ionic strength and mutations on the association rates of four protein complexes. These are formed by barnase (Bn) and barstar (Bs), DNase E9 and its cognate immunity protein Im9, interlukin-4 (IL4) and interlukin-4-binding protein (IL4BP), and fasciculin 2 (Fas) and acetylcholinesterase (AChE). They represent a broad structural and functional spectrum. Bn is a ribonuclease produced by *Bacillus amyloliquefaciens* and is excreted as a weapon against predators or competitors. Binding of the inhibitor Bs protects the host cell from the potentially lethal ribonuclease activity of the enzyme. The binding of Im9 to E9 plays an analogous role in *Escherichia coli*. Binding of the IL4 cytokine to IL4BP, the extracellular domain of the IL4 receptor α chain, is a critical event in the regulation of immune responses. Fas is a snake toxin from the venom of the green mamba that inhibits the enzyme, AChE, responsible for the breakdown of the neural transmitter acetylcholine.

Effects of mutations and ionic strength on the binding affinities of these complexes have been studied recently (Dong et al., 2003; Dong and Zhou, 2006). An initial test of the transition-state theory on the ionic strength dependence of the association rate of the Bn-Bs complex appeared promising (Alsallaq and Zhou, 2007). To our knowledge, the study presented here, covering all 4 of the complexes and 23 of their mutants over wide ranges of ionic strength, is the most comprehensive theoretical investigation of protein-protein association rates. The overall good agreement with experimental data, obtained without human intervention, indicates that the transition-state theory holds truly predictive power.

RESULTS

The transition state for protein-protein association separates the bound state, with numerous short-range interactions but restricted translational and rotational freedom, and the unbound state, with, at most, a small number of interactions but expanded configurational freedom. Specifically, it is located at the outer boundary of the bound state and is uniquely identified by the onset of a sharp increase in translational and rotational freedom. In particular, the sampling range of the rotational angle, χ (Figure 1B), experiences a sharp increase as the two proteins move out of the bound state. This increase is reflected by the magnitude of σ_χ, the standard deviation of χ sampled at a given level of interaction energy (Figure 1C).

The interaction energy was actually mimicked by N_c, the total number of native and nonnative contacts. The transition state was identified by the level of N_c that maximized

$$\Xi(N_c) = <\sigma_\chi(N'_c)>_{N'_c<N_c} - \sigma_\chi(N_c), \quad (2)$$

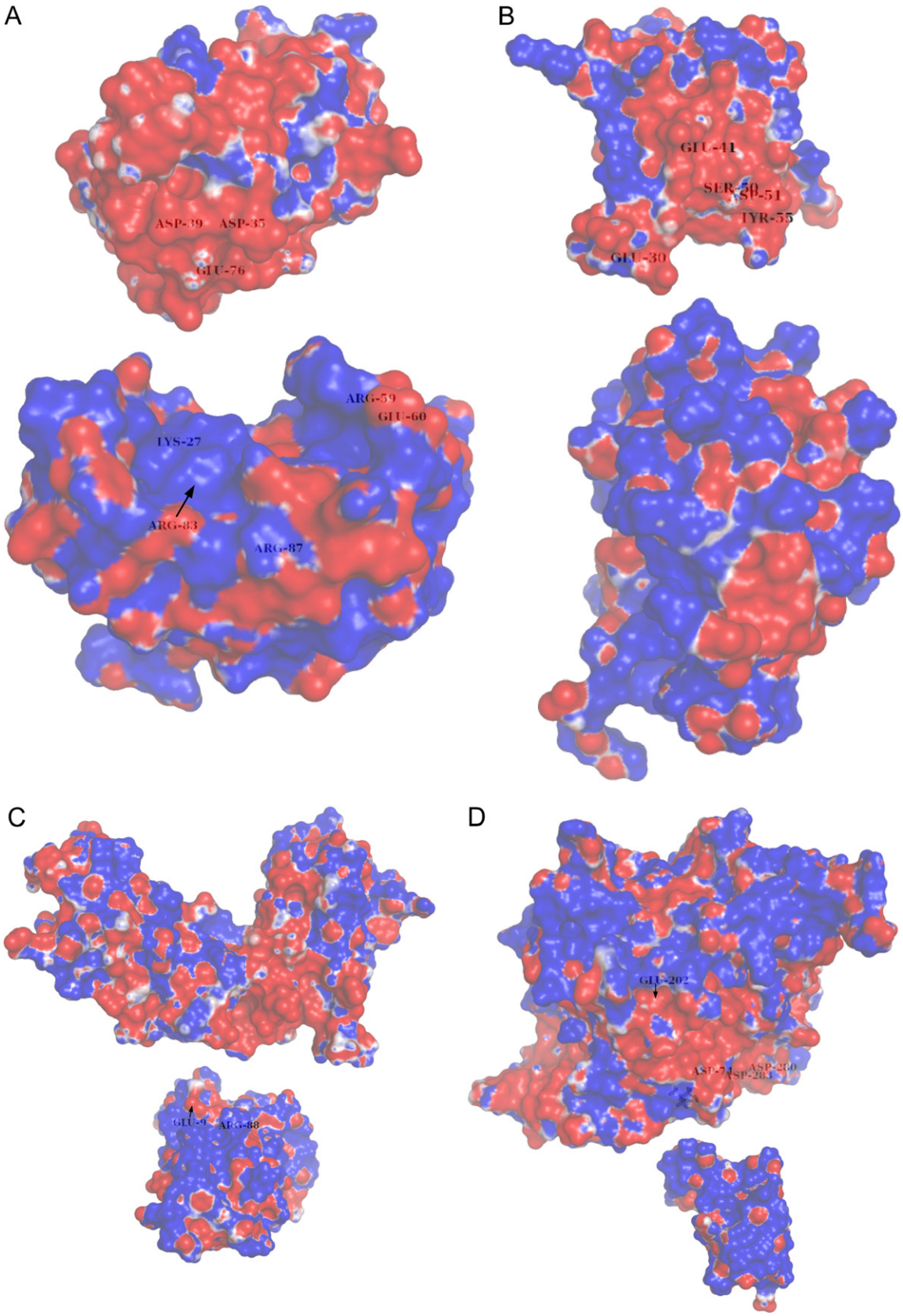
which represents the difference between the standard deviation of χ at contact level N_c and the average for all lower contact levels. The transition-state contact level is denoted as $N_c^\ddagger$ (Figure 1C).

