

Kinetics of Desolvation-Mediated Protein–Protein Binding

Carlos J. Camacho, S. R. Kimura, Charles DeLisi, and Sandor Vajda

Department of Biomedical Engineering, Boston University, Boston, Massachusetts 02215, USA

ABSTRACT The role of desolvation in protein binding kinetics is investigated using Brownian dynamics simulations in complexes in which the electrostatic interactions are relatively weak. We find that partial desolvation, modeled by a short-range atomic contact potential, is not only a major contributor to the binding free energy but also substantially increases the diffusion-limited rate for complexes in which long-range electrostatics is weak. This rate enhancement is mostly due to weakly specific pathways leading to a low free-energy attractor, i.e., a precursor state before docking. For α -chymotrypsin and human leukocyte elastase, both interacting with turkey ovomucoid third domain, we find that the forward rate constant associated with a collision within a solid angle φ around their corresponding attractor approaches 10^7 and 10^6 $\text{M}^{-1}\text{s}^{-1}$, respectively, in the limit $\varphi \sim 2^\circ$. Because these estimates agree well with experiments, we conclude that the final bound conformation must be preceded by a small set of well-defined diffusion-accessible precursor states. The inclusion of the otherwise repulsive desolvation interaction also explains the lack of aggregation in proteins by restricting nonspecific association times to ~ 4 ns. Under the same reaction conditions but without short range forces, the association rate would be only $\sim 10^3$ $\text{M}^{-1}\text{s}^{-1}$. Although desolvation increases these rates by three orders of magnitude, desolvation-mediated association is still at least 100-fold slower than the electrostatically assisted binding in complexes such as barnase and barstar.

INTRODUCTION

Diffusion of the reactants is frequently the rate-limiting step in the association of two proteins. The maximum rate constant for this process, 10^9 – 10^{10} $\text{M}^{-1}\text{s}^{-1}$, is given by the Smoluchowski equation that describes the collision rate for two uniformly reactive spherical molecules in solution (Smoluchowski, 1917). A successful reaction between two proteins must meet the additional constraint that small reactive patches on a particular face of each protein are properly aligned. The probability of satisfying this condition in a random collision is very small, suggesting that the reaction rates should be several orders of magnitude lower than the diffusion-limited collision rates. Nevertheless, it is not uncommon to find reaction rates as high or higher than 10^9 $\text{M}^{-1}\text{s}^{-1}$. Although this at first seems puzzling, analysis indicates that long-range electrostatic effects can heavily bias the approach of the molecules to favor reactive conditions. This effect was shown to be important for many association processes, including those of proteins with DNA (von Hippel and Berg, 1986), proteins with highly charged small molecules (Sharp et al., 1987), and proteins with oppositely charged protein substrates (Stone et al., 1989; Eltis et al., 1991; Schreiber and Ferscht, 1996; Gabdouliline and Wade, 1997; Vijayakumar et al., 1998). These systems have been thoroughly studied, and are frequently regarded as typical examples of binding phenomena.

Electrostatics is clearly not the only force that can affect the association rate. In addition to electrostatics, the most

important process contributing to the binding free energy is desolvation, i.e., the removal of solvent both from nonpolar (hydrophobic) and polar atoms (Chothia and Janin, 1975). It is generally accepted that partial desolvation is always a significant contribution to the free energy in protein–protein association, and it becomes dominant for complexes in which the long-range electrostatic interactions are weak (Camacho et al., 1999). In this paper, we perform Brownian dynamics simulations to study the effects of desolvation on the rates of diffusion-limited protein–protein association.

Brownian dynamics treats each protein as a rigid body, generally a sphere, and the solvent as a viscous Newtonian liquid (Ermak and McCammon, 1978; DeLisi, 1980; Northrup et al., 1984; Northrup and Erickson, 1992; Luty et al., 1993). The method has led to a number of important results. In particular, Gabdouliline and Wade (1997) have modeled the association of the barnase–barstar complex, and found that long-range electrostatic forces alone can reconcile the high rates (10^9 $\text{M}^{-1}\text{s}^{-1}$) observed by Schreiber and Ferscht (1996). Short-range interactions have been considered by Northrup and Erickson (1992), who have shown that a short-range locking potential can increase the association rate from 1×10^5 $\text{M}^{-1}\text{s}^{-1}$ to 2×10^6 $\text{M}^{-1}\text{s}^{-1}$, but did not attempt to explain the physical origin of this potential.

To establish the expected magnitude of desolvation effects on the association rates, we first perform simulations in which the interacting proteins are described by a simple model that assumes a hydrophobic interaction uniformly distributed over the entire protein surface. Results show that such short-range nonspecific interactions can significantly enhance the diffusion entrapment and thus increase the association rate. However, they also yield lengthy collisions and large nonspecific affinities that are rarely seen in real

Received for publication 30 April 1999 and in final form 7 December 1999.

Send reprint requests to Carlos J. Camacho, Department of Biomedical Engineering, Boston University, 44 Cummington St., Boston, MA 02215. Tel.: 617-353-4842; Fax: 617-353-6766; E-mail: ccamacho@bu.edu.

© 2000 by the Biophysical Society

0006-3495/00/03/1094/12 \$2.00

proteins (Northrup and Erickson, 1992). Therefore, we continue the analysis using a more realistic interaction potential that includes both atomic-level desolvation and short-range electrostatics.

The central assumption of the paper is that desolvation is a major contributor to the binding free energy of proteins. The magnitude of desolvation free energy is relatively well established on the basis of the free energy of transferring small molecules from water into organic solvents (Eisenberg and McLachlan, 1986; Vajda et al., 1994). We model this contribution using a structure-based atomic contact potential that has been independently validated by comparing it to various thermodynamic data, and has been shown to provide values of the desolvation free energy of proteins with remarkable accuracy (Zhang et al., 1997). Thus, there is little doubt that the desolvation force is real and, for the first time, it is included in a Brownian dynamic simulation.

Simulations are performed for two complexes in which turkey ovomucoid third domain (OMTKY) binds to α -chymotrypsin and human leukocyte elastase. OMTKY has zero net charge, and, although it has higher order electrostatic multipoles, both systems represent the class of complexes in which the long-range electrostatic interactions are weak, but, under normal conditions, the binding process is still diffusion limited (Camacho et al., 1999). Typical binding rates for this type of complexes are on the order of 10^5 – 10^7 M⁻¹s⁻¹, about 100-fold lower than for complexes that are electrostatically steered toward the binding pocket. The simulations identify weakly specific pathways leading to a low free energy attractor of well-oriented encounter complexes, embedded in an otherwise repulsive environment. Due to these attractors, the desolvation can increase the association rate by several orders of magnitude.

As it is always the case in Brownian dynamics, the calculated absolute rates also depend on the reaction condition (Gabdouline and Wade, 1997), and this may reduce the value of the method as a predictive tool. In our study, the simulations provide important information on the reaction condition itself. It is shown that, accounting for desolvation, the calculated and observed association rates agree if and only if the reaction condition is defined as a small ensemble of diffusion-accessible encounter complexes, suggesting that the final bound conformation is preceded by an almost unique precursor state. This interesting observation will be discussed further in the paper.

METHODS

Reference coordinates and reaction condition

The receptor and the ligand are treated as spheres of radii R_r and R_l diffusing in a viscous liquid. The coordinate systems XYZ and xyz are fixed to the receptor and ligand, respectively, and their origins coincide with the corresponding centers of mass (see Fig. 1 A). The Euler angles θ and ϕ define the vector pointing to the center of the ligand, i.e., to the origin of the coordinate system xyz . Three further Euler angles θ , ϕ , and ψ

determine the relative orientation of the ligand axes xyz in the reference coordinate system XYZ .

The calculation of the association rate requires a reaction condition that defines the last stage of the diffusion process, after which binding would be expected to be certain. The reaction condition we use is given in terms of a reactive patch around an optimal state on each protein surface, determined by a particular position and orientation of the two molecules in contact. This orientation is given by five angles θ_0 , ϕ_0 , θ_{10} , ϕ_{10} , and ψ_{10} , where the first two angles determine the position of the ligand's center of mass, and the last three specify the optimal relative orientation of the ligand denoted by the axes x_0 , y_0 , and z_0 in Fig. 1 A. According to this condition, a simulation trajectory leads to receptor-ligand association if the receptor and ligand collide, and, at the time of the collision, the two molecules are oriented in such a way that the solid angle φ around the direction (θ_0, ϕ_0) in Fig. 1 A is less than a threshold φ_r , and, similarly, each axis of the coordinate system xyz is rotated by less than a threshold φ_l from the optimal axes x_0 , y_0 , and z_0 . The relationship of this surface patch to the various reaction criteria used by other groups in Brownian dynamics simulations will be discussed further in the paper.

