J. Mol. Biol. (1998) **278**, 1015–1024

JMB

Electrostatic Enhancement of Diffusion-controlled Protein–Protein Association: Comparison of Theory and Experiment on Barnase and Barstar

M. Vijayakumar[1], Kwan-Yin Wong[1], Gideon Schreiber[2], Alan R. Fersht[3]* Attila Szabo[4]* and Huan-Xiang Zhou[1]*

[1]*Department of Biochemistry Hong Kong University of Science and Technology Clear Water Bay, Kowloon Hong Kong*

[2]*Department of Biochemistry Weizmann Institute of Science Rehovot 76100, Israel*

[3]*Cambridge Centre for Protein Engineering, Medical Research Council Centre, Hills Road Cambridge CB2 2QH, UK*

[4]*Laboratory of Chemical Physics, National Institute of Diabetes and Digestive Kidney Diseases, National Institutes of Health, Bethesda MD 20892, USA*

**Corresponding authors*

The electrostatic enhancement of the association rate of barnase and barstar is calculated using a transition-state theory like expression and atomic-detail modeling of the protein molecules. This expression predicts that the rate enhancement is simply the average Boltzmann factor in the region of configurational space where association occurs instantaneously in the diffusion-controlled limit. Based on experimental evidence, this ''transition state'' is defined by configurations in which, relative to the stereospecifically bound complex, the two proteins are shifted apart by ~8 Å (so a layer of water can be accommodated in the interface) and the two binding surfaces are rotated away by 0° to 3°. The values of the average Boltzmann factor, calculated by solving the Poisson–Boltzmann equation, for the wild-type complex and 16 complexes with single mutations are found to correlate well with experimental results for the electrostatic rate enhancement. The predicted rate enhancement is found to be somewhat insensitive to the precise definition of the transition state, due to the long-range nature of electrostatic interactions. The experimental ionic strength dependence of the rate enhancement is also reasonably reproduced.

Keywords: protein–protein association; electrostatic rate enhancement; average Boltzmann factor; barnase; barstar

Introduction

Before two proteins associate to form a complex, they have to find the proper relative orientation by translational and rotational diffusion. The orientational constraints will severely limit the rate of association. For example, for two proteins A and B modeled as spheres with closest approach distance R and relative translational diffusion constant D, ''reactive'' patches spanned by polar angles in (0, δ_A) and (0, δ_B) will reduce the diffusion-controlled association rate constant from the Smoluchowski result $4\pi DR$ (Smoluchowski, 1917) by a factor of $\delta_A\delta_B(\delta_A+\delta_B)/8$ (Berg, 1985; Zhou, 1993). For $\delta_A=\delta_B=3^\circ$, the resulting association rate constant is $\sim 10^5$ M^{-1} s^{-1}, which is roughly the magnitude of association rate constants observed for a number of protein complexes at extremely high ionic strengths (Stone *et al.*, 1989; Eltis *et al.*, 1991; Schreiber & Fersht, 1993, 1996; Wallis *et al.*, 1995; Wendt *et al.*, 1997). At low ionic strengths, these protein complexes are found to be formed three to four orders of magnitude faster, indicating the effect of electrostatic interactions between associating proteins. Recently, we found that electrostatic rate enhancement can be calculated from the expression (Zhou, 1996, 1997; Zhou *et al.*, 1996, 1997):

$$k_1/k_1^0 = \langle \exp(-\beta U) \rangle \qquad (1)$$

where $\beta = 1/k_BT$, U is the electrostatic interaction energy between the associating proteins, $\langle \cdots \rangle$ signifies averaging over the configurational space of the ''reactive'' region, and k_1 and k_1^0 are the association rate constants in the presence and absence of electrostatic interactions. In the diffusion-controlled limit, the reactive region becomes an absorbing surface, where association occurs instantaneously.

Present address: H.-X. Zhou and M. Vijayakumar, Department of Physics and Atmospheric Science, Drexel University, Philadelphia, PA 19104, USA.
E-mail: hxzhou@einstein.drexel.edu

0022–2836/98/201015–10 $25.00/0/mb981747

If this surface is viewed as defining a ''transition state'', then equation (1) superficially resembles the transition state theory. Here we apply this equation to account for the electrostatic enhancement of the association rate of barnase and barstar. Comparison to the extensive experimental data of Fersht and co-workers (Schreiber & Fersht, 1993, 1995, 1996; Schreiber *et al.*, 1997) will allow us to gain structural and mechanistic insight on the association of barnase and barstar.

It should be noted at the outset that there is some arbitrariness in the choice of the transition state. It is clear that this must consist of configurations with relative orientations near that of the final stereospecific complex. Without such severe orientational constraints, the rate constant at high ionic strengths will be calculated to be much too large. The experimental study by Schreiber & Fersht (1995) using double mutant cycles also indicates that in the transition state there is a distinct correlation between charged residues that have stereospecific contact in the bound complex. On the other hand, the interaction energies between these residues in the transition state are much smaller than those in the bound complex. This suggests that in the transition state the two proteins are separated by solvent.

Mechanistically, the associating proteins will probably first form much ''looser'' complexes, and these will then slowly be rearranged *via* translational and rotational diffusion to reach the orientationally constrained transition state. We will refer to these much looser complexes collectively as ''encounter complex.'' The transition state enters the specific calculation of the association rate constant by equation (1), whereas the encounter complex is a tool designed for a mechanistic interpretation of the association process. In any event, if the formation of the encounter complex is fast and the rearrangement into the transition state is slow then:

$$k_1 = K_{ec} k_{e\to t} \tag{2}$$

where K_{ec} is the equilibrium constant for forming the encounter complex from the separated proteins and $k_{e\to t}$ is the rate constant for rearranging the encounter complex into the transition state. When the interaction potential U is smooth over the configurational space of the encounter complex, K_{ec} will be proportional to the average Boltzmann factor $\langle \exp(-\beta U)\rangle$ in the transition state and $k_{e\to t}$ will be insensitive to the interaction potential. Then equation (2) is equivalent to equation (1). However, while this analysis does capture some of the key physics of diffusion-controlled protein–protein association, it is heuristic at best. The precise nature of the encounter complex is not specified. Moreover, the formalism cannot be used to actually calculate the numerical values of rate constants for well-defined microscopic models of associating proteins. Interestingly, Mrabet *et al.* (1986) used it to examine the effect of electrostatic interactions on the rate of hemoglobin αβ dimer assembly. On the basis of equation (2) and a description of the interaction energy between α and β chains in the encounter complex by Coulomb's law, they rationalized the difference in rates for normal and variant β chains.