The number of contacts, N_c, was 38, 42, 35, and 52, respectively, in the X-ray structures of the Bn-Bs, E9-Im9, IL4-IL4BP, and Fas-AChE complexes. The corresponding transition-state contact level, $N_c^\ddagger$, was 14, 24, 16, and 23. Among the transition-state ensembles of the four complexes, the relative separation, r, averaged 4.9, 3.9, 4.2, and 5.1 Å, respectively, with standard deviations all at 0.5 Å. The relative rotation angle, χ, distributed around 0° with standard deviations at 18°, 9°, 11°, and 6°, respectively.

In each of the four protein complexes, the two subunits have opposite net charges. As shown in Figure 2, the electrostatic surfaces are highly complementary. The complementarity underlies the electrostatic rate enhancement observed on these complexes (Wallis et al., 1995; Schreiber and Fersht, 1996; Shen et al., 1996; Radic et al., 1997).

Using 100 representative configurations of the transition-state ensemble, the electrostatic interaction free energy was calculated by solving the linearized Poisson-Boltzmann (PB) equation. In Figure 3, association rates predicted by Equation 1 are compared with experimental results for the four complexes (Wallis et al., 1995; Schreiber and Fersht, 1996; Shen et al., 1996; Radic et al., 1997) (data listed in Table S1; see the Supplemental Data available with this article online). In this comparison, the basal association rate, k_a^0, was left as an adjustable parameter. With k_a^0 set to 1.4×10^6, 5×10^5, 5×10^4, and 5×10^4 $M^{-1}s^{-1}$, respectively, for Bn-Bs, E9-Im9, IL4-IL4BP, and Fas-AChE, the ionic-strength (I) dependences of k_a for the complexes are well predicted overall by the theory. The sharp decreases of k_a with increasing ionic strength are in line with electrostatic rate enhancement for these complexes.

For three of the four complexes (Bn-Bs being the exception), the predictions overestimated the rates of association at the lowest ionic strengths. To investigate possible reasons for the overestimation, we also calculated the electrostatic interaction free energy by solving the nonlinear PB equation. The magnitudes of $\Delta G_{el}^\ddagger$ were found to be reduced upon using the nonlinear PB equation at the lowest ionic strengths. Specifically, $\Delta G_{el}^\ddagger$ changed from −3.55 to −3.30 kcal/mol at I = 13 mM for Bn-Bs, but from −6.77 to −5.85 kcal/mol at I = 25 mM for E9-Im9, from −4.31 to −3.66 kcal/mol at I = 150 mM for IL4-IL4BP, and from −4.93 to −3.80 kcal/mol at I = 50 mM for Fas-AChE. The differences in $\Delta G_{el}^\ddagger$ calculated from the linearized and nonlinear PB equations diminished at higher ionic



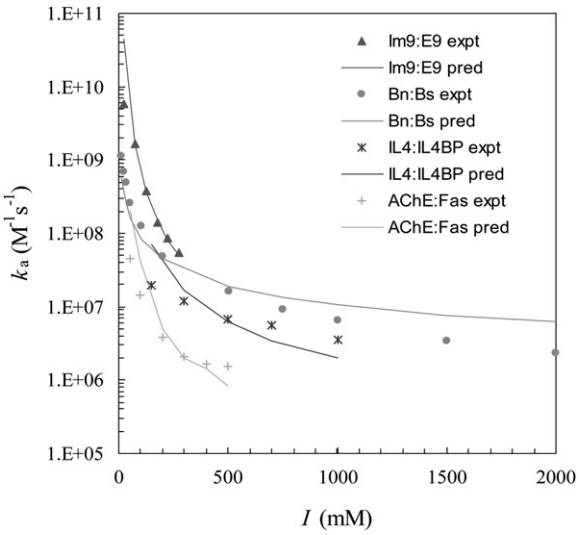


Figure 3. Dependences of k_a on Ionic Strength for Wild-Type Complexes

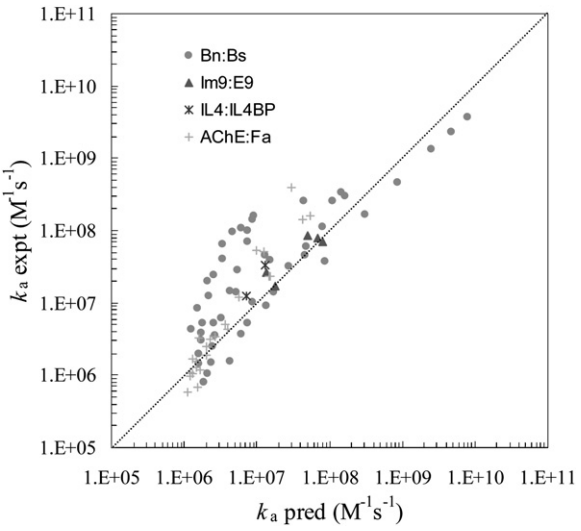


Figure 4. Comparison of k_a between Prediction and Experiment for 23 Mutants

Mutants were studied over different ranges of ionic strength, resulting in a total of 81 data points. A diagonal line is drawn to indicate ideal agreement.

strengths. The values changed from −0.87 to −0.82 kcal/mol at I = 2000 mM for Bn-Bs, from −2.74 to −2.43 kcal/mol at I = 275 mM for E9-Im9, from −2.19 to −1.94 kcal/mol at I = 1000 mM for IL4-IL4BP, and from −1.66 to −1.42 kcal/mol at I = 500 mM for Fas-AChE. If the nonlinear PB equation was used instead of the linearized PB equation, the predicted rate of association for Bn-Bs would be almost unchanged, but would be reduced by 5-, 3-, and 7-fold, respectively, for E9-Im9, IL4-IL4BP, and Fas-AChE at the lowest ionic strengths. These reductions would bring predictions into much closer agreement with experimental data.