Brownian dynamics

The transport properties of proteins are calculated by assuming a spherical shape. Although structural asymmetries are known to yield anisotropic diffusion, it is generally accepted that, for globular proteins, these corrections should be small, and hence the spherical approximation and the resulting isotropic diffusion constants are appropriate. The translational and rotational diffusion coefficients for a sphere of radius R are given by the Stokes–Einstein relations $D^{\text{trans}} = k_B T / 6\pi\eta R$ and $D^{\text{rot}} = k_B T / 8\pi\eta R^3$, where k_B is Boltzmann constant, T is temperature and η is solvent viscosity.

According to the Brownian dynamics algorithm developed by Ermak and McCammon (1978), the time evolution of the relative displacement, Δr , between the reactants centers of mass of the reacting molecules is given by

$$\Delta r = D\Delta t F / k_B T + S, \quad (1)$$

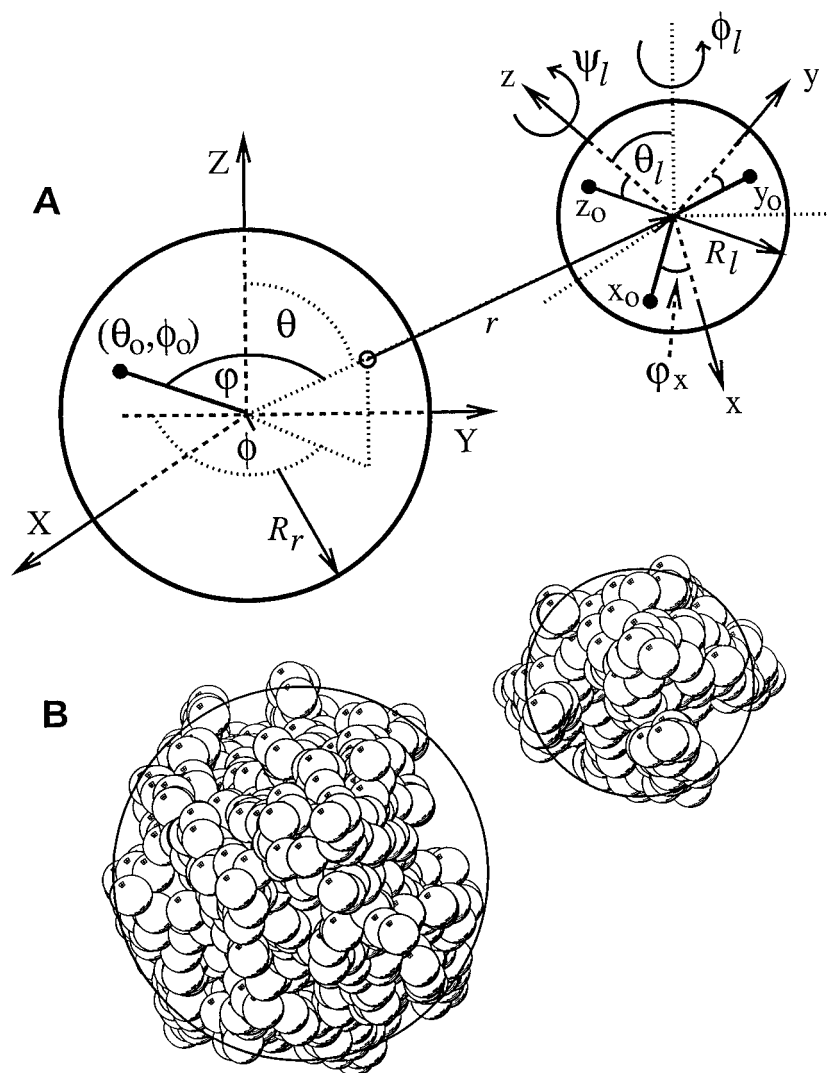
where Δt is the time step, F is the interparticle force, and $D = D_r^{\text{trans}} + D_l^{\text{trans}}$ denotes the sum of the unimolecular diffusion constants. S is a stochastic component of the displacement arising from random collisions of particles with solvent molecules, and is generated by taking normally distributed random numbers obeying the relationship $\langle S_k^2 \rangle = 2D\Delta t$. A similar expression governs the independent rotational Brownian motion. For each molecule α , ($\alpha = r, l$), the angular change $\Delta\Phi^\alpha$ around each of the orthogonal axes is given by

$$\Delta\Phi^\alpha = D_\alpha^{\text{rot}} \Delta t K^\alpha / k_B T + W^\alpha, \quad (2)$$

where K^α is the total torque on protein α , and W is the stochastic term such that $\langle (W_k^\alpha)^2 \rangle = 2D_\alpha^{\text{rot}} \Delta t$. The time step Δt decreases monotonically from 160 ps at 500-Å separation to 0.5 ps within the desolvation layer. If a given time step leads to a protein overlap, instead of voiding the move altogether (as in Gabdouline and Wade, 1997), we rescale Δt to avoid the overlap.

Throughout this paper, we study the binding of turkey ovomucoid third domain to α -chymotrypsin or human leukocyte elastase. The effective radius of these molecules shown in Fig. 1 B is estimated based on the solvent-accessible surface area (Lee and Richards, 1971) of the free molecules, calculated with a water radius of 0.6 Å. The radii of chymotrypsin and leukocyte elastase differ by less than 2%, thus, for simplicity, we assume the same radius $R_r = 23.24$ Å for both molecules. The effective radius of turkey ovomucoid third domain is $R_l = 14.15$ Å. The sum $R_r + R_l$ also agrees with the average center-to-center distance of the diffusion-accessible encounter complexes. At the temperature $T = 25^\circ\text{C}$ and viscosity coefficient $\eta = 1$ cP, the diffusion constants are $D_r^{\text{trans}} = 0.0094$ Å²/ps, $D_l^{\text{trans}} = 0.0154$ Å²/ps, $D_r^{\text{rot}} = 0.0000131$ rad²/ps, and $D_l^{\text{rot}} = 0.0000578$ rad²/ps.

FIGURE 1 (A) XYZ and xyz are coordinate systems fixed to the receptor and ligand centers. The position of the origin of xyz with respect to XYZ is defined by the angles (θ, ϕ) and the center-to-center distance r . $(\theta_l, \phi_l, \psi_l)$ are the Euler angles describing the ligand orientation. The optimal relative orientation of receptor and ligand is given by the angles (θ_0, ϕ_0) and by the coordinate system $x_0y_0z_0$. Notice that (θ_0, ϕ_0) defines a vector $\mathbf{r}_0$ pointing to the center of the ligand, and the coordinate axes x_0, y_0 , and z_0 define the orientation of the ligand. The reaction condition is defined as a patch around this optimal point where the solid angles $\varphi, \varphi_x, \varphi_y$, and φ_z measure the deviations from the angle (θ_0, ϕ_0) (i.e., from the vector $\mathbf{r}_0$, and from the coordinate axes x_0, y_0 , and z_0 , respectively). (B) The space-filling models show the van der Waals surfaces of the receptor (α -chymotrypsin) and that of the ligand (OMTKY) to illustrate that these surfaces can be relatively well approximated by spheres. The overall shape and size of the second receptor considered in this work (human leukocyte elastase) is very similar to that of α -chymotrypsin, and hence both molecules will be represented by spheres of effective radius $R_r = 23.24$ Å. The effective radius of turkey ovomucoid third domain is $R_l = 14.15$ Å.



Determination of the association rate

Brownian dynamics trajectories start with the ligand randomly placed and oriented on the surface of a sphere of radius $R_0 = 60$ Å in the XYZ coordinate system. The 60-Å threshold has been chosen because it is larger than the range of any intermolecular interaction. Independent trajectories are run until they either meet the reaction condition, or leave the finite diffusion space defined by a larger concentric sphere of radius $R_\infty = 500$ Å. The corresponding association rate constant is obtained from the equation (Northrup et al., 1984)

$$k_{\text{on}} = 4\pi DR_0\beta/[1 - (1 - \beta)R_0/R_\infty], \quad (3)$$

where β is the fraction of trajectories that satisfy the reaction condition. We note that the denominator corrects for those trajectories that reenter the diffusion space at R_0 having left at R_∞ .