To illustrate equation (1) and to introduce the model that we will use for calculating electrostatic interaction energies between two proteins, let us first consider a simplified version of the model. This consists of two spherical proteins (with radii R_A and R_B), each carrying a point charge (q_A or q_B) at the center. The proteins are assumed to be low dielectric regions (dielectric constant ε_p) and the electrostatic potential ϕ in the solvent arising from the protein charges is supposed to be governed by the Poisson–Boltzmann equation:

$$\nabla^2\phi - \kappa^2\phi = 0 \tag{3}$$

The Debye–Huckel screening parameter is given by $\kappa^2 = 8\pi N_A e^2 I/k_B T\varepsilon_s$, where N_A is Avogadro's number, e is the unit charge, I and ε_s are the solvent ionic strength and dielectric constant, respectively. Calculation of the electrostatic interaction energy U is complicated even for this simplified version of the model, involving an infinite series with coefficients specified by recursive relations (Zhou, 1993). For the purpose of illustration, let us assume that the asymptotic expression of U at large protein–protein separation r is applicable to an arbitrary r. This is given by (Debye, 1942):

$$U(r) = \frac{q_A q_B}{2\varepsilon_s}\left(\frac{e^{\kappa R_A}}{1+\kappa R_A} + \frac{e^{\kappa R_B}}{1+\kappa R_B}\right)\frac{e^{-\kappa r}}{r} \tag{4}$$

If in the transition state the two proteins are in contact, then the average Boltzmann factor is given by:

$$\langle \exp(-\beta U)\rangle = \exp\left[-\frac{q_A q_B}{2k_B T\varepsilon_s R}\left(\frac{e^{-\kappa R_B}}{1+\kappa R_A} + \frac{e^{-\kappa R_A}}{1+\kappa R_B}\right)\right] \tag{5}$$

where $R = R_A + R_B$. Equation (1) predicts:

$$\ln k_1 = \ln k_1(I=\infty) - \frac{q_A q_B}{2k_B T\varepsilon_s R}\left(\frac{e^{-\kappa R_B}}{1+\kappa R_A} + \frac{e^{-\kappa R_A}}{1+\kappa R_B}\right) \tag{6}$$

where we have identified k_1^0 with $k_1(I=\infty)$. This can be rewritten as:

$$\begin{aligned}\ln k_1 &= \ln k_1(I=0) \\ &\quad + \frac{q_A q_B}{2k_B T\varepsilon_s R}\left(2 - \frac{e^{\kappa R_B}}{1+\kappa R_A} - \frac{e^{-\kappa R_A}}{1+\kappa R_B}\right) \\ &\approx \ln k_1(I=0) + \frac{q_A q_B}{k_B T\varepsilon_s}\frac{\kappa}{1+\kappa R}\end{aligned} \tag{7}$$

to second order in κR. Equation (7) is essentially a formula used by Schreiber & Fersht (1996) to analyze their experimental data, but the above derivation makes it clear that R should be the sum

of the radii of the associating proteins rather than the radii of the electrolyte ions, as assumed originally.

In the diffusion-controlled limit, protein–protein association occurs instantaneously on a reactive surface. For the practical calculation of the association rate constant through Brownian dynamics simulations, it is more convenient to allow the association to occur at a finite rate rather than instantaneously in a region with finite volume. The diffusion-controlled rate constant can then be obtained by a limiting procedure, as we now illustrate. Suppose that association occurs with a rate constant k_{in} in the region bounded by the absorbing surface and the contact surface between the two proteins. The formation of the final stereospecific complex AB from two associating proteins A and B in separation can be described by the following kinetic scheme:

$$A + B \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} A \cdot B \xrightarrow{k_{in}} AB \tag{8}$$

where A·B represents complexes in the new reactive region with finite volume and k_{-1} is the diffusion-controlled rate constant for dissociating A·B. Now the overall association rate constant is:

$$k = \frac{k_{in}}{k_{-1} + k_{in}} k_1 \tag{9}$$

if A·B is in steady state. The ratio of the diffusion-controlled association and dissociation rate constants k_1 and k_{-1} is an equilibrium constant (given by Shoup & Szabo, 1982):

$$\frac{k_1}{k_{-1}} = V_{rr} \langle \exp(-\beta U) \rangle \tag{10}$$

where V_{rr} is the volume of the reactive region. This allows us to write:

$$\frac{1}{k} = \frac{1}{k_{in} V_{rr} \langle \exp(-\beta U) \rangle} + \frac{1}{k_1} \tag{11}$$

Hence from the association rate constant k for a finite value of k_{in}, one can obtain the diffusion-controlled association rate constant k_1. Interestingly, this analysis also provides an alternative way to rationalize equation (1). In analogy to $k_{e \to t}$, when the interaction potential is smooth around the reaction region, k_{-1} is expected to be insensitive to the interaction potential. Then according to equation (10), k_1 is proportional to the average Boltzmann factor in the reactive region. For a finite value of k_{in}, the association rate constant can be written as:

$$k = \frac{k_{-1} k_{in} V_{rr}}{k_{-1} + k_{in}} \langle \exp(-\beta U) \rangle \tag{12}$$

and is proportional to $\langle \exp(-\beta U) \rangle$ under the same conditions.

Results

Configurational space of the transition state

The transition state consists of configurations from which the stereospecific complex form instantaneously. These configurations were taken to be ones in which, relative to the stereospecific complex, the two proteins are shifted apart by a distance d and the two binding surfaces are rotated away by angles $\theta_{A,B} < \delta$. The value of d was fixed so that a layer of water can be accommodated in the interface, incorporating experimental evidence of solvent separation in the transition state (Schreiber & Fersht, 1995). The value of δ was fixed so that the experimental result for the association rate constant of barnase and barstar in the absence of electrostatic interactions (i.e. at infinite ionic strength) is reproduced. The procedure for the determination of d and δ and the generation of configurations in the transition state is described in Methods.