The values of k_a^0 obtained from the ionic strength dependences of k_a nearly fall in the expected range of 10^5–10^6 $M^{-1}s^{-1}$ (Northrup and Erickson, 1992; Zhou, 1997; Schlosshauer and Baker, 2004). These were used below for the mutants of the respective complexes without further adjustments. We also obtained values for k_a^0 by Brownian dynamics simulations, and these were 1.8×10^5, 3.4×10^4, 4.2×10^4, and 3.6×10^4 $M^{-1}s^{-1}$. The values required for optimal reproduction of the ionic strength effects on k_a were all underestimated by the Brownian dynamics simulation results. The underestimation is not unexpected, since in the Brownian dynamics simulations the proteins were treated as rigid, which could lead to an underestimate of the volume of the bound state and hence the basal association rate. Interestingly, in both sets of results, the basal rate for the Bn-Bs complex is roughly an order of magnitude higher than those of the other three complexes.

Association rates for 23 mutants of the 4 complexes were calculated because of availability of experimental data (Schreiber and Fersht, 1993, 1995, 1996; Shen et al., 1996; Radic et al., 1997; Wallis et al., 1998; Frisch et al., 2001). There were 12 mutant Bn-Bs complexes (mutations before and after the hyphen refer to those on Bn and Bs, respectively): K27A-Bs, R59A-Bs, E60A-Bs, R83Q-Bs, R87A-Bs, Bn-D35A, Bn-D39A, Bn-E76A, K27A-D39A, R59A-D35A, R83Q-D39A, and R87A-D39A. The E9-Im9 complex had five mutants (all on the Im9 subunit): E9-E30A, E9-E41A, E9-S50A, E9-D51A, and E9-Y55A. Two IL4-IL4BP mutants were E9Q-IL4BP and R88A-IL4BP, and four Fas-AChE mutants were Fas-D74N, Fas-E202Q, Fas-D280V, and Fas-D283Q. The mutated residues are shown in Figures 2A–2D. Except for the E202Q-Fas mutant, all mutations were in or around the interfaces of the complexes. Most of the mutations reduced the net charges of the wild-type proteins, but two actually increased the net charges (E60A-Bs and E9Q-IL4BP).

Figure 4 shows the comparison of predicted and experimental results on k_a for the 23 mutants of the 4 complexes. The mutants were studied over different ionic strengths. In all, there are 81 data points in the comparison. Over four orders of magnitude, the predictions closely track the

Figure 2. Electrostatic Potential Surfaces of the Subunits in the Four Protein Complexes Studied

(A) Bn-Bs.
(B) E9-Im9.
(C) IL4-IL4BP.
(D) Fas-AChE.

Positive and negative potentials are shown in blue and red, respectively. Mutated residues are labeled. Potential surfaces were generated by the APBS program (Baker et al., 2001) and were displayed by PyMOL (http://www.pymol.org).

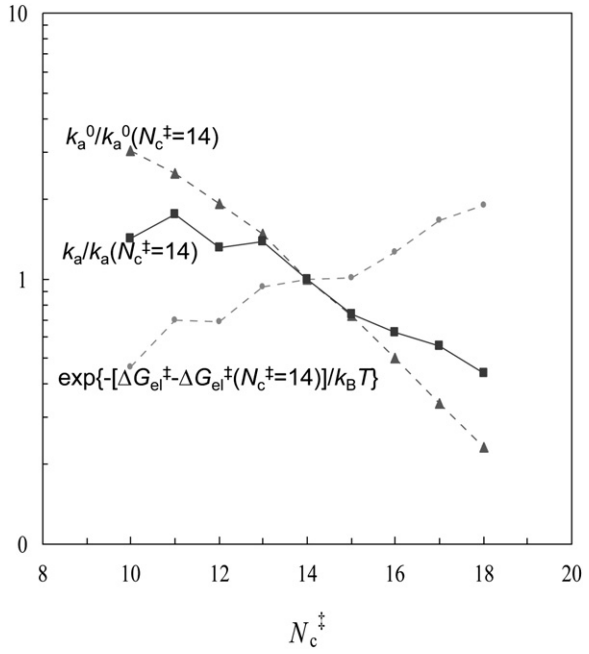


Figure 5. Insensitivity of the Predicted Rate to the Choice of the Transition-State Contact Level

Dependences of k_a^0, $\exp(-\Delta G_{el}^{\ddagger}/k_BT)$, and $k_a = k_a^0\exp(-\Delta G_{el}^{\ddagger}/k_BT)$ on $N_c^{\ddagger}$ for the Bn-Bs complex. Values are scaled by those at $N_c^{\ddagger} = 14$.

experimental results. However, for some of the Bn-Bs mutants, the rate of association was underestimated. Possible explanations for the underestimation will be discussed in the next section.

The sensitivity of the predicted rate of association to the choice of the transition-state contact level was examined by calculating $\Delta G_{el}^{\ddagger}$ and the basal rate, k_a^0, from Brownian dynamics simulations over a range of $N_c^{\ddagger}$ values. Figure 5 shows the variations of $\Delta G_{el}^{\ddagger}$ and k_a^0 for the Bn-Bs complex over the $N_c^{\ddagger}$ range of 10–18. Relative to the contact level of 14 determined by Equation 2, $\exp(-\Delta G_{el}^{\ddagger}/k_BT)$ decreased by 2.2-fold at $N_c^{\ddagger} = 10$ and increased by 1.9-fold at $N_c^{\ddagger} = 18$. The relatively small variation is a consequence of the long-range nature of electrostatic interactions. The variation of k_a^0 expectedly showed the opposite trend, increasing by 3.0-fold at $N_c^{\ddagger} = 10$ and decreasing by 4.3-fold at $N_c^{\ddagger} = 18$. Multiplying the two factors together (see Equation 1), the rate of association would be relatively constant, varying at most by a factor of 2 within the $N_c^{\ddagger}$ range of 10–18. Near constancy of k_a is generally expected since k_a^0 and $\exp(-\Delta G_{el}^{\ddagger}/k_BT)$ have opposite dependences on $N_c^{\ddagger}$.