Forces and torques

Let U denote the potential that determines the receptor–ligand interactions in a particular model of the receptor–ligand system. The potential depends on distance r and five angles $\theta, \phi, \theta_l, \phi_l$, and ψ_l , i.e., $U = U(r; \theta, \phi, \theta_l, \phi_l, \psi_l)$. Forces and torques are defined as translational and rotational deriva-

tives of U . Using finite differences, the force component along an axis, say, X , is calculated by

$$F_X = -[U(r + aX; \theta, \phi, \theta_l, \phi_l, \psi_l) - U(r - aX; \theta, \phi, \theta_l, \phi_l, \psi_l)]/2a, \quad (4)$$

where $a = 0.05$ Å. The torque on the ligand K_X^l is a function of the same variables, and is calculated by

$$K_X^l = -[U(r; \theta, \phi, \theta_l, \phi_l, \psi_l) - U(r; \theta, \phi, \mathcal{R}_X^{\Delta\phi}\{\theta_l, \phi_l, \psi_l\})]/\Delta\phi, \quad (5)$$

where $\mathcal{R}_X^{\Delta\phi}$ denotes the operator that rotates the ligand axes by $\Delta\phi = 0.02$ rad around its center of mass along the vector X . The torque on the receptor is computed by using conservation of angular momentum.

Interaction potential

We will use the binding free energy of the receptor–ligand system as the interaction potential $U \equiv \Delta G$. Let G_0 denote the free energy in the unbound

state. The free energy difference $\Delta G = G - G_0$ is calculated by

$$\Delta G = \Delta E_{\text{coul}} + \Delta G_{\text{des}}, \quad (6)$$

where ΔE_{coul} and ΔG_{des} denote the direct electrostatic (Coulombic) and desolvation contributions, respectively. In this work, we use two different desolvation models.

Model I: Uniform desolvation

We set $\Delta E_{\text{coul}} = 0$ and assume that the desolvation contribution is proportional to the change ΔA in the solvent-accessible surface area, i.e., $\Delta E_{\text{des}} = \lambda \Delta A$, where λ is an atomic solvation parameter. Notice that this relationship reflects the classical description of hydrophobicity (Hermann, 1972; Wolfenden et al., 1981). On the basis of the free energy of transfer between an organic liquid and water, λ is close to zero for polar atoms, and its value has been estimated to be in the range of 16–33 cal/mol/Å² for nonpolar atoms (Eisenberg and McLachlan, 1986; Vajda et al., 1994).

The actual reach of the hydrophobic and desolvation effects is not fully established, but most evidence seems to indicate that it must be at least one or two water layers from the protein surface (Israelachvili and Wennerstrom, 1996). In other words, this is the distance at which the steric hindrance of the first layers of water molecules become relevant. A similar model, with attractive forces acting only in a boundary layer, has been studied by Zhou (1979), but he assumed that the forces are due to van der Waals rather than hydrophobic interactions.

Model II: Atomic-level interaction potential

The electrostatic term is given by the expression $\Delta E_{\text{coul}} = \sum_{i=1}^n \Phi_i q_i$, where q_i is the charge of atom i of the ligand, Φ_i is the electrostatic potential of the solvated receptor at the position of the same atom, and n is the number of ligand atoms. We use a semi-Coulombic approximation, i.e., the potential Φ is calculated for the receptor (dielectric 2) by solving the linearized Poisson–Boltzmann equation by a finite difference method. The solvent dielectric constant is set to 40, empirically accounting for the effects of the low dielectric cavity of the ligand on the potential of the receptor. The charges and partial charges of the ligand are not changed. We have shown that the effective solvent dielectric of 40 accounts for the long-range electrostatic energy, smoothly extrapolating this energy to the partially desolvated interface.

The important feature of Model II is an atomic-level description of desolvation using the structure-based atomic contact energy (ACE) developed by Zhang et al. (1997). We have devoted substantial efforts to show that the calculated desolvation free energies are consistent with the available thermodynamic data (Zhang et al., 1997). In particular, it was shown that the free energy of solvating amino-acid side chains obtained by this method correlates to a high degree ($r = 0.975$) with the experimentally determined free energies of transferring the side chains between water and octanol.

Spherical approximation

As shown in Fig. 1 B, we model the van der Waals surface of the proteins using spheres. Geometric effects are included, up to a certain degree, by precalculating the force fields using an all-atom description of the proteins prior to the spherical approximation. Although this approach does not reflect the fine details of steric complementarity, it is consistent with the overall accuracy of this type of simulations. Indeed, Brownian dynamics already assumes hard walls and rigid body diffusion, and the hydrodynamic parameters are calculated for spheres. Notice that the excluded solvent-accessible surface areas of typical encounter complexes are on the order of 300–500 Å² (Camacho et al., 1999), similar to the excluded surface area between two interacting spheres of the effective radii shown in Fig. 1. This

implies that, in an encounter complex, there is neither significant shape complementarity nor steric conflict between the two proteins, at least nothing close to the one found in the fully formed complex where the desolvated interface is 1400–1600 Å². We recognize that the spherical approximation affects the length of the path as the two molecules approach each other. However, although the reaction rates are independent of any local entanglement of the model proteins (i.e., we do not account for hydrodynamics), a small increase or decrease of some of the pathways by ± 3 Å should barely have an impact on our numerical results.

Free energy landscapes

The basic idea of Brownian dynamics using free energy landscapes is precalculating the potential for a large set of interacting ligand–receptor pairs in various orientations using all-atom protein models, projecting this free energy landscape onto the surface of spheres representing the proteins, extending the surface potential to the whole space, and then using it for the calculation of forces and torques in the simulation. The construction of free energy landscapes of encounter complexes has been reported elsewhere (Camacho et al., 1999). Briefly, the receptor and the ligand are placed in the coordinate system such that the orientation of the native complex is given by $(\theta, \phi; \theta_l, \phi_l, \psi_l) = (90^\circ, 90^\circ; 0^\circ, 0^\circ, 0^\circ)$. Encounter complexes are generated by sampling the six-dimensional space of ligand translations and rotations, and setting the surface-to-surface distance $d_{\text{S-S}} = 0$. The five degrees of freedom $\theta, \phi, \theta_l, \phi_l$, and ψ_l are first sampled at every 20°, where the θ angles vary between 0° and 180°, and the others vary between 0° and 360°. A finer angular grid is used around the region of low free energy (already identified as the binding region in Camacho et al., 1999). Namely, for $70^\circ \leq \theta, \phi \leq 110^\circ$ the sampling is every 5°.

We note that a vector of the form $(\theta, \phi; \theta_l, \phi_l, \psi_l)$ specifies both an encounter complex in the all-atom representation, and its spherical approximation. We calculate the electrostatic energy ΔE_{coul} and desolvation term ΔG_{ACE} for each encounter complex, and assign the resulting values to the contact points on the surfaces of both spheres. The potentials are extended to the whole surface by using a standard linear interpolation based on the ten nearest neighbor sites in the 5-dimensional space $(\theta, \phi; \theta_l, \phi_l, \psi_l)$. As shown in Fig. 2, this restricted sampling yields a reliable estimate of the interaction energy for the whole receptor–ligand interface.

Crystal structures

Due to the induced fit, the conformations of proteins in a complex can differ from the conformations in their monomeric states. For the computation of the free energy landscapes, we use the bound (co-crystallized) conformations of the two proteins, rather than their unbound, (separately crystallized) forms. The bound conformations have been selected to account for the fact that, in our analysis, all interactions are restricted to surface-to-surface separation of less than 4.2 Å (see below). Under these conditions, the bound conformation is likely to be a better approximation than the unbound one, which assumes that the molecules do not interact at all. Using the unbound rather than the bound conformations could slightly change the free energy landscape and hence the calculated association rates. However, in the systems studied here, the main source of interaction is the desolvation free energy ΔG_{ACE} , described by a smoothly varying contact potential that is much less affected by conformational changes than electrostatics. In addition, as we will show, a change in the calculated rates by as much as 50% could be easily compensated by a small increase in the solid angle of the reaction condition, without affecting our main results. Therefore, we conclude that the difference between bound and unbound conformations is not a major concern in the present work.