The center–center distance between barnase and barstar in the X-ray structure of the bound complex (see Figure 1) is $r_x = 23.4$ Å (Buckle *et al.*, 1994). Using $\delta = 3°$, the smallest shift distance that allows for one layer of solvent to be accommodated in the interface is $d = 8$ Å. Out of 200,000 configurations uniformly distributed in the region specified by center–center distance $r < r_x + d = 31.4$ Å and angles $\theta_{A,B} < \delta = 3°$, 31,761 have non-overlapping proteins. The volume of the reactive region is thus:

$$\begin{aligned} V_{rr} &= \frac{4\pi}{3}(31.4^3 - 23.4^3) \times \frac{1 - \cos 3°}{2} \\ &\quad \times \frac{1 - \cos 3°}{2} \times \frac{31,761}{200,000} \text{ Å}^3 \\ &= 5.67 \times 10^{-3} \text{ Å}^3 \end{aligned} \tag{13}$$

Using $k_{in} = 160$ ns^{-1}, the long-time limit of the survival fraction $S(t)$ of trajectories started in the configurational space of the reactive region is found to be 0.44. Equation (21) then gives $k_1^0 = 4.3 \times 10^5$ M^{-1} s^{-1}, which matches the experimental result for the association rate constant of barnase and barstar at infinite ionic strength (Schreiber & Fersht, 1996).

Electrostatic rate enhancement at $I = 0$

Using the above definition of the transition state, 20 configurations were randomly selected and the electrostatic interaction energy for each of them was calculated. From these the average Boltzmann factor was obtained. The results, expressed as the free energy of the transition state:

$$G^{\ddagger} \equiv -k_B T \ln \langle \exp(-\beta U) \rangle \tag{14}$$

at $I = 0$ for the wild-type barnase–barstar complex and 16 complexes with single mutations are listed in Table 1, along with the experimental

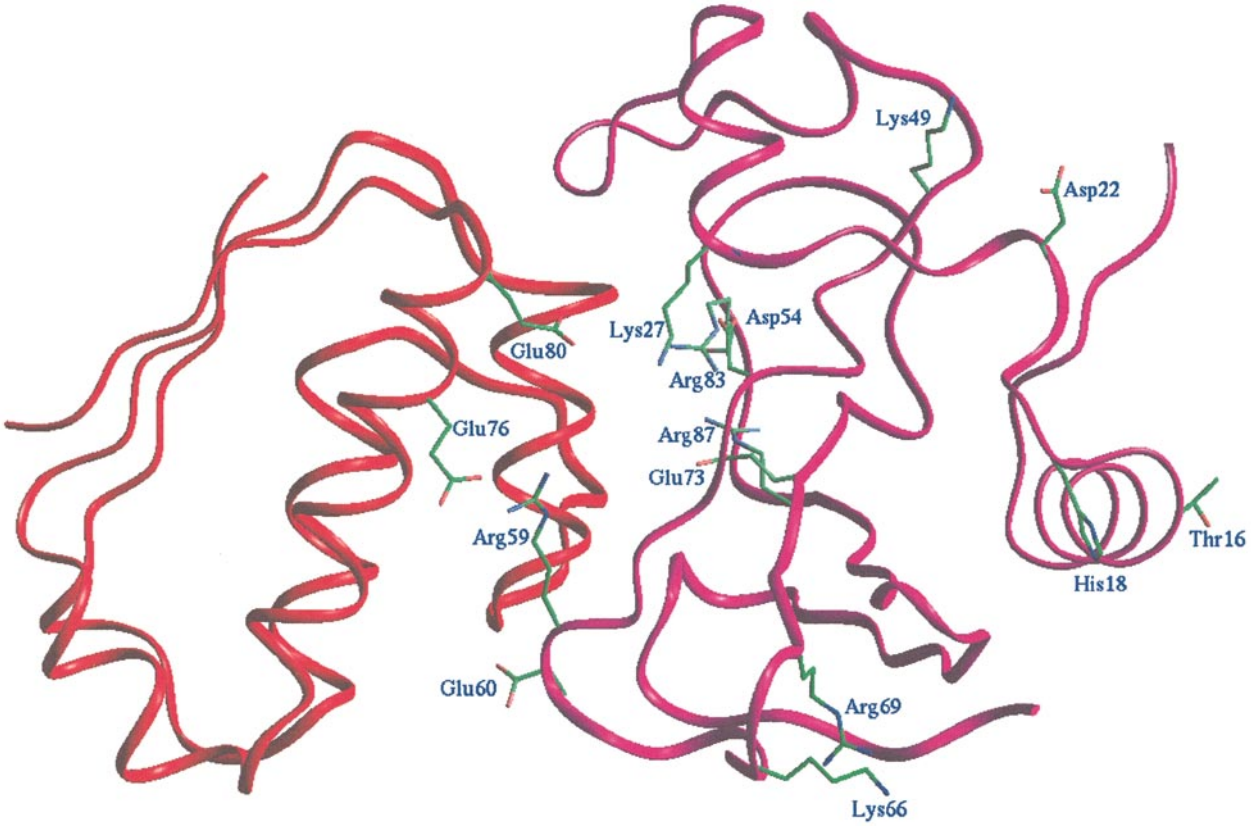

Figure 1. The X-ray structure of the bound complex of barnase (on the right and in pink) and barstar (in red). The side-chains of the 16 mutated residues studied here are also shown.

data for $k_1(I=0)$ and $k_1(I=\infty)$ from Fersht and co-workers (Schreiber & Fersht, 1993, 1995, 1996; Schreiber, Frisch & Fersht, unpublished results). The comparison of calculated results for $G^{\ddagger}(I=0)$ and experimental results for: $k_BT\ln[k_1(I=0)/k_1(I=\infty)]$, will put a severe test on the utility of equation (1). As Figure 2 shows, there is good agreement between calculation and experiment. In three cases (barnase D54A, E60A and E73W), the magnitudes of $G^{\ddagger}$ are overestimated by more than 1 kcal/mol. The largest error is ~25%. Possible reasons for the errors will be examined in the Discussion.

To check the sensitivity of the calculated electrostatic rate enhancement on the precise definition of the transition state, the wile-type barnase–barstar complex in a particular orientation was further shifted apart from a distance of 8 Å to 9 and 10 Å.