DISCUSSION

We have shown that the transition-state theory predicts well association rates for diverse protein-protein complexes and a large number of their mutants over wide ranges of ionic strength. These results suggest that the theory will have wide applicability.

For three of the complexes studied, i.e., Bn-Bs, E9-Im9, and Fas-AChE, the need for electrostatic rate enhancement is clear, as the association occurs in situations in which speed is of the essence (Zhou, 2005a). Indeed, for Bn-Bs, rapid association is such a priority that the inhibitor Bs has a cluster of acidic residues that facilitate association with the nuclease Bs, even though the clustered charges reduce folding stability (Schreiber et al., 1994). Speed is also required in the blocking of ion channels by toxins (Terlau et al., 1996) and along the signaling pathway leading to the stimulation of actin polymerization. As actin polymerization is a nucleated event, the speed of upstream signaling has a critical impact on the rate of polymer formation. It is thus not surprising that high association rates have been observed between many partners along the signaling pathway, such as between Cdc42 and the Wiskott-Aldrich syndrome protein (WASp)—leading to the latter's activation—and between activated WASp and actin (Marchand et al., 2001; Hemsath et al., 2005). The high association rate between Cdc42 and WASp has been found to be essential for the latter to stimulate actin polymerization, as another Rho GTPase sharing 70% sequence identity, TC10, with an identical dissociation rate but a 1000-fold lower association rate, failed to stimulate actin polymerization (Hemsath et al., 2005). In all of these examples, the transition-state theory may be used to gain insight into the biological roles of protein association rates.

Several other compelling arguments can be made for the requirement of rapid protein association. First, high binding affinity is often achieved through slow dissociation. However, for proteins involved in signaling, slow dissociation is not an option, since a long-lasting bound state effectively corresponds to a permanent off- or on-switch. Therefore, even if not for a direct reason (such as in nucleation of actin polymerization), the affinity requirement alone calls for fast association. This may partly explain why the association between IL4 and IL4BP is electrostatically enhanced. Second, enzyme-substrate binding is a determining factor for the overall turnover rate and becomes the rate-limiting step for catalytically perfect enzymes. Third, when several proteins compete for the same receptor or when one protein is faced with alternative pathways, kinetic control, not thermodynamic control, dominates for much of the time; this is especially true when dissociation is slow. For example, newly synthesized proteins potentially face aggregation if not isolated by a chaperonin. From the point of view of kinetic control, it is easy to see why rapid binding of denatured proteins to GroEL has been observed (Perrett et al., 1997). Fourth, related proteins, such as Cdc42 and TC10, can have significantly different association rates, suggesting that association rate may serve as an additional mechanism for specificity. Much of the focus of studies on protein association has been on the binding affinity. The examples presented here suggest that binding rate deserves as much attention as binding affinity.

Electrostatic rate enhancement requires complementary charge distributions on the two associating molecules.

For some proteins, there are obvious reasons to take up positive or negative charges. For example, proteins targeting nucleic acids generally have enriched distributions of positively charged residues on the nucleic acid-binding sites (Luscombe et al., 2001; Tjong and Zhou, 2007). The positive charges on the nucleic acid-binding sites can be easily exploited by inhibitors (like Bs): fast inhibition can be achieved through a concentration of negatively charged residues on the latter molecules. Toxins blocking the *Shaker* potassium channel apparently follow a similar strategy (Huang et al., 2005). Apparently, for facilitating the conduction of the positively charged potassium ion, the mouth of the channel pore is lined with two rings of negative charge (Doyle et al., 1998). To complement the resulting negative electrostatic surface, channel toxins have excess positively charged residues. The charge-complementarity argument can also be made to rationalize the excess positive charges on Fas, which targets AChE. The latter uses a negative electrostatic surface around the entrance to the active site gorge for the fast binding of its positively charged substrate (Radic et al., 1997; Zhou et al., 1998). In other cases, such as in complexes formed by WASp and Cdc42 (Abdul-Manan et al., 1999) and by IL4 and its receptor (Hage et al., 1999), the reasons for a particular subunit to take up either positive or negative charges are not obvious. That they nonetheless show charge complementarity highlights the biological roles of their fast association.

The transition-state theory offers important advantages over Brownian dynamics simulations. First, the transition state is uniquely specified, so the theory holds truly predictive power. Second, the theory is rigorous and yet fast to implement. The need for expensive simulations of translational and rotational diffusion in the presence of intermolecular forces is eliminated.

Schreiber and coworkers have used Equation 1, but they calculated $\Delta G^{\ddagger}_{el}$ from an empirical energy function on the bound state, thus avoiding the specification of the transition state (Shaul and Schreiber, 2005). Though they seem to have had some success with the simplified approach, an approach with theoretical rigor is still more desirable. As we have shown here, the rigorous transition-state theory makes accurate predictions for the association rate.

Based on an expression similar to Equation 1, Miyashita et al. (2004) have used experimental data to derive the transition state for the association of cytochrome *c*2 with bacterial reaction center. Their experiment-based transition state appears to be similar to the transition states that we have determined from theory for the four complexes. In particular, the standard deviation of the rotation angle, χ, in their transition-state ensemble was 9°, which falls within the range of corresponding values that we found. Based on experimental results for mutational effects on association rates, we and others (Vijayakumar et al., 1998; Frisch et al., 2001) have suggested a transition state in which the two proteins are prealigned and solvent separated. The transition states determined here conform to this description.