Three-dimensional mapping

So far, the components ΔE_{coul} and ΔG_{ACE} of the interaction potential have been defined only for encounter complexes, i.e., for surface-to-surface

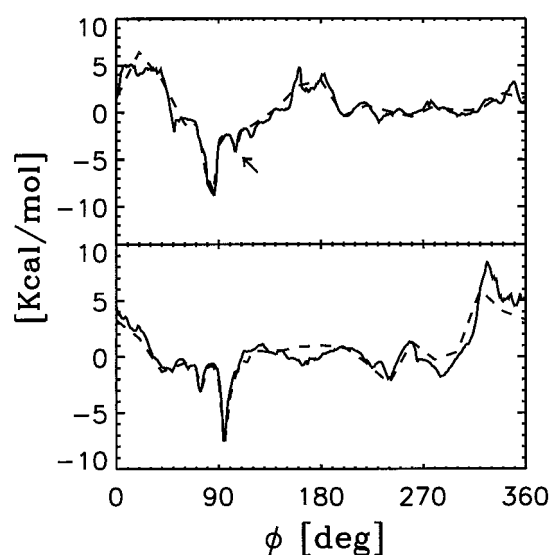


FIGURE 2 Profiles of the diffusion-accessible interaction potential of α -chymotrypsin (*top*) and human leukocyte elastase (*bottom*) with OMTKY as a function of ϕ . The profiles include the low free energy attractor for both complexes. The solid lines show the potential calculated at every 1° , whereas the dashed lines correspond to the potential as used in the Brownian dynamics simulations, i.e., calculated on a coarse grid and then extended by interpolation onto the entire surface. *Top*: $\theta = 85^\circ$ and $(\theta_i, \phi_i, \psi_i) = (20^\circ, 180^\circ, 180^\circ)$. In the reaction condition, the reactive patch is centered at $\phi = 85^\circ$. The arrow indicates a nearby minimum for which we also compute the association rate. *Bottom*: $\theta = 80^\circ$ and $(\theta_i, \phi_i, \psi_i) = (20^\circ, 300^\circ, 40^\circ)$. The reaction condition is centered at $\phi = 95^\circ$.

distance $d_{S-S} = 0$. For Brownian dynamics simulations, the potential must be defined over the entire space, and thus we need to study the radial dependence of the free energy components, thereby extending the potentials to $d_{S-S} > 0$. Figure 3 shows some typical profiles of the desolvation and electrostatic terms as functions of d_{S-S} .

The ACE-based desolvation, ΔG_{ACE} , depends almost linearly on d_{S-S} , and vanishes at ~ 4.2 Å. It should be noted that the solvent-accessible area excluded at close proximity also depends linearly on d_{S-S} . Indeed, the solvent-excluded area for two regular spheres is a linear function of d_{S-S} . Hence, in what follows, we compute an effective interaction potential at an arbitrary d_{S-S} from the equation,

$$\Delta G_{des}(d_{S-S}) = \begin{cases} \Delta G_{des}(0)(4.2 - d_{S-S})/4.2, & \text{if } d_{S-S} < 4.2 \text{ Å} \\ 0, & \text{otherwise,} \end{cases} \quad (7)$$

where $\Delta G_{des}(0)$ denotes the desolvation free energy on the molecular surface, i.e., at $d_{S-S} = 0$. As shown in Fig. 3, the linear approximation reflects the general behavior of the desolvation energy very well.

To a first approximation, the electrostatic potential can also be represented by a linear function. In particular, if $|\Delta E_{coul}| \leq 4$ Kcal/mol at $d_{S-S} = 0$, which is the case for the overwhelming number of encounter conformations (more than 99.7% of those sampled), the long-range tail of the interaction is < 1 Kcal/mol. For the small number of encounter complexes for which the electrostatic interaction is larger than 4 Kcal/mol, the desolvation interaction is usually repulsive. For example, in Fig. 3 we plot the electrostatic and desolvation energy for the encounter pairs with the highest and lowest electrostatic interactions as they are moved apart along the vector connecting their centers of mass. Even for these complexes, the linear approximation captures the right behavior within one water layer. On

the basis of these observations, we use linear approximations both for the desolvation and electrostatic terms, which substantially simplifies the Brownian dynamics simulations. It is also important to remember that our calculations are aimed primarily at uncovering the role of short-range forces, acting within the 4.2-Å-thick desolvation shell, where the linear approximation is clearly adequate.

Why do we need a spherical approximation?

Gabdouline and Wade (1997) have shown that Brownian dynamics simulations can be carried out using all-atom protein models; thus, a similar approach might appear to be feasible also in the present study. However, Gabdouline and Wade studied the association of barnase and barstar, which, due to the strong electrostatic steering, is at least two orders of magnitude faster than the desolvation-mediated association reactions considered here. Therefore, it is critical to use the spherical approximation, which can speed up the calculation by more than 100-fold.

RESULTS

Model I: Uniform desolvation

Potential

Recall that, in this model, $\Delta E_{coul} = 0$ and $\Delta E_{des} = \lambda \Delta A$, where ΔA denotes the change in the solvent-accessible surface area, and λ is the solvation parameter. To account for a desolvation interaction between 1–2 water layers, we assume an effective water radius $R_{H_2O} = 2.1$ Å. Then the loss of solvent-accessible surface area in the association of two spheres with radii R_r and R_l varies between zero (at $d_{S-S} > 2R_{H_2O} = 4.2$ Å) to a maximum of $|\Delta A|_{max} = 485$ Å² (at $d_{S-S} = 0$). These numbers are consistent with the all-atom model, because the formation of an encounter complex between α -chymotrypsin and human leukocyte elastase with ovomucoid turkey third domain decreases the solvent-accessible surface area by 300–500 Å², and the ACE model yields desolvation interactions that vanish for $d_{S-S} > 4.2$ Å (see Fig. 3).

Reaction condition

To incorporate orientational constraints, an optimal receptor–ligand encounter pair is defined at $\theta_0 = \phi_0 = \theta_{i0} = \phi_{i0} = \psi_{i0} = 0^\circ$. The reaction condition is met if a collision occurs when the relative orientation of the receptor and the ligand is within the reactive patch defined by solid angles $\varphi_r = 8^\circ$ and $\varphi_l = 20^\circ$, both deviations from the optimal orientation (see Methods and Fig. 1).

Association rates

The rate of collisions between spheres under a centrosymmetric potential is well known (Debye, 1942; DeLisi and Wiegel, 1981) and is shown as a solid line in the top region of Fig. 4 as a function of the solvation parameter λ . The rates calculated by Brownian dynamics simulations, shown as triangles, agree very well with the theory, indicating the