Table 1. Calculated and experimental results for electrostatic rate enhancement at $I=0$

Barnase	Barstar	$G^{\ddagger}(I=0)$ (kcal/mol)	$k_1(I=0)$ (10^9 M^{-1} s^{-1})	$k_1(I=\infty)$ (10^5 M^{-1} s^{-1})	$-k_BT\ln[k_1(I=0)/k_1(I=\infty)]$ (kcal/mol)
wt	wt	−7.0	7.2	1.2	−6.5
T16R	wt	−7.4	13	1.0	−7.0
H18D	wt	−6.5	11	0.7	−7.1
D22M	wt	−7.6	18	0.7	−7.4
K27A	wt	−4.7	0.5	1.6	−4.8
K49A	wt	−6.3	9.7	0.9	−6.9
D54A	wt	−8.6	32	0.7	−7.7
R59A	wt	−4.6	0.36	0.8	−5.0
E60A	wt	−9.9	65	1.2	−7.8
K66A	wt	−6.4	5.9	1.8	−6.2
R69S	wt	−6.2	6.5	1.2	−6.5
E73W	wt	−9.3	33	0.5	−7.9
R83A	wt	−4.7	1.1	4.0	−4.7
R83Q	wt	−4.5	1.0	3.6	−4.7
R87A	wt	−5.0	2.2	1.3	−5.8
wt	E76A	−6.6	2.2	1.3	−6.0
wt	E80A	−6.4	2.9	1.1	−5.5

wt, wild-type.

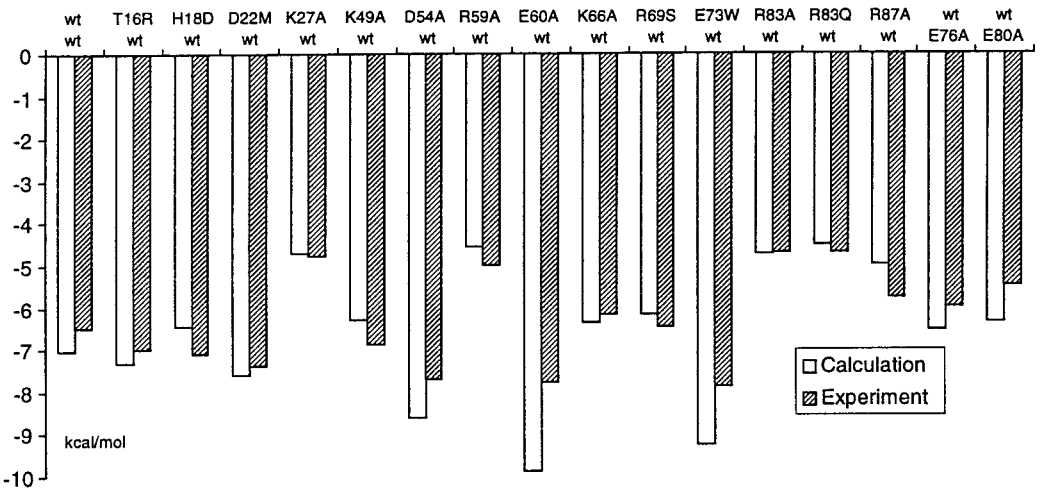

Figure 2. Comparison of calculated and experimental results for the free energy $G^{\ddagger}$ of the transition state. Top row refers to variants of barnase.

The calculated electrostatic interaction energy changed from −6.89 kcal/mol to −6.83 and −6.78 kcal/mol, respectively. This can be rationalized as a manifestation of the long-range nature of electrostatic interactions. This is very fortunate, for the accuracy of equation (1) relies on the smoothness of the interaction potential U around the configurational space of the transition state (Zhou, 1996; Zhou *et al.*, 1997).

Ionic strength dependence of electrostatic rate enhancement

The ionic strength dependence of the electrostatic rate enhancement is also reasonably reproduced. This can be illustrated by results on the wild-type complex. In Figure 3, calculated results for $-\beta[G^{\ddagger}(I) - G^{\ddagger}(I=0)]$ and experimental results for $\ln[k_1(I)/k_1(I=0)]$ as functions of $-(\beta e^2/2\varepsilon_s)\kappa/(1+\kappa a)$ are plotted. This is essentially the same way in which Schreiber & Fersht (1996) plotted their data, except for the division of $k_1(I)$ by the constant $k_1(I=0)$. As noted by Schreiber & Fersht, the experimental curve is fairly straight, with a slope of 17.6. There is a reasonable agreement between calculation and experiment for $I < 0.1$ M, i.e. $-(\beta e^2/2\varepsilon_s)/\kappa/(1+\kappa a) > -0.23$. At higher ionic strengths, the calculated results exhibit much weaker ionic strength dependence. Similar trends are also found for the 16 mutant complexes.

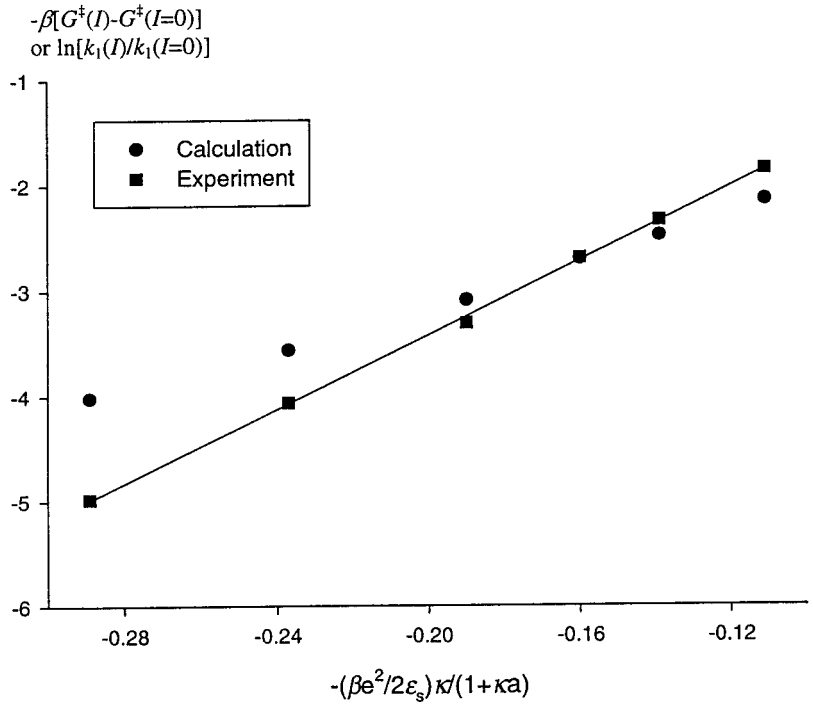

Figure 3. Comparison of calculated and experimental ionic strength dependence of the electrostatic rate enhancement for the wild-type barnase–barstar complex. For the abscissa $-(\beta e^2/2\varepsilon_s)\kappa/(1+\kappa a)$, the fitting parameter a has a value of 5.6 Å. The ionic strength increases as the magnitude of the abscissa increases.