Mechanistically, long-range electrostatic interactions bias configurational sampling around the transition state. Once reaching this state, the two proteins will undergo further conformational rearrangements to achieve stereospecific fit. Molecular dynamics simulations have shown that the formation of native contacts between two proteins is a very cooperative event (Huang et al., 2005), much like what is seen with protein folding.

The rates of association are predicted well for most of the 23 mutants studied here, but the rates for Bn-Bs mutations involving K27, R59, R83, and R87 of Bn and D39 of Bs are systematically underpredicted. These residues form an ion cluster across the interface (see Figure 2A). In the present study, modeling of mutation was restricted to the mutated residue. In reality, neighboring residues will respond to the mutation by conformational rearrangement. Such rearrangement will optimize interactions. For example, upon mutating the bulky side chain of K27 to Ala, neighboring R83 and R87 may adjust to make better contacts with D39 and D35 across the interface. The neglect of conformational rearrangement of neighboring residues in the present study appears to be the most likely reason for the underprediction of the Bn-Bs mutants. Conformational rearrangement can be modeled by molecular dynamics simulations (Huang et al., 2005) and will be considered in the future.

Like in earlier Brownian dynamics studies (Gabdoulline and Wade, 1997, 2001; Elcock et al., 1999), we have modeled the effects of salts totally in terms of the ionic strength, which appears in the Debye-Hückel screening parameter in the linearized PB equation. The effects of salts on biomolecular interactions are actually very complex. Broadly speaking, salt ions can exert influences by specifically binding to the biomolecules and by indirect means. We have no way of ruling out specific ion binding in the four protein complexes studied, but there is no particular reason to suspect that it plays a significant role either. One indirect means is through the redistribution around the proteins throughout the association process. The redistribution was modeled by the PB equation in the present study. Another indirect means gives rise to the Hofmeister effect, which itself is not well understood. According to an electrostatic theory (Kirkwood, 1943; Zhou, 2005b), the origin of the Hofmeister effect is the tendency of ions to stay away from protein surfaces in order to optimize hydration. Since the two proteins are solvent separated in our model for the transition state, we may expect the differential Hofmeister effect between the transition state and the unbound state not to be significant. The overall good agreement between the predicted and experimental dependences of the rate of association on salt concentration suggests that the (nonlinear) PB equation captures the bulk of salt effects on protein-protein association rates.

It is now increasingly recognized that proteins function in the context of multicomponent complexes. Binding rate likely plays as important roles as binding affinity in the proper functioning. The transition-state theory presents new opportunities for uncovering molecular bases of variations in binding rates among proteins and for designing proteins with desired binding rates.

THEORY AND IMPLEMENTATION

Specification of the Transition State

Both the theoretical basis and the explicit procedure for specifying the transition state have been presented in a recent paper (Alsallaq and Zhou, 2007). Here, we outline this procedure.

The transition state was determined by sampling configurations. The proteins were treated as rigid; therefore, there were only six relevant degrees of freedom: three for relative translation and three for relative rotation (Figure 1B). Configurations in the bound state and the transition region to the unbound state were sampled by randomly generating coordinates for the six degrees of freedom; the only restriction was that the separation, r, be within 6 Å. Configurations that did not involve steric collision between the two proteins were saved, and the contact number was then calculated (described below). For the purpose of detecting collision, atoms were classified into three types: hydrogen, polar (nitrogen and oxygen), and nonpolar (carbon and others). The collision distance within one type or between two types of atoms was set to the minimum distance of such contacts in the X-ray structure of the bound complex. The resulting values were 2.5–2.7 Å between polar atoms, 3.2–3.5 Å between nonpolar atoms, 2.8–3.1 Å between polar and nonpolar pairs, 1.6–2.1 Å between hydrogens, 1.6–1.7 Å between polar and hydrogen atoms, and 2.5 Å between nonpolar and hydrogen atoms.

Contacts, native or nonnative, were formed between interaction locus atoms. The latter were a representative set of interface atoms, which, in turn, were heavy atoms making crossinterface contacts below 5 Å in the X-ray structure of the complex. All crossinterface contacts were sorted in ascending order of contact distances. If a contact-forming atom was within 3.5 Å of an atom on the same protein that formed a shorter contact, the longer contact was eliminated from the list. The final remaining list constituted the cognate pairs of interaction locus atoms; the contact radius of each interaction locus atom was assigned to be half of the contact distance. When configurations of the proteins were sampled, a native contact was formed when an interaction locus atom was separated from its cognate partner by a distance that was not longer by 3.5 Å than the corresponding value in the X-ray structure. A nonnative contact was formed when a noncognate pair of interaction locus atoms were within a distance that was not longer by 2.5 Å than the sum of their contact radii.

For configurations at each contact level, N_c, the standard deviation in the rotation angle, χ, was calculated (Figure 1C). The transition state was specified by the contact level at which the function $\Xi(N_c)$ given in Equation 2 was a maximum.

The bound structures of the four protein complexes were from the following Protein Data Bank entries: Bn-Bs, 1brs (Buckle et al., 1994); E9-Im9, 1emv (Kuhlmann et al., 2000); IL4-IL4BP, 1iar (Hage et al., 1999); and Fas-AChE, 1mah (Bourne et al., 1995). Hydrogens were added and were energy minimized.

Calculation of the Electrostatic Interaction Free Energy

For each complex, the transition-state ensemble was represented by 100 configurations with the contact level at $N_c^{\ddagger}$. For each configuration, the electrostatic interaction energy was calculated as

$$U_{el} = U_{el}(AB) - U_{el}(A) - U_{el}(B). \quad (3)$$

The three terms on the right-hand side of the equation represent the electrostatic energies of the complex and each of the two proteins by itself, respectively. Electrostatic energies were calculated by the UHBD program (Madura et al., 1995); the boundary between the protein low dielectric and the solvent high dielectric was specified by the protein van der Waals surface. Other details of the electrostatic interactions can be found in a recent study of the effects of mutations and ionic strength on the binding affinities of the four protein complexes (Dong et al., 2003; Dong and Zhou, 2006). Unless otherwise indicated, electrostatic energies were calculated by solving the linearized PB equation. For comparison, the nonlinear PB equation was also solved (by adding the option "full" in the input script for the UHBD program).