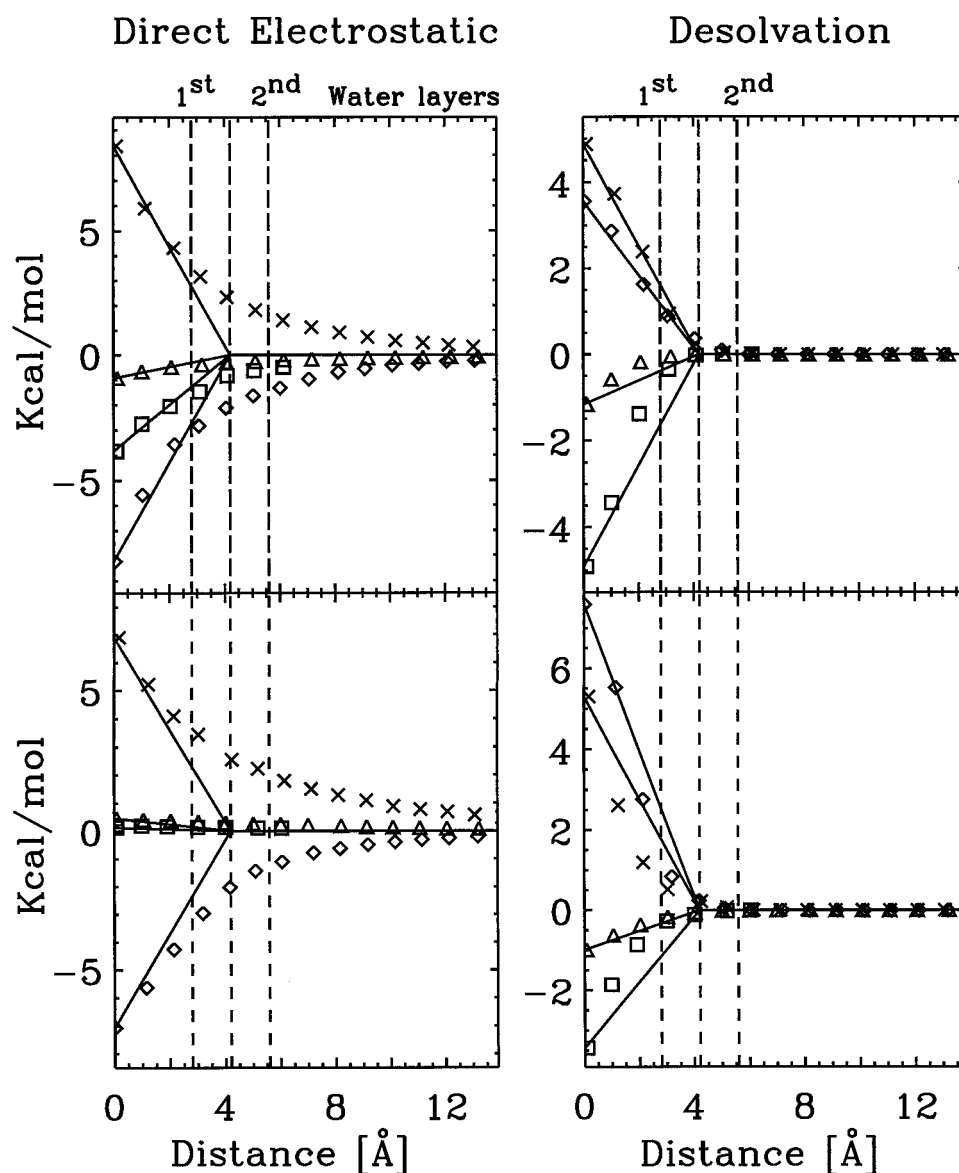


FIGURE 3 Direct electrostatic energy (i.e., electrostatic interaction energy without including the desolvation of polar atoms) and desolvation free energy as functions of the surface-to-surface separation for four encounter complexes. Top plots are for α -chymotrypsin and OMTKY; bottom plots are for human leukocyte elastase and OMTKY. Symbols $\times$ and $\diamond$ correspond to encounter pairs with the largest and lowest electrostatic energy, respectively. Note that, for both systems, some encounter complexes have highly unfavorable desolvation energies. Solid lines correspond to the linear approximation assumed in the Brownian dynamic simulations. Dashed lines indicate the distances defined by one, one and a half, and two layers, respectively, using a water radius of 1.4 Å.

accuracy of the simulations. The maximum rate for this bimolecular reaction, enhanced by the interaction shell of 4.2 Å, is given by the Smoluchowski rate $k_{\text{coll}} = 4\pi D(R_r + R_l + 4.2) = 7.8 \times 10^9 \text{ M}^{-1}\text{s}^{-1}$, and is almost independent of λ .

We now proceed to determine the association rate for the case of geometric constraints, i.e., taking into account the reaction condition. These are calculated by Eq. 3 on the basis of productive collisions. As shown by the straight lines in Fig. 4, the rate grows exponentially with λ (+ symbols),

i.e.

$$k_{\text{on}}(\lambda) \approx k_{\text{on}}(0) \exp(\lambda/\lambda_0), \quad (8)$$

where $\lambda_0 = 1.47 \pm 0.15 \text{ cal}/(\text{mol } \text{\AA}^2)$. This parameter is primarily determined by the Boltzmann factor of the centrosymmetric potential, $RT/|\Delta A|_{\text{max}} \approx 1.2 \text{ cal}/(\text{mol } \text{\AA}^2)$.

Figure 4 shows that nonspecific attractive interactions due to hydrophobicity could significantly enhance the association rate. For the selected reaction condition, the associ-

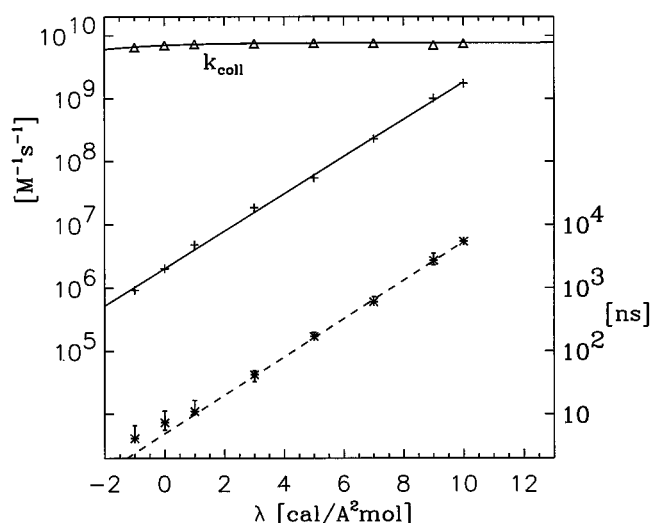


FIGURE 4 Model I: Uniform potential. Association rate as a function of atomic solvation parameter λ . The number of independent runs performed to estimate the rates range from 182,000 for $\lambda = -1$ to 722 for $\lambda = 10$, the corresponding errors vary between 30% and 10%, respectively. Random collision rate constant is indicated by Δ (the solid line is the theoretical value). The + symbols denote the collision rate defined by a reactive surface patch with solid angles $\varphi_r = 8^\circ$ around (θ_0, ϕ_0) , and $\varphi_l = 20^\circ$ around $x_0 y_0 z_0$. The * symbols show the typical time that productive pathways spend within the desolvation layer of 4.2 Å, (time scale is on the right axis).

ation rate increases from $5 \times 10^5 \text{ M}^{-1} \text{ s}^{-1}$ to $2 \times 10^9 \text{ M}^{-1} \text{ s}^{-1}$ as the atomic solvation parameter λ increases from -2 to $10 \text{ cal}/(\text{mol } \text{\AA}^2)$. The high rates are achieved due to what Sommer et al. (1982) described as “lengthy collisions between proteins.” Indeed, the average time the productive trajectories spend within the molecule’s narrow desolvation shell also grows exponentially with λ , and, for $\lambda \geq 8 \text{ cal}/(\text{mol } \text{\AA}^2)$, it is already on the order of microseconds. However, as pointed out by Berg (1985), lengthy collision on a microsecond time scale would yield nonspecific affinities on the order of 10^4 M^{-1} , whereas, experimentally, they do not seem to be larger than 10^2 M^{-1} . Northrup and Erickson (1992) estimated that, to reconcile the experimental evidence that shows no nonspecific protein–protein association, the collisions should not last more than 10 ns. According to our simulations, the hydrophobicity parameter λ must be less than $1 \text{ cal}/(\text{mol } \text{\AA}^2)$ to keep the average surface-on-surface diffusion time below 10 ns, confirming that protein binding cannot be governed solely by nonspecific interactions.

As will be further discussed, the rates heavily depend on the selected reaction condition. For example, assuming no steering forces (i.e., $\lambda = 0$), the prefactor $k_{\text{on}}(0)$ in Eq. 8 depends on the solid angle φ according to the expression

$$k_{\text{on}}(\varphi) \approx 10\kappa(\varphi)k_{\text{coll}}, \quad (9)$$

where

$$\kappa(\varphi) = \frac{1}{4} [1 - \cos(\varphi_r)][1 - \cos(\varphi_l)] \frac{\varphi_l}{180} \quad (10)$$

corresponds to the fraction of angular orientational space, and k_{coll} is the collision rate constant at $\lambda = 0$ (Janin, 1997). As shown by Eq. 9, the association rate constant exceeds the value that would follow from geometric considerations by a factor of 10, which is due to diffusion entrapment (see Berg, 1985; Zhou, 1997).