Discussion

We have applied a recently derived expression, given in equation (1), to account for the electrostatic enhancement of the association rate of barnase and barstar. Experimentally observed effects of charge mutation and ionic strength are reasonably reproduced. A transition state for protein–protein association, from which the association occurs instantaneously, is identified explicitly. The magnitude of $\sim 10^5\ M^{-1}\ s^{-1}$ of the association rate constant at infinite ionic strength dictates that this must consist of configurations with relative orientations near that of the final stereospecific complex. This is further supported by the experimental observation that in the transition state there is distinct correlation between charged residues that have stereospecific contact in the bound complex. On the other hand, in the transition state the two proteins are not in direct contact but separated by solvent. In the context of the findings of the present study, we now discuss two previous studies of barnase–barstar association.

Work of Schreiber & Fersht (1996)

Schreiber & Fersht (1996) used the following formula to analyze the ionic strength dependence of the association rate constant:

$$\ln k_1 = \ln k_1(I=0) - \alpha \frac{e^2}{2k_B T \varepsilon_s} \frac{\kappa}{1+\kappa a} \qquad (15)$$

where α and a are fitting parameters. This is the same as equation (7) if $\alpha = -2q_A q_B/e^2$ and $a = R$. Remarkably, Schreiber & Fersht found that equation (15) reproduced their data for wild-type barnase and barstar over an extremely wide range of ionic strengths (up to 2 M NaCl) with $\alpha = 17.6$ and $a = 5.6$ Å (see Figure 3). Similar results were also found for mutant barnase–barstar complexes. We can rationalize this finding by noting that the charges that contribute most significantly to the average Boltzmann factor and hence electrostatic rate enhancement are the ones in the interfaces of the protein complexes constituting the transition state. Because of the small (i.e. 3°) relative rotations, these interfaces must consist of essentially the same atoms that consist of the interface of the stereospecific complex. We will thus base our discussion on the interface identified in the X-ray structure of the bound complex. Adding the charges of either protein in the interface of the bound complex, one has $q_{barnase} = +3e$ (from Lys27, Arg59, Glu60, Arg83 and Arg87) and $q_{barstar} = -3e$ (from Asp35, Asp39, Glu76). These give $\alpha = -2q_{barnase}q_{barstar}/e^2 = 18$, close to the fitted value of 17.6. Furthermore, the distances between these two groups of charges in the configurations constituting the transition state are ~7 Å, also close to the fitted value of $a = 5.6$ Å. Hence the fact that experimental data can be fitted to equation (15) with a small value of a (relative to the center–center separation of the proteins) lends further support to our proposed transition state. This consists of configurations with relative orientations near that of the final stereospecific complex but with relative separations increased so a layer of solvent can be accommodated in the interface.

It is perhaps fortuitous but nonetheless interesting that the above simplified electrostatic model (with appropriate interpretation of the parameters) is also able to rationalize mutational effects observed by Schreiber & Fersht (1995, 1996). For barnase mutants that preserve the charged residues Lys27, Arg59, Glu60, Arg83 and Arg87 in the interface with barstar, the values of the fitting parameter α are generally close to that for the wild-type complex. Neutralization of Lys27, Arg59, Arg83 or Arg87 is expected to decrease $q_{barnase}$ by $+e$ and α by 6, and the observed decrease in α is ~5.

Schreiber & Fersht (1995, 1996) designed double mutant cycle experiments to gain structural information on the transition state for barnase–barstar association. In terms of the measured association rate constants for the wild-type complex, the single mutant (X → A or Y → B) complexes, and the double mutant (X → A and Y → B) complexes, they defined the interaction energy between residues X and Y in the transition state as:

$$\Delta\Delta G^{\ddagger} = -k_B T \ln \frac{k_{1wt} k_{1X\to A, Y\to B}}{k_{1X\to A} k_{1Y\to B}} \qquad (16)$$

The maximum value of $-\Delta\Delta G^{\ddagger}$ was ~0.5 kcal/mol (for mutations in the interface of the X-ray complex). In contrast, the analogous quantity $-\Delta\Delta G_{eq}$ obtained from the equilibrium constants for forming the stereospecific complex could be as large as 6 kcal/mol. Partly on the basis of this observation, it was suggested that the transition state was an early, weakly specific complex, which was dominated by non-specific long-range electrostatic interactions. On the other hand, from equation (1), we have specifically identified the transition state by the configurations from which the stereospecific complex forms instantaneously. We have taken these configurations to be the ones in which the two proteins are orientationally constrained though solvent separated. We now argue using the simplified electrostatic model that this is also consistent with the double-mutant results of Schreiber & Fersht. Suppose that the mutation X → A changes the binding interface charge on barnase from q_A to $q_A + \delta_A$ and the mutation Y → B changes the binding interface charge on barstar from q_B to $q_B + \delta_B$. From equations (5) and (14), one finds that at zero ionic strength the interaction energy between the two residues

in our version of the transition state is:

$$\Delta\Delta G^{\ddagger} = (G^{\ddagger}_{X\to A,Y\to B} - G^{\ddagger}_{wt}) - (G^{\ddagger}_{X\to A} - G^{\ddagger}_{wt}) - (G^{\ddagger}_{Y\to B} - G^{\ddagger}_{wt}) = \frac{\delta_A \delta_B}{k_B T \varepsilon_s R} \quad (17)$$