The electrostatic interaction free energy, $\Delta G_{el}^{\ddagger}$, technically should be calculated from

$$\exp\left(-\Delta G_{el}^{\ddagger}/k_B T\right) = \langle \exp(-U_{el}/k_B T) \rangle^{\ddagger}, \quad (4)$$

where $\langle \ldots \rangle^{\ddagger}$ signifies averaging over the transition-state ensemble. However, for a finite sample of the transition-state configurations, such an averaging could be dominated by the configuration with the lowest interaction energy, which appears as a Boltzmann factor. As a compromise, we instead calculated $\Delta G_{el}^{\ddagger}$ simply as the average of the interaction energies in the transition-state configurations:

$$\Delta G_{el}^{\ddagger} = \langle U_{el} \rangle^{\ddagger}. \quad (5)$$

Determination of Basal Rate by Brownian Dynamics Simulations

The association rate in the absence of interaction force was obtained from Brownian dynamics simulations by using an algorithm developed previously (Zhou, 1993). Briefly, the associating proteins were randomly started from the bound state, defined by $N_c > N_c^{\ddagger}$ (the volume of the bound state is denoted as V_b subsequently). The same contact levels used for calculating the transition-state electrostatic interaction free energy were used for defining the reaction criteria. The trajectories were then propagated, with translation by the Ermak-McCammon algorithm (Ermak and McCammon, 1978) and rotation by an algorithm of Fernandes and de la Torre (2002). The relative translational diffusion constants of the protein complexes were estimated from the molecular weights of the subunits (Zhou, 1995). The values were 24, 23, 21, and 19 Å^2/ns, respectively, for Bn-Bs, E9-Im9, IL4-IL4BP, and Fas-AChE. Collisions between the two proteins

were treated like reflecting boundary conditions. While in the bound state, the proteins could form a complex with finite reactivity, κ. All trajectories were either terminated due to reaction or stopped after reaching a cutoff time.

The lengths of the trajectories were converted to the survival probability, $S(t)$, which equals the time-dependent rate, $k(t)$, scaled by the initial value $k(0) = V_b\kappa$. Extrapolation of $k(t)$ to infinite time gave the association rate, $k(\infty)$, obtained with finite reactivity. The diffusion-limited association rate, k_a^0, was obtained with the formula (Zhou and Szabo, 1996)

$$\frac{1}{k_a^0} = \frac{1}{k(0)} + \frac{1}{k(\infty)}. \quad (6)$$

Supplemental Data

Supplemental Data include $\Delta G_{el}^{\ddagger}$ and predicted and experimental results of k_a at various ionic strengths for the four protein complexes and are available at http://www.structure.org/cgi/content/full/15/2/215/DC1/.

ACKNOWLEDGMENTS

This work was supported in part by National Institutes of Health grant GM058187.

Received: October 24, 2006
Revised: December 27, 2006
Accepted: January 2, 2007
Published: February 13, 2007

REFERENCES

Abdul-Manan, N., Aghazadeh, B., Liu, G.A., Majumdar, A., Ouerfelli, O., Siminovitch, K.A., and Rosen, M.K. (1999). Structure of Cdc42 in complex with the GTPase-binding domain of the 'Wiskott-Aldrich syndrome' protein. Nature *399*, 379–383.

Alsallaq, R., and Zhou, H.-X. (2007). Energy landscape and transition state of protein-protein association. Biophys. J., in press.

Baker, D. (2000). A surprising simplicity to protein folding. Nature *405*, 39–42.

Baker, N.A., Sept, D., Joseph, S., Holst, M.J., and McCammon, J.A. (2001). Electrostatics of nanosystems: application to microtubules and the ribosome. Proc. Natl. Acad. Sci. USA *98*, 10037–10041.

Bourne, Y., Taylor, P., and Marchot, P. (1995). Acetylcholinesterase inhibition by fasciculin: crystal structure of the complex. Cell *83*, 503–512.

Bryngelson, J.D., and Wolynes, P.G. (1987). Spin-glasses and the statistical-mechanics of protein folding. Proc. Natl. Acad. Sci. USA *84*, 7524–7528.

Buckle, A.M., Schreiber, G., and Fersht, A.R. (1994). Protein-protein recognition: crystal structural analysis of a barnase-barstar complex at 2.0-Å resolution. Biochemistry *33*, 8878–8889.

Daggett, V., and Levitt, M. (1993). Protein unfolding pathways explored through molecular-dynamics simulations. J. Mol. Biol. *232*, 600–619.

Darling, R.J., Kuchibhotla, U., Glaesner, W., Micanovic, R., Witcher, D.R., and Beals, J.M. (2002). Glycosylation of erythropoietin affects receptor binding kinetics: role of electrostatic interactions. Biochemistry *41*, 14524–14531.

Dill, K.A., and Chan, H.S. (1997). From Levinthal to pathways to funnels. Nat. Struct. Biol. *4*, 10–19.

Dong, F., and Zhou, H.-X. (2006). Electrostatic contribution to the binding stability of protein-protein complexes. Proteins *65*, 87–102.

Dong, F., Vijayakumar, M., and Zhou, H.-X. (2003). Comparison of calculation and experiment implicates significant electrostatic contributions to the binding stability of barnase and barstar. Biophys. J. *85*, 49–60.

Doyle, D.A., Morais Cabral, J., Pfuetzner, R.A., Kuo, A., Gulbis, J.M., Cohen, S.L., Chait, B.T., and MacKinnon, R. (1998). The structure of the potassium channel: molecular basis of K^+ conduction and selectivity. Science *280*, 69–77.