Model II: Atomic-level interaction potential

Potential

As described in the Methods, in Model II, the desolvation term is calculated using an extension of the ACE. From the atomic interaction potential at close proximity, we find that the average interaction over all possible encounter conformations is *repulsive* for both complexes studied in this paper (see also Camacho et al., 1999). Indeed, the average values of the desolvation free energy ΔG_{ACE} are $1.1 \pm 0.1 \text{ Kcal/mol}$, and $1.2 \pm 0.1 \text{ Kcal/mol}$, respectively, for the chymotrypsin–OMTKY and the elastase–OMTKY complexes. The electrostatic interactions are negligible, $\langle \Delta E_{\text{coul}} \rangle = -0.003 \text{ Kcal/mol}$ and $\langle \Delta E_{\text{coul}} \rangle = -0.01 \text{ Kcal/mol}$, respectively. Thus, the average short-range potential is repulsive and roughly equivalent to $\lambda \lesssim -2 \text{ cal}/(\text{mol } \text{\AA}^2)$. As described for Model I, this value of λ implies that typical macrocollisions last 10 ns or less, which is in good agreement with experiments.

Reaction condition

To introduce a reaction condition, we first need to identify the location of the free energy attractor on each protein surface, i.e., an encounter complex at the center of the binding region, around which the reactive patch will be placed. Table 1 lists the lowest free energy encounter conformations. We recall that, in the selected coordinate system, the relative orientation in the native complex is given by $(90^\circ, 90^\circ; 0^\circ, 0^\circ, 0^\circ)$.

For α -chymotrypsin and OMTKY, all low free energy conformations are relatively close to the crystal orientation. The energy difference between the low energy states is less than 1 Kcal/mol. Thus, from an energetic point of view, there is little or no difference in picking any of the minima on the free energy landscape as the optimal point in the reaction condition. In contrast, from a kinetic point of view, cooperativity is important, and hence it makes sense to choose the geometric center of the low free energy attractor as the optimal site around which to define the surface patch. A trivial angular comparison indicates that the third lowest free energy at $(85^\circ, 85^\circ; 20^\circ, 180^\circ, 180^\circ)$ is closest to the center of the cluster formed by the top nine structures. Thus,

TABLE 1 Free energy ranking of diffusion-accessible encounter complexes

Ranking	θ	ϕ	θ_1 (Degrees)	ϕ_1	ψ_1	ΔG_{des}	ΔE_{coul} (Kcal/mol)	Free Energy
α -chymotrypsin and OMTKY3								
1	80	85	20	240	120	-3.18	-5.841	-9.021
2	85	90	40	120	220	-7.28	-1.523	-8.803
3	85	85	20	180	180	-4.91	-3.839	-8.749
4	80	75	20	200	140	-3.97	-4.327	-8.297
5	85	90	20	220	160	-2.63	-5.532	-8.162
6	85	85	20	220	140	-3.25	-4.890	-8.140
7	90	85	0	0	0	-5.29	-2.784	-8.074
8	85	90	20	200	180	-3.19	-4.823	-8.013
9	85	85	20	200	160	-3.67	-4.277	-7.947
10	80	85	40	120	200	-5.77	-2.157	-7.927
11	85	70	60	60	220	-5.05	-2.682	-7.732
>3500	85	90	20	120	220	-2.52	-1.000	-3.520
...	85	90	20	180	180	-1.35	-2.147	-3.497
>14000	80	85	20	260	120	0.81	-3.118	-2.308
...	90	90	0	0	0	-1.17	-0.923	-2.093
Human leukocyte elastase and OMTKY3								
1	80	95	20	300	40	-7.91	0.529	-7.381
2	90	105	80	0	300	-6.21	-1.068	-7.278
3	75	95	160	100	80	-7.08	-0.177	-7.257
4	75	95	160	80	60	-6.89	-0.325	-7.215
5	85	95	20	340	0	-7.50	0.320	-7.180
6	85	95	20	0	340	-7.54	0.401	-7.139
7	90	95	100	340	280	-4.98	-1.910	-6.890
8	85	90	20	0	340	-7.06	0.278	-6.782
9	95	100	80	340	280	-5.20	-1.507	-6.707
10	70	80	140	140	120	-7.36	0.669	-6.691
11	80	90	20	320	20	-7.04	0.363	-6.677

somewhat arbitrarily, we choose this configuration as the optimal attractor. Notice that the actual phase space encompassed by the binding region is not as large as one might expect on the basis of Table 1. Indeed, an artificial feature of the Euler angle representation is that, for values of θ_1 close to 0° (or 180°), there is almost no orientational change for rotations if $\chi_1 +$ (or $-$) $\psi_1 \approx 360^\circ$, as it is the case with several top ranked encounter pairs. To emphasize the complexity of the landscape, we also list four states located at grid sites adjacent to the top structures, but with much higher energies (see Table 1).

For human leukocyte elastase and OMTKY, Table 1 indicates at least three different low free energy states. The binding pocket is defined around the lowest free energy region that includes the orientations ranked 1, 5, 6, 8, and 11 in Table 1. We choose the top structure at (80° , 95° ; 20° , 300° , 40°) as the attractor, because it is surrounded by the other low-energy structures.

Association rates

Figure 5 shows the association rates obtained by Brownian dynamics simulations for the complexes of α -chymotrypsin with OMTKY and human leukocyte elastase with OMTKY using the interaction potential and the reaction condition just described. Both rates are shown as functions of the solid

angle $\varphi = \varphi_r = \varphi_l$ that specifies the size of the reactive patch in the reaction condition. The association rate calculated for the same reaction condition, but without any interaction force, is also shown.

The most striking difference between the three curves in Fig. 5 is their behavior when the size of the reactive patch is reduced, i.e., when the solid angle φ approaches zero. Without interaction forces, the rate becomes very small. With the potential, however, the rates approach their observed values of $10^7 \text{ M}^{-1}\text{s}^{-1}$ and $10^6 \text{ M}^{-1}\text{s}^{-1}$ for α -chymotrypsin and human leukocyte elastase, respectively. The discrete nature of the time steps and interaction potential prevents reliable calculation of the rates for patch sizes with $\varphi < 2^\circ$. However, the rate constants calculated for $\varphi = 2^\circ$ are in good agreement with the experimental values of $k_{\text{on}} = 1.2 \times 10^7$ and 1.1×10^6 for (bovine) α -chymotrypsin and (porcine) elastase I with OMTKY3, respectively (Ardelt and Laskowski, 1985). These high rates should be compared to the rate calculated for random collisions ($\times$ symbols). Using the given reaction condition, but assuming no interactions, at $\varphi = 2^\circ$, the rate is on the order of $3 \times 10^2 \text{ M}^{-1}\text{s}^{-1}$.

Association rates at nearby sites

As we described, if the interactions are not uniform, the definition of a reaction condition involves selecting a point

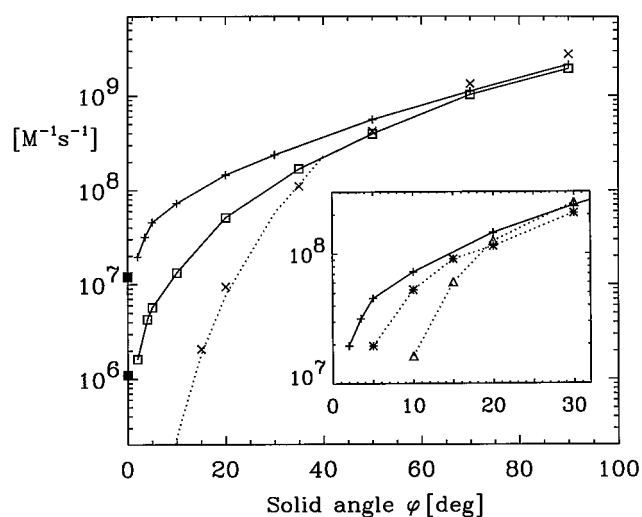


FIGURE 5 Model II: Atomic-level interaction potential. Desolvation mediated binding: association rate for α -chymotrypsin (+) and human leukocyte elastase ($\square$) with OMTKY as a function of $\varphi = \varphi_r = \varphi_l$ (see text). A total of 40,000 and 100,000 runs were made for chymotrypsin and elastase, respectively. The errors of the calculations range from 15–30% at $\varphi = 2^\circ$ to 3% at $\varphi = 90^\circ$ for both chymotrypsin and elastase. The rate calculated for the same reaction condition but without any force field is indicated by the $\times$ symbols (dotted line correspond to Eq. 9). Filled squares indicate the experimental values. Inset: For comparison, we show the association rate for the lowest free-energy encounter complex in Table 1 (*), and for the local minimum indicated by an arrow in Fig. 2 (Δ) as a function of φ .