If either mutation occurs away from the interface, then δ_A or $\delta_B = 0$ and $\Delta\Delta G^{\ddagger} = 0$. The maximum value of $-\Delta\Delta G^{\ddagger}$ is obtained when both mutations occur in the interface. Then, e.g., $-\delta_A = \delta_B = e$ and $-\Delta\Delta G^{\ddagger} = e^2/\varepsilon_s R$. This predicts that $-\Delta\Delta G^{\ddagger} = 0.6$ kcal/mol using $\varepsilon_s = 78.5$ and $R = 7$ Å, which is appropriate for the distance between the interface charges on the two proteins in our transition state. Note that the solvent dielectric constant enters equation (17), consistent with the fact that the two proteins in the transition state, though close to each other, are separated by solvent. An analogous calculation for $\Delta\Delta G_{eq}$ would not only involve a smaller R but, more importantly, a much smaller dielectric constant ε_p for the protein medium because of solvent exclusion from the interface of the stereospecific complex. Thus the fact that $-\Delta\Delta G^{\ddagger}$ is much smaller than $-\Delta\Delta G_{eq}$ is not unexpected even when the orientationally constrained but solvent-separated configurations are taken to be the transition state.

Work of Gabdoulline & Wade (1997)

Gabdoulline & Wade (1997) have carried out extensive Brownian dynamics simulations of the association between barnase and barstar. One of the focuses of this work is also the definition of the reactive surface. The results are qualitatively similar to those found in the present paper. There are several technical differences. In the Gabdoulline–Wade work, the electrostatic interaction energy was calculated using an effective charge model derived from the Poisson–Boltzmann equation (Gabdoulline & Wade, 1996), whereas we directly used the Poisson–Boltzmann equation. The effective charge model, though much improved over the test charge model, appears to still substantially underestimate the electrostatic interaction energy. For example, for one particular configuration of the wild-type barnase–barstar complex, solution of the Poisson–Boltzmann equation gave −6.0 kcal/mol for the interaction energy between the proteins. The result using the effective charge model was −3.8 kcal/mol, which is only moderately better than the result, −3.0 kcal/mol, calculated by the test charge model. Consistent with this, the electrostatic rate enhancement obtained by Gabdoulline & Wade is significantly smaller. At $I = 0.05$ M, their enhancement was 450-fold, which corresponds to $G^{\ddagger} = -3.6$ kcal/mol. In comparison, our calculation using the Poisson–Boltzmann equation gave $G^{\ddagger} = -5.2$ kcal/mol. It appears that the accuracy of the effective charge model requires further assessment.

Gabdoulline & Wade obtained the association rate constant from Brownian dynamics simulations, whereas we have obtained the association rate constant *via* calculating the average Boltzmann factor and applying equation (1). Incidentally, their algorithm for obtaining the association rate constant is different from the one we described here for calculating k_1^0 and causes difficulty when calculating small rate constants (such as k_1^0). The configurational space of the reactive surface in the Gabdoulline–Wade work is somewhat larger than ours, as indicated by the tenfold higher value for k_1^0 (10^6 M^{-1} s^{-1} compared to our calculated and experimental value of 10^5 M^{-1} s^{-1}). This is compensated by their underestimation of the electrostatic interaction energy, so the magnitude of the electrostatically enhanced rate constant k_1 obtained in that work is comparable to ours.

Finally, let us discuss the possible reasons for the errors of our calculations (see Table 1). The first is the neglect of polarization effect (i.e. the charges used in the electrostatic calculations were derived from small fragments of the proteins in gas phase). This tends to overestimate the magnitude of the electrostatic interaction energy (York *et al.*, 1996). Perhaps a more significant reason, as noted by Gabdoulline & Wade, is the neglect of structural flexibility. Fixing the side-chains in the conformations that occur in the X-ray structure of the bound complex will favour electrostatic enhancement and thus lead to overestimation of the magnitude of the electrostatic interaction energy.

The deviation between calculation and experiment is increased at high ionic strengths (around 0.2 M; see Figure 3). This perhaps is an indication of the breakdown of the Poisson–Boltzmann equation. It should be noted that using the full rather than linearized Poisson–Boltzmann equation will not alleviate this problem. It has been shown that the electrostatic interaction energies calculated using the two versions of the Poisson–Boltzmann equations are essentially the same, as long as the proteins are treated as low dielectric regions (Zhou, 1994). Gabdoulline & Wade have drawn the same conclusion.

In summary, we suggest that the transition state for diffusion-controlled protein–protein association, from which the stereospecific complex forms instantaneously, consists of configurations with three features. First, the relative orientations of the associating proteins are constrained. Second, the proteins are separated by a layer of solvent. Third, the side-chain of the proteins are flexible.

Methods

Electrostatic model

The model we used to calculate the electrostatic interaction energy between barnase and barstar is more general than the spherical-protein version described in the

Introduction, in two respects. First, the protein surfaces are defined with atomic-level details using the structures of the proteins in the X-ray structure of the complex (Buckle *et al.*, 1994). Second, each protein atom has a partial charge. The sum of the partial charges for an Asp or Glu residue or each C terminus is $-e$, whereas the sum of partial charges for an arginine or lysine residue or each N terminal is $+e$. All other residues have zero net charges. The total charges of wild-type barnase and barstar are $+2e$ and $-6e$, respectively.

Every protein atom was included in the electrostatic calculations. The heavy atoms missing in the X-ray structure of the barnase–barstar complex were added by model building using the InsightII program from Molecular Simulations, Inc. Hydrogen atoms were then added by energy minimization while keeping the heavy atoms fixed. Mutations of residues were also generated by model building. The mutations studied here are shown in Figure 1.

Solution of the Poisson–Boltzmann equation

The electrostatic interaction energy between barnase and barstar in a given relative configuration is calculated as:

$$U = \frac{1}{2}\sum_{\text{barnase-barstar}} q_i\phi_i - \frac{1}{2}\sum_{\text{barnase}} q_i\phi_i - \frac{1}{2}\sum_{\text{barstar}} q_i\phi_i \qquad (18)$$

where q_i are the partial charges of the protein atoms and ϕ_i are the electrostatic potential at the atomic sites for a protein specified under the summation sign. The electrostatic potential is obtained by the program UHBD (Madura *et al.*, 1995), in which the Poisson–Boltzmann equation is solved by finite difference. The protein and solvent dielectric constants are $\varepsilon_p = 4$ and $\varepsilon_s = 78.5$.