Duan, Y., and Kollman, P.A. (1998). Pathways to a protein folding intermediate observed in a 1-microsecond simulation in aqueous solution. Science *282*, 740–744.

Elcock, A.H., Gabdoulline, R.R., Wade, R.C., and McCammon, J.A. (1999). Computer simulation of protein-protein association kinetics: acetylcholinesterase-fasciculin. J. Mol. Biol. *291*, 149–162.

Ermak, D.L., and McCammon, J.A. (1978). Brownian dynamics with hydrodynamic interactions. J. Chem. Physiol. *69*, 1352–1360.

Fernandes, M.X., and de la Torre, J.G. (2002). Brownian dynamics simulation of rigid particles of arbitrary shape in external fields. Biophys. J. *83*, 3039–3048.

Frisch, C., Fersht, A.R., and Schreiber, G. (2001). Experimental assignment of the structure of the transition state for the association of barnase and barstar. J. Mol. Biol. *308*, 69–77.

Gabdoulline, R.R., and Wade, R.C. (1997). Simulation of the diffusional association of barnase and barstar. Biophys. J. *72*, 1917–1929.

Gabdoulline, R.R., and Wade, R.C. (2001). Protein-protein association: investigation of factors influencing association rates by Brownian dynamics simulations. J. Mol. Biol. *306*, 1139–1155.

Garcia, A.E., and Sanbonmatsu, K.Y. (2001). Exploring the energy landscape of a β hairpin in explicit solvent. Proteins *42*, 345–354.

Haddadian, E.J., and Gross, E.L. (2006). A Brownian Dynamics study of the interactions of the luminal domains of the cytochrome b6f complex with plastocyanin and cytochrome c6: the effects of the Rieske FeS protein on the interactions. Biophys. J. *91*, 2589–2600.

Hage, T., Sebald, W., and Reinemer, P. (1999). Crystal structure of the interleukin-4/receptor α chain complex reveals a mosaic binding interface. Cell *97*, 271–281.

Hemsath, L., Dvorsky, R., Fiegen, D., Carlier, M.F., and Ahmadian, M.R. (2005). An electrostatic steering mechanism of Cdc42 recognition by Wiskott-Aldrich syndrome proteins. Mol. Cell *20*, 313–324.

Huang, X., and Zhou, H.-X. (2006). Similarity and difference in the unfolding of thermophilic and mesophilic cold shock proteins studied by molecular dynamics simulations. Biophys. J. *91*, 2451–2463.

Huang, X., Dong, F., and Zhou, H.-X. (2005). Electrostatic recognition and induced fit in the κ-PVIIA toxin binding to *Shaker* potassium channel. J. Am. Chem. Soc. *127*, 6836–6849.

Jucovic, M., and Hartley, R.W. (1996). Protein-protein interaction: a genetic selection for compensating mutations at the barnase-barstar interface. Proc. Natl. Acad. Sci. USA *93*, 2343–2347.

Kirkwood, J.G. (1943). The theoretical interpretation of the properties of solutions of dipolar ions. In Proteins, Amino Acids and Peptides as Ions and Dipolar Ions, E.J. Cohn and J.T. Edsall, eds. (New York: Reinhold Publishing Corporation), pp. 586–622.

Kuhlmann, U.C., Pommer, A.J., Moore, G.R., James, R., and Kleanthous, C. (2000). Specificity in protein-protein interactions: the structural basis for dual recognition in endonuclease colicin-immunity protein complexes. J. Mol. Biol. *301*, 1163–1178.

Lin, J., and Beratan, D.N. (2005). Simulation of electron transfer between cytochrome C2 and the bacterial photosynthetic reaction center: Brownian dynamics analysis of the native proteins and double mutants. J. Phys. Chem. B *109*, 7529–7534.

Luscombe, N.M., Laskowski, R.A., and Thornton, J.M. (2001). Amino acid-base interactions: a three-dimensional analysis of protein–DNA interactions at an atomic level. Nucleic Acids Res. *29*, 2860–2874.

Madura, J.D., Briggs, J.M., Wade, R.C., Davis, M.E., Luty, B.A., Ilin, A., Antosiewicz, J., Gilson, M.K., Bagheri, B., Scott, L.R., et al. (1995). Electrostatics and diffusion of molecules in solution: simulations with the University of Houston Brownian Dynamics program. Comput. Phys. Comm. *91*, 57–95.

Marchand, J.-B., Kaiser, D.A., Pollard, T.D., and Higgs, H.N. (2001). Interaction of WASP/Scar proteins with actin and vertebrate Arp2/3 complex. Nat. Cell Biol. *3*, 76–82.

Mayor, U., Guydosh, N.R., Johnson, C.M., Grossmann, J.G., Sato, S., Jas, G.S., Freund, S.M.V., Alonso, D.O.V., Daggett, V., and Fersht, A.R. (2003). The complete folding pathway of a protein from nanoseconds to microseconds. Nature *421*, 863–867.

Miyashita, O., Onuchic, J.N., and Okamura, M.Y. (2004). Transition state and encounter complex for fast association of cytochrome c2 with bacterial reaction center. Proc. Natl. Acad. Sci. USA *101*, 16174–16179.

Munoz, V., Thompson, P.A., Hofrichter, J., and Eaton, W.A. (1997). Folding dynamics and mechanism of β-hairpin formation. Nature *390*, 196–199.

Northrup, S.H., and Erickson, H.P. (1992). Kinetics of protein-protein association explained by Brownian dynamics computer simulation. Proc. Natl. Acad. Sci. USA *89*, 3338–3342.

Onuchic, J.N., Luthey-Schulten, Z., and Wolynes, P.G. (1997). Theory of protein folding: the energy landscape perspective. Annu. Rev. Phys. Chem. *48*, 545–600.