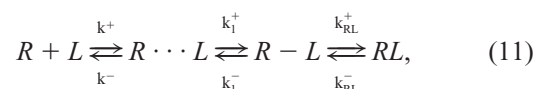
on each protein surface that will become the center of the active site. This selection may affect the calculated reaction rates. The sensitivity of rates to reaction conditions is studied by comparing results for two alternative reaction criteria: one defined around the lowest free energy encounter complex (top structure in Table 1), and another around the local minimum indicated by the arrow in Fig. 2. These rates are shown in the inset of Fig. 5. For the lowest free energy complex, we find that, upon decreasing φ , the rate starts to drop at $\varphi = \sim 15^\circ$, 5° earlier than for the main attractor. This difference reflects the fact that the lowest free energy minimum is closer to the edge of the binding pocket (see Table 1). For the local minimum indicated in Fig. 2, the rate drops sharply for $\varphi < 10^\circ$ because the reaction patch no longer includes the low free energy attractor. For $\varphi > 20^\circ$, the calculated rates are virtually independent of the particular site in the reaction condition.

DISCUSSION

Background: General mechanism of association

To discuss the results of the simulations, it is useful to recall that, in the most general case, the process of diffusion-

limited protein-protein association can be described by a three-step reaction mechanism,



where R denotes the receptor, L the ligand, $R \cdots L$ the nonspecific encounter pairs, and $R - L$ the precursor state(s) leading to the docked conformation (DeLisi and Wiegel, 1981).

If long-range interactions can be neglected, the first reaction step is the random collision of the two proteins R and L , resulting in a nonspecific encounter complex $R \cdots L$ within the desolvation layer. As already mentioned, to a good approximation, the limiting rate k^+ of this first regime is given by the Smoluchowski limit k_{coll} . Indeed, the overall repulsion of the force fields has little effect on k^+ . We have found that the typical lifetime of a nonspecific encounter complex $R \cdots L$ diffusing within the desolvation layer is about 4 ± 1 ns. This value is consistent with the nonspecific affinity between proteins that is estimated to be 10^2 M^{-1} or less (Northrup and Erickson, 1992).

The third reaction step in Eq. 11, i.e., the late transition between the favorable intermediate(s) $R - L$ and the bound state RL , substantially differs from the first two steps. The onset of the late transition coincides with the need to remove steric clashes and charge overlaps in the binding mechanism. Although the first two steps are governed by diffusion, the third is a process of induced fit that requires structural rearrangements involving mostly side chains. Although the actual modeling of this regime is beyond the scope of the present paper, it is safe to assume that this late transition is not diffusive. For proteins that bind in a diffusion-controlled (or diffusion limited) reaction, the rate-limiting step must be the diffusive search for the partially desolvated intermediate(s) or precursor state(s) rather than the third step, and thus $k_{RL}^+ \gg k_1^-$.

In this paper, we focus on the kinetics of the second reaction step and on the nature of the precursor state(s) $R - L$. This step consists of a two-dimensional diffusive transition, driven by desolvation and short-range electrostatic forces, from a nonspecific encounter complex to the precursor state. Depending on the interaction potential, this step may or may not be important. In the well-studied association of barnase and barstar, Step 1 of the mechanism is affected by long-range electrostatic steering toward the binding site, and the encounter complex $R \cdots L$ is likely to be close to the precursor state $R - L$. Therefore, the role of any two-dimensional search is very limited. Without long-range electrostatics, however, any observed rate enhancement is due to short range forces that increase the probability of transition between $R \cdots L$ and $R - L$ states.

Reaction criteria

An uncertainty inherent to any Brownian dynamics simulation is the somewhat arbitrary definition of the reaction condition. In fact, the methods described in the literature show more variation in this condition than in all the other aspects of the algorithm. A simple condition, most frequently used for model proteins, assumes that each sphere bears a small reactive patch around some point on each protein surface. The relative positions and orientations of the receptor and the ligand are defined by the direction and length of the center-to-center vector $\mathbf{r}$ and by the torsion angle χ of rotation around $\mathbf{r}$. The size of the patch on each protein is given by a solid angle φ around $\mathbf{r}$.

Berg (1985) and Zhou (1997) have used the above condition with no constraint on the torsion angle χ . It was shown that, without an interaction potential, the solid angle $\varphi = 5^\circ$ yields a rate constant $k_{\text{on}} = 6 \times 10^5 \text{ M}^{-1}\text{s}^{-1}$ (Zhou 1997). The remaining degree of freedom associated with the free rotation χ has been removed by Janin (1997). If we again assume no interaction potential, we find $k_{\text{on}} = 6 \times 10^5 \text{ M}^{-1}\text{s}^{-1}$ for $\varphi = \chi = 19^\circ$, whereas $k_{\text{on}} = 1 \times 10^5 \text{ M}^{-1}\text{s}^{-1}$ requires about $\varphi = \chi = 14^\circ$ (Janin, 1997). All these calculations account for the repeated microcollisions during the association of encounter complexes.

A contact-based reaction condition has been defined by Northrup and Erickson (1992) for spherical protein models. Each sphere with a radius of 18 Å had a set of four contact points mounted in a 17×17 Å square arrangement tangent to the surface. Each point on one molecule had a partner on the other molecule, and a reaction was considered to occur when N ($N = 1, 2$, or 3) of the four points were simultaneously within 2 Å of their partners. Notice that the cases $N = 1$ and $N = 2$ leave freedom in some orientational degrees, and hence we restrict consideration to the $N = 3$ case, which, apart from the 2-Å mismatch, fully specifies the relative orientations of the two molecules. Simple geometric arguments show that this mismatch can be described in terms of the surface patch model if the solid angle φ and the threshold on the rotational angle χ both are around 10° . Under this condition, Northrup and Erickson obtained the rate constant of $k_{\text{on}} = 1 \times 10^5 \text{ M}^{-1}\text{s}^{-1}$ for purely diffusive association, in relatively good agreement with the results of Janin (1997). They have also shown that a short-range locking potential increases the rate constant to $k_{\text{on}} = 2 \times 10^6 \text{ M}^{-1}\text{s}^{-1}$.

Our reaction condition, defined in terms of the vector $(\theta, \phi; \theta_l, \phi_l, \psi_l)$, fully specifies the precursor state within a solid angle $\varphi = \varphi_r = \varphi_l$, and, in this sense, is similar to the condition used by Janin (1997) or to the one by Northrup and Erickson (1992) at $N = 3$. As shown in Fig. 5, in purely diffusive association, the condition with $\varphi = 10^\circ$ yields a rate constant of $k_{\text{on}} = 1 \times 10^5 \text{ M}^{-1}\text{s}^{-1}$. Thus, despite their formal differences, the various reaction criteria generally

provide similar values for the association rate without any interactions.

The nature of the precursor state

As we already mentioned, in a complex without strong long-range electrostatic interactions, one can find association rates as high as $k_{\text{on}} = 1 \times 10^7 \text{ M}^{-1}\text{s}^{-1}$. This rate could be easily explained by using more permissive reaction criteria. For example, according to Fig. 5, the rate $k_{\text{on}} = 1 \times 10^7 \text{ M}^{-1}\text{s}^{-1}$ can be obtained if we neglect all interactions but assume $\varphi = 24^\circ$ in the reaction condition. This explanation, however, does not take into account that the short-range interactions due to desolvation definitely exist. Indeed, the thermodynamic role of desolvation is generally accepted, and its magnitude is well known. The substantial change in the desolvation free energy, occurring mostly within the first one or two water layers around the protein, necessarily yields forces associated with desolvation.

Although the mean desolvation forces are repulsive, desolvation and short-range electrostatics can substantially increase the association rate for the two complexes studied in this paper (Fig. 5). Agreement with experimentally determined rate constants can be attained under two very different assumptions. The first is assuming a very small solid angle $\varphi \approx 2^\circ$, and a necessarily fast last step to the complex RL . The second mechanism assumes a much larger ensemble of intermediates, and a slow rate to the final bound state, i.e., $k_{\text{RL}}^+ \sim k_1^-$. However, this second mechanism contradicts the diffusion-limited assumption. Thus, for a diffusion-limited process, the reaction condition must correspond to the first mechanism, which involves a small set of well-defined precursor states.