For either barnase or barstar, we used a $60 \times 60 \times 60$ grid with a spacing of 1.5 Å followed by a $140 \times 140 \times 140$ grid with a spacing of 0.35 Å for focusing. This was judged satisfactory by calculating the electrostatic energy of barnase with the grids imposed in two different orientations relative to the protein. For a barnase–barstar complex, the dimensions of these grids are too small. Instead we used a $100 \times 100 \times 100$ grid wit a spacing of 1.5 Å followed first by a $140 \times 140 \times 140$ grid with an intermediate spacing of 0.5 Å and then by a $60 \times 60 \times 60$ grid with a fine spacing of 0.25 Å. The last grid was centered on the N atom of the mutated residue. For the wild-type barnase–barstar complex, the fine grid was centered on the N atom of barnase Arg59. To reduce errors arising from distributing charges to a cubic grid, ϕ_i values were calculated as $\phi_i - \phi_i^{\text{ref}} + \Sigma_{j \neq i}\, q_j/\varepsilon_p r_{ij}$, where ϕ_i^{ref} is the finite-difference solution of ϕ_i for the reference state with $\varepsilon_p = \varepsilon_s = 4$ and $\kappa = 0$.

Specification of the transition state

The binding surface on barnase (barstar) was defined by those atoms that are closer than 4 Å from any atom of barstar (barnase) in the X-ray structure of the bound complex. The atoms on the two binding surfaces were combined to form a set of interface atoms. The total number of interface atoms is 304. A plane that passes through the geometric center of these atoms and gives the least sum of squared distances to the atoms was found. The axis that passes through the geometric center of the interface atoms and along the normal of the least-squares plane is called the main axis. For either protein, a point on the main axis that has the mean coordinate along the axis was chosen as the center. The vector from the center of either protein to the geometric center of the interface atoms is called its main vector. Roughly speaking, the main vector points from the center of the protein to the center of its binding surface. Both proteins were then allowed to reorient by rotating their main vectors away as well as rotating around the main vectors. The protein centers were also allowed to move away from each other. The vector pointing from the center of barnase (barstar) to the center of barstar (barnase) was defined as its polar axis. In total, six variables are required to specify a relative configuration of the protein complex: the polar angle θ_A of, and rotation angle ψ_A around, the main vector of protein A, the polar angle θ_B and azimuthal angle ϕ_B of, and rotation and ψ_B around, the main vector of protein B, and the separation r between the two protein centers (Figure 4). The azimuthal angle ϕ_A is unnecessary because a rotation of the complex as a whole does not change the relative configuration of the two proteins.

Based on the magnitude of $10^5\ M^{-1}\ s^{-1}$ of the barnase–barstar association rate constant at infinite ionic strength (Schreiber & Fersht, 1996), a provisional value of $\delta = 3^\circ$ was taken. To determine the shift distance d, trial values starting at 5 Å and with a 1 Å increment were tested sequentially. At each trial value d_t, configurations with angles $\theta_{A,B}$ restricted to δ (and ψ_A, ϕ_B and ψ_B unrestricted) and d restricted to ± 0.5 Å of d_t were generated. For each configuration, 3 Å (roughly the diameter

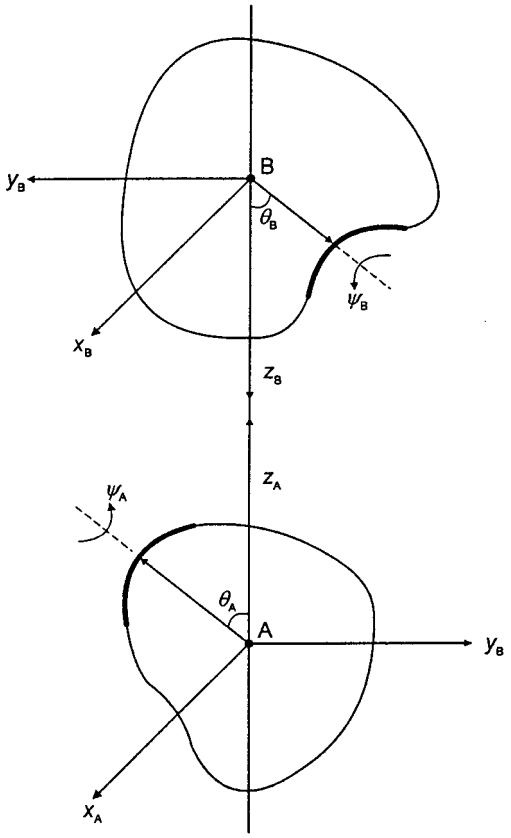

Figure 4. Variables that define the relative configurations of two proteins.

of a water molecule) was added to the van der Waals radii of the atoms of barnase (barstar) and the atoms with inflated radii were checked against barstar (barnase) for overlap. A configuration is said to allow for one layer of water in the interface if the total number of atoms that cause overlap after the 3 Å inflation of radii is less than 20% of the number of interface atoms in the X-ray complex. The smallest d_t value that allowed for 90% of the corresponding configurations to accommodate one layer of water was taken to be the value of d in the transition state.

If the value of k_1^0 calculated for the transition state thus defined is different from the experiment result, the values of δ and d have to be adjusted accordingly.