Perrett, S., Zahn, R., Stenberg, G., and Fersht, A.R. (1997). Importance of electrostatic interactions in the rapid binding of polypeptides to GroEL. J. Mol. Biol. *269*, 892–901.

Radic, Z., Kirchhoff, P.D., Quinn, D.M., McCammon, J.A., and Taylor, P. (1997). Electrostatic influence on the kinetics of ligand binding to acetylcholinesterase. J. Biol. Chem. *272*, 23265–23277.

Schlosshauer, M., and Baker, D. (2004). Realistic protein-protein association rates from a simple diffusional model neglecting long-range interactions, free energy barriers, and landscape ruggedness. Protein Sci. *13*, 1660–1669.

Schreiber, G., and Fersht, A.R. (1993). Interaction of barnase with its polypeptide inhibitor barstar studied by protein engineering. Biochemistry *32*, 5145–5150.

Schreiber, G., and Fersht, A.R. (1995). Energetics of protein-protein interactions: analysis of the barnase-barstar interface by single mutations and double mutant cycles. J. Mol. Biol. *248*, 478–486.

Schreiber, G., and Fersht, A.R. (1996). Rapid, electrostatically assisted association of proteins. Nat. Struct. Biol. *3*, 427–431.

Schreiber, G., Buckle, A.M., and Fersht, A.R. (1994). Stability and function: two constraints in the evolution of barstar and other proteins. Structure *2*, 945–951.

Shaul, Y., and Schreiber, G. (2005). Exploring the charge space of protein–protein association: a proteomic study. Proteins *60*, 341–352.

Shen, B.J., Hage, T., and Sebald, W. (1996). Global and local determinants for the kinetics of interleukin-4/interleukin-4 receptor α chain interaction. A biosensor study employing recombinant interleukin-4-binding protein. Eur. J. Biochem. *240*, 252–261.

Simmerling, C., Strockbine, B., and Roitberg, A.E. (2002). All-atom structure prediction and folding simulations of a stable protein. J. Am. Chem. Soc. *124*, 11258–11259.

Snow, C.D., Nguyen, N., Pande, V.S., and Gruebele, M. (2002). Absolute comparison of simulated and experimental protein-folding dynamics. Nature *420*, 102–106.

Spaar, A., Dammer, C., Gabdoulline, R.R., Wade, R.C., and Helms, V. (2006). Diffusional encounter of barnase and barstar. Biophys. J. *90*, 1913–1924.

Terlau, H., Shon, K.-J., Grilley, M., Stocker, M., Stuhmer, W., and Baldomero, O.M. (1996). Strategy for rapid immobilization of prey by a fish-hunting marine snail. Nature *381*, 148–151.

Tjong, H., and Zhou, H.X. (2007). Dros. Inf. Serv.PLAR: an accurate method for predicting DNA-binding sites on protein surfaces. Nucleic Acids Res., in press.

Vijayakumar, M., Wong, K.-Y., Schreiber, G., Fersht, A.R., Szabo, A., and Zhou, H.-X. (1998). Electrostatic enhancement of diffusion-controlled protein-protein association: comparison of theory and experiment on barnase and barstar. J. Mol. Biol. *278*, 1015–1024.

Wallis, R., Moore, G.K., James, R., and Kleanthous, C. (1995). Protein-protein interactions in colicin E9 DNase-immunity protein complexes. 1. Diffusion-controlled association and femtomolar binding for the cognate complex. Biochemistry *34*, 13743–13750.

Wallis, R., Leung, K.-Y., Osborne, M.J., James, R., Moore, G.R., and Kleanthous, C. (1998). Specificity in protein-protein recognition: conserved Im9 residues are the major determinations of stability in the colicin E9 Dnase-Im9 complex. Biochemistry *37*, 476–485.

Wells, J.A. (1996). Binding in the growth hormone receptor complex. Proc. Natl. Acad. Sci. USA *93*, 1–6.

Zhou, H.-X. (1993). Brownian dynamics study of the influences of electrostatic interaction and diffusion on protein-protein association kinetics. Biophys. J. *64*, 1711–1726.

Zhou, H.-X. (1995). Calculation of translational friction and intrinsic viscosity. II. Application to globular proteins. Biophys. J. *69*, 2298–2303.

Zhou, H.-X. (1996). Effect of interaction potentials in diffusion-influenced reactions with small reactive regions. J. Chem. Physiol. *105*, 7235–7237.

Zhou, H.-X. (1997). Enhancement of protein-protein association rate by interaction potential: accuracy of prediction based on local Boltzmann factor. Biophys. J. *73*, 2441–2445.

Zhou, H.-X. (2001). Disparate ionic-strength dependencies of on and off rates in protein–protein association. Biopolymers *59*, 427–433.

Zhou, H.-X. (2003). Association and dissociation kinetics of colicin E3 and immunity protein 3: convergence of theory and experiment. Protein Sci. *12*, 2379–2382.

Zhou, H.-X. (2005a). How do biomolecular systems speed up and regulate rates of processes? Phys. Biol. *2*, R1–R25.

Zhou, H.X. (2005b). Interactions of macromolecules with salt ions: an electrostatic theory for the Hofmeister effect. Proteins *61*, 69–78.

Zhou, H.-X., and Szabo, A. (1996). Theory and simulation of the time-dependent rate coefficients of diffusion-influenced reactions. Biophys. J. *71*, 2440–2457.

Zhou, H.X., Wong, K.Y., and Vijayakumar, M. (1997). Design of fast enzymes by optimizing interaction potential in active site. Proc. Natl. Acad. Sci. USA *94*, 12372–12377.

Zhou, H.-X., Wlodek, S.T., and McCammon, J.A. (1998). Conformation gating as a mechanism for enzyme specificity. Proc. Natl. Acad. Sci. USA *95*, 9280–9283.

Zou, G., and Skeel, R.D. (2003). Robust biased Brownian dynamics for rate constant calculation. Biophys. J. *85*, 2147–2157.