We emphasize that the above results remain valid despite the potential uncertainties in the calculated rate constants. The error bars, estimated following Gabdoulline and Wade (1997), are $\sim 30\%$ (see Captions for Figs. 4 and 5). Comparing our results to the experimentally determined rate constants shows that agreement can be attained only by assuming a very small solid angle φ in the reaction condition. In this region, the curves are so steep that even a 100% error in the rate constants has only minor effects on the required value of φ , and hence the conclusion concerning the nature of the precursor state is rather robust, i.e., independent of the potential errors in the calculation.

Having an almost unique precursor state suggests a relatively narrow pathway from the very restricted set of precursor states to the also unique high-affinity complex. Indeed, this late transition to the bound conformation involves going from a partially desolvated interface of $300\text{--}500 \text{ Å}^2$ to a fully desolvated interface of $1400\text{--}1600 \text{ Å}^2$, where shape complementarity and steric effects are most important. The trajectory between these two states includes a translation of $4\text{--}6 \text{ Å}$, i.e., the distance separating the encounter pairs from the complex structure. It is reasonable to

assume that the binding (unbinding) pathway between these two more-or-less unique states is restricted to a narrow channel in conformational space. It is only outside this channel, at partially solvated intermediates, that one could start to envisage the possibility for multiple binding (unbinding) pathways. Thus, we conclude that the late transition is a highly collimated pathway from the small set of precursor states to the complex structure. This transition should be mostly driven by a fast downhill enthalpic reduction, consistent with a locking process on the order of milliseconds (Laskowski, private communication; Lombardi et al., 1992).

CONCLUSIONS

We have performed Brownian dynamics simulations of receptor–ligand association without strong long-range electrostatic interactions. Simulations with a uniform interaction potential have shown that short-range attractive interactions substantially increase the association rate, but also result in large nonspecific affinities not seen in real proteins (Northrup and Erickson, 1992). The analysis of a realistic interaction potential, including both a well-established atomic-level desolvation and short-range electrostatics, resolves this contradiction by showing that, on average, these forces result in repulsive interactions that prevent nonspecific association. The simulations also identify weakly specific pathways leading to a low free energy attractor embedded in the otherwise repulsive environment. Along these pathways, the association rates are enhanced by the desolvation force field, both by locally increasing the diffusion entrapment and by guiding the proteins towards a *small* set of well-oriented kinetic intermediates. The results of Brownian dynamics simulations show that a diffusion-limited process can be reconciled with experimentally determined association rates only by assuming that the final bound conformation is preceded by an almost unique precursor state. Although desolvation can increase the association rates by several orders of magnitude, these rates are still 100-fold smaller than the ones observed for complexes in which long-range electrostatics provide the binding specificity.

We thank Dr. Michael Laskowski, Jr. for helpful comments. This work was supported by National Science Foundation grant DBI-9630188 and Department of Energy grant DE-F602-96ER62263.

REFERENCES

- Ardelt, W., and M. Laskowski, Jr. 1985. Turkey ovomucoid third domain inhibits eight different serine proteinases of varied specificity on the same $\cdots\text{Leu}^{18}\text{--Glu}^{19}\cdots$ reactive site. *Biochemistry*. 24:5313–5320.
- 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.
- Camacho, C. J., Z. Weng, S. Vajda, and C. DeLisi. 1999. Free energy landscapes of encounter complexes in protein–protein association. *Biophys. J.* 76:1166–1178.
- Chothia, C., and J. Janin. 1975. Principles of protein–protein recognition. *Nature*. 256:705–708.
- Debye, P. 1942. Reaction rates in ionic solutions. *Trans. Electrochem. Soc.* 82:265–272.
- DeLisi, C. 1980. The biophysics of ligand–receptor interactions. *Q. Rev. Biophys.* 13:201–230.
- DeLisi, C., and F. Wiegand. 1981. Effect of nonspecific forces and finite receptor number on rate constants of ligand–cell bound–receptor interactions. *Proc. Natl. Acad. Sci. USA*. 78:5569–5572.
- Eisenberg, D., and A. D. McLachlan. 1986. Solvation energy in protein folding and binding. *Nature*. 319:199–203.
- Eltis, L. D., R. G. Herbert, P. D. Barker, A. G. Mauk, and S. H. Northrup. 1991. Reduction of horse ferricytochrome *c* by bovine liver ferrocyanochrome *b*₅. Experimental and theoretical analysis. *Biochemistry*. 30:3663–3674.
- Ermak, D. L., and J. A. McCammon. 1978. Brownian dynamics with hydrodynamic interactions. *J. Chem. Phys.* 69:1352–1360.
- Gabdoulline, R. R., and R. C. Wade. 1997. Simulation of the diffusional association of barnase and barstar. *Biophys. J.* 72:1917–1929.
- Hermann, R. B. 1972. Theory of hydrophobic bonding. II. The correlation of hydrocarbon solubility in water with solvent cavity surface area. *J. Phys. Chem.* 76:2754–2759.
- Israelachvili, J., and H. Wennerstrom. 1996. Role of hydration and water structure in biological and colloidal interactions. *Nature*. 379:219–225.
- Janin, J. 1997. The kinetics of protein–protein recognition. *Proteins: Struct. Funct. Genet.* 28:153–161.
- Lee, B., and F. M. Richards. 1971. The interpretation of protein structures: estimation of static accessibility. *J. Mol. Biol.* 55:379–400.
- Lombardi, V., G. Piazzesi, and M. Linari. 1992. Rapid regeneration of the actin myosin power stroke in contracting muscle. *Nature*. 355:638–641.
- Luty, B. A., R. C. Wade, J. D. Madura, M. E. Davis, J. M. Briggs, and J. A. McCammon. 1993. Brownian dynamics simulations of diffusional encounters between triose phosphate isomerase and glyceraldehyde phosphate–electrostatic steering of glyceraldehyde phosphate. *J. Phys. Chem.* 97:233–237.
- Northrup, S. H., S. A. Allison, and J. A. McCammon. 1984. Brownian dynamics simulation of diffusion-influenced bimolecular reactions. *J. Chem. Phys.* 80:1517–1524.
- Northrup, S. H., and H. P. Erickson. 1992. Kinetics of protein–protein association explained by Brownian dynamics computer simulations. *Proc. Natl. Acad. Sci. USA*. 89:3338–3342.
- Schreiber, G., and A. R. Fersht. 1996. Rapid, electrostatically assisted association of proteins. *Nature Struct. Biol.* 3:427–431.
- Sharp, K., R. Fine, and B. Honig. 1987. Computer simulations of the diffusion of a substrate to an active site of an enzyme. *Science*. 236:1460–1463.
- Smoluchowski, M. V. 1917. Versuch einer mathematischen Theorie der Koagulationskinetik kolloider Loesungen. *Z. Phys. Chem.* 92:129–168.
- Sommer, J., C. Jonah, R. Fukuda, and R. Bersohn. 1982. Production and subsequent second-order decomposition of protein disulfide anions. *J. Mol. Biol.* 159:721–744.
- Stone, R. S., S. Dennis, and J. Hofsteenge. 1989. Quantitative evaluation of the contribution of ionic interactions to the formation of thrombin–hirudin complex. *Biochemistry*. 28:6857–6863.
- Vajda, S., Z. Weng, R. Rosenfeld, and C. DeLisi. 1994. Effect of conformational flexibility and solvation on receptor–ligand binding free energies. *Biochemistry*. 33:13977–13988.

- Vijayakumar, M., K.-Y. Wong, G. Schreiber, A. R. Fersht, A. Szabo, and H.-X. Zhou. 1998. Electrostatic enhancement of diffusion-controlled protein-protein association: comparison of theory and experiment on Barnase and Barstar. *J. Mol. Biol.* 278:1015–1024.
- von Hippel, P. H., and O. G. Berg. 1986. On the specificity of DNA-protein interactions. *Proc. Natl. Acad. Sci. USA.* 83:1608–1612.
- Wolfenden, R., Andersson, L., Cullis, P. M., and C. C. B. Southgate. 1981. Affinities of amino acid side chains for solvent water. *Biochemistry.* 20:849–855.
- Zhang, C., G. Vasmatazis, J. L. Cornette, and C. DeLisi. 1997. Determination of atomic desolvation energies from the structures of crystallized proteins. *J. Mol. Biol.* 267:707–726.
- Zhou, G. 1979. Influences of van der Waals force upon diffusion-controlled reaction rate. *Scientia Sinica.* 22:845–858.
- 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.