Brownian dynamics simulation of k_1^0

To find the absolute value of the association rate constant k_1 at a given ionic strength from equation (1), we need the association rate constant k_1^0 in the absence of interaction between the associating proteins (i.e. k_1 at infinite ionic strength). This was obtained from Brownian dynamics simulations by using a previously developed algorithm (Zhou, 1990, 1993; Zhou & Szabo, 1996) and taking the reactive region to be defined by configurations with $\theta_{A,B} < \delta$ and $r < r_x + d$, where r_x is the center–center separation in the X-ray structure of the bound complex. According to the algorithm, the two proteins are started uniformly in the reactive region. This was implemented by sampling 200,000 configurations randomly distributed within the allowed ranges of the six variables: $\theta_{A,B}$, ϕ_B, $\psi_{A,B}$ and r, and keeping only those that do not have protein–protein overlap. Within the reactive region, the proteins can form the stereospecific complex with a first-order rate constant k_{in} (see equation (8)). The Brownian dynamics trajectories are terminated either when the stereospecific complex is formed or when a cutoff time has been reached. Along the way the survival fraction $S(t)$ of the trajectories at time t is recorded. The association rate constant k^0 is given by the long-time limit of $k_{in}V_{rr}S(t)$. This is obtained by fitting the long-time portion of the simulated $S(t)$ to its analytical asymptotic behavior (Zhou & Szabo, 1996):

$$S(t) = \frac{k^0}{k_{in}V_{rr}}\left(1 + \frac{k^0}{4\pi D}\frac{1}{\sqrt{\pi D t}}\right) \tag{19}$$

The relative translational diffusion constant for barnase and barstar is taken to be 30 Å^2/ns, and the rotational diffusion constants for the two proteins are 0.04 and 0.045 ns^{-1}, respectively. These values are the same as those used by Gabdoulline & Wade (1977) in their Brownian dynamics simulation study of barnase–barstar association.

When $k_{in} \to \infty$, the association process is diffusion-controlled. The corresponding association rate constant k_1^0 can be obtained from k^0 *via* the approximate relation (Zhou & Szabo, 1996):

$$\frac{1}{k^0} = \frac{1}{k_{in}V_{rr}} + \frac{1}{k_1^0} \tag{20}$$

which is equivalent to equation (11). The diffusion-controlled association rate constant k_1^0 is what is relevant for the barnase–barstar complex (Schreiber & Fersht, 1996). In terms of $S(\infty)$, one has:

$$k_1^0 = k_{in}V_{rr}\frac{S(\infty)}{1 - S(\infty)} \tag{21}$$

During the Brownian dynamics simulation, the protein surfaces are reflecting to each other. This condition was treated by moving back to the original configuration whenever a step leads to overlapping between the proteins (Zhou, 1990). To accelerate the detection of overlapping, atoms on the surface of each protein are found using a method described by Zhou (1995). Only the surface atoms are used in checking for overlapping between the proteins. This reduces the number of atoms that need to be considered from 1727 to 952 for barnase and from 1434 to 738 for barstar.

Acknowledgements

The authors thank J. A. McCammon for providing a copy of the UHBD program and R. C. Wade for providing a copy of the program implementing the effective charge model. This work was supported in part by grant HKUST 638/96 M from the Research Grants Council of Hong Kong.

References

Berg, O. G. (1985). Orientation constraints in diffusion-limited macromolecular association. The role of surface diffusion as a rate-enhancing mechanism. *Biophys J.* **47**, 1–14.

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

Debye, P. (1942). Reaction rates in ionic solutions. *Trans. Electrochem. Soc.* **82**, 265–272.

Eltis, L. D., Herbert, R. G., Barker, P. D., Mauk, A. G. & Northup, S. H. (1991). Reduction of horse ferricytochrome *c* by bovine liver ferrocytochrome b_5. Experimental and theoretical analysis. *Biochemistry*, **20**, 3663–3674.

Gabdoulline, R. R. & Wade, R. C. (1996). Effective charges for macromolecules in solvent. *J. Phys. Chem.* **100**, 3868–3878.

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

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. & McCammon, J. A. (1995). Electrostatics and diffusion of molecules in solution: simulations with the University of Houston Brownian Dynamics program. *Comp. Phys. Comm.* **91**, 57–95.

Mrabet, N. T., McDonald, M. J., Turci, S., Sarkar, R., Szabo, A. & Bunn, H. F. (1986). Electrostatic attraction governs the dimer assembly of human hemoglobin. *J. Biol. Chem.* **261**, 5222–5228.

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

Schreiber, G. & 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. & Fersht, A. R. (1996). Rapid, electrostatically assisted association of proteins. *Nature Struct. Biol.* **3**, 427–431.
Schreiber, G., Frisch, C. & Fersht, A. R. (1997). The role of Glu73 of barnase in catalysis and the binding of barstar. *J. Mol. Biol.* **270**, 111–122.
Shoup, D. & Szabo, A. (1982). Role of diffusion in ligand binding to macromolecules and cell-bound receptors. *Biophys. J.* **40**, 33–39.
Smoluchowski, M. V. (1917). Versuch einer mathematischen theorie der koagulationskinetik kolloider losungen. *Z. Phys. Chem.* **92**, 129–168.
Stone, R. S., Dennis, S. & Hofsteenge, J. (1989). Quantitative evaluation of the contribution of ionic interactions to the formation of the thrombin-hirudin complex. *Biochemistry,* **28**, 6857–6863.
Wallis, R., Moore, G. R., James, R. & 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.
Wendt, H., Leder, L., Harma, H., Jelesarov, I., Baici, A. & Bosshard, H. R. (1997). Very rapid, ionic strength-dependent association and folding of a heterodimeric leucine zipper. *Biochemistry,* **36**, 204–213.
York, D. M., Lee, T.-S. & Yang, W. (1996). Quantum mechanical study of aqueous polarization effects on biological macromolecules. *J. Am. Chem. Soc.* **118**, 10940–10941.
Zhou, H.-X. (1990). Kinetics of diffusion-influenced reactions studied by Brownian dynamics. *J. Phys. Chem.* **94**, 8794–8800.
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. (1994). Macromolecular electrostatic energy within the nonlinear Poisson-Boltzmann equation. *J. Chem. Phys.* **100**, 3152–3162.
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. Phys.* **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.* **13**, 2441–2445.
Zhou, H.-X. & Szabo, A. (1996). Theory and simulation of the time-dependent rate coefficients of diffusion-influenced reactions. *Biophys. J.* **71**, 2440–2457.
Zhou, H.-X., Briggs, J. M. & McCammon, J. A. (1996). A 240-fold electrostatic rate-enhancement for acetylcholinesterase-substrate binding can be predicted by the potential within the active site. *J. Am. Chem. Soc.* **118**, 13069–13070.
Zhou, H.-X., Wong, W.-Y. & Vijayakumar, M. (1997). Design of fast enzymes by optimizing interaction potential in active site. *Proc. Natl Acad. Sci. USA,* **94**, 12372–12377.

Edited by B. Honig

(Received 16 October 1997; accepted 20 February 1998)