PROTEINS: Structure, Function, and Genetics 28:153–161 (1997)

The Kinetics of Protein-Protein Recognition

Joël Janin*
Laboratoire d' Enzymologie et de Biologie Structurales, UPR 9063 CNRS, Bât. 34, 91198 Gif-sur-Yvette, France

ABSTRACT **We examine a simple kinetic model for association that incorporates the basic features of protein-protein recognition within the rigid body approximation, that is, when no large conformation change occurs. Association starts with random collision at the rate k_{coll} predicted by the Einstein-Smoluchowski equation. This creates an encounter pair that can evolve into a stable complex if and only if the two molecules are correctly oriented and positioned, which has a probability p_r. In the absence of long-range interactions, the bimolecular rate of association is p_r k_{coll}. Long-range electrostatic interactions affect both k_{coll} and p_r. The collision rate is multiplied by q_t, a factor larger than 1 when the molecules carry net charges of opposite sign as coulombic attraction makes collisions more frequent, and less than 1 in the opposite case. The probability p_r is multiplied by a factor q_r that represents the steering effect of electric dipoles, which preorient the molecules before they collide. The model is applied to experimental data obtained by Schreiber and Fersht (Nat. Struct. Biol. 3:427–431, 1996) on the kinetics of barnase-barstar association. When long-range electrostatic interactions are fully screened or mutated away, $q_t q_r \approx 1$, and the observed rate of productive collision is p_r $k_{coll} \approx 10^5\ M^{-1} \cdot s^{-1}$. Under these conditions, $p_r \approx 1.5 \cdot 10^{-5}$ is determined by geometric constraints corresponding to a loss of rotational freedom. Its value is compatible with computer docking simulations and implies a rotational entropy loss $\Delta S_{rot} \approx 22$ e.u. in the transition state. At low ionic strength, long-range electrostatic interactions accelerate barnase-barstar association by a factor $q_t q_r$ of up to 10^5 as favorable charge-charge and charge-dipole interactions work together to make it much faster than free diffusion would allow. Proteins 28: 153–161, 1997.**

Key words: collision theory; protein-protein association; electrostatic interactions; dipole steering; barnase-barstar complex; rotational-translational entropy

INTRODUCTION

Protein-protein recognition results from the assembly of complementary surfaces on two molecules that form a stable, albeit noncovalent, specific complex. The physical chemical basis of the stability[1,2] is reasonably well understood for systems, such as antigen-antibody or enzyme-inhibitor complexes, that undergo no large conformation change. The interfaces are large and close-packed; they exclude water; buried polar groups form hydrogen bonds and salt bridges. Stability is estimated at thermodynamic equilibrium by the dissociation constant and dissociation-free enthalpy. Here, we deal with the physical mechanism of association and the kinetic aspects of recognition. The relevant parameter is the bimolecular rate constant for association k_a. Its value can be derived from collision theory, taking into account the effect of long-range electrostatic interactions on the diffusion and orientation of the protein molecules.[3,4] A simple model of the rigid-body association reaction is presented and shown to yield an accurate description of kinetic data obtained by Schreiber and Fersht[5–7] on the barnase-barstar system,[8] a typical and well-documented example of specific protein-protein recognition.

A Mechanism for Complex Formation Between Proteins

The reaction pathway

The kinetic model shown in Figure 1 is the simplest that retains the essential feature of protein-protein recognition in the rigid-body approximation: the precise assembly of two sterically and chemically complementary surfaces leaving no degree of freedom to the component molecules. Protein molecules A and B come together by translational diffusion; if they are in the right position and orientation, an interface may develop as part of the molecular surface becomes dehydrated and internal degrees of freedom relax to optimize short-range interactions.

The reaction starts with a collision yielding the encounter pair of collision theory,[9] noted A . . . B here. Because the components are free to diffuse away, the half-life of the encounter pair is a fraction

*Correspondence to: Joël Janin, Laboratoire d'Enzymologie et de Biologie Structurales, UPR 9063 CNRS, Bât. 34, 91198 Gif-sur-Yvette, France.
Received 20 September 1996; Accepted 5 December 1996

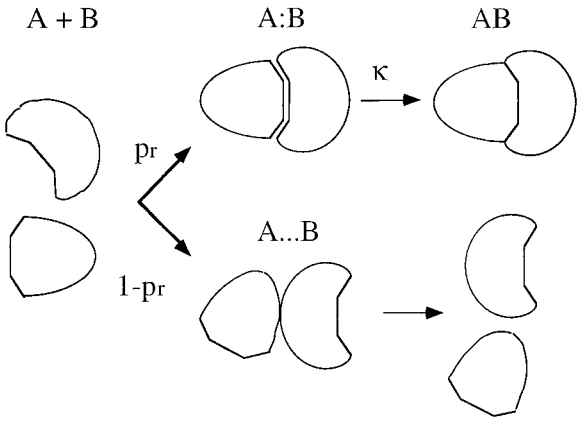

Fig. 1. A reaction pathway for protein-protein recognition. Molecules A and B diffuse freely in solution (A+B) and undergo random collision. This yields the encounter pair A . . . B, which has a probability p_r of being in the particular orientation that characterizes the precomplex A:B. Although other encounter pairs are short-lived and must dissociate, A:B has a high probability $\kappa \approx 1$ to yield the proper complex AB.

of a nanosecond, the time it takes for A or B to move by a few angstrom units at T = 300 K. The encounter pair is formed at a rate equal to k_{coll} [A][B] where k_{coll} is the bimolecular collision rate constant, [A] and [B], the concentrations in molar units. If A and B are correctly positioned and oriented to start with, the encounter pair can evolve into the final product without further global motion of the components. We call it a 'precomplex' noted A:B, AB being the final stable product. With macromolecules, rotational diffusion takes place in nanoseconds, a time much longer than the half-life of the encounter pair.[9] Thus, the precomplex forms only if the two components happen to be in the correct orientation when they collide, whereas rotation within the encounter pair should also be envisaged for a small molecule ligand.

In A:B, translational and rotational movements of the two components are correlated, but no proper interface has yet formed, and the two molecules are still hydrated. Thus, there is little enthalpy change between A+B, A . . . B, and A:B, yet there is an entropy and therefore a free enthalpy (Gibbs free energy) change, which we break into two components: ΔG_{trans} between A+B and A . . . B, ΔG_{rot} between A . . . B and A:B. ΔG_{trans} reflects the loss of translational entropy of two molecules forming an encounter pair; ΔG_{rot} reflects the rotational entropy cost of selecting a particular orientation in the precomplex. Because the number of free particles changes from two to one at the first step, ΔG_{trans} depends on the concentrations of A and B. Assuming all chemical species to be at 1 M concentration, ΔG_{trans} is defined in the same standard state as the free enthalpy of dissociation ΔG_d.

The precomplex A:B has lost entropy with little or no compensation from favorable interactions and hydration changes. It is a high free enthalpy species, which we consider to be the transition state of the reaction. Furthermore, we assume that there is a smooth path down an energy well leading from A:B to AB. Solvent is excluded and side chain rearranged at the interface with no significant activation barrier. A:B has a high probability κ of yielding AB and, therefore, the bimolecular rate constant for forming A:B is close to that for forming AB, which is the association rate constant k_a. It is less than k_{coll} by a factor equal to the probability p_r that the two components are correctly oriented.

Geometric constraints and long-range interactions

For two hard spheres of equal size, the rate of collision is given by the Einstein-Smoluchowski equation.[9] At 300 K and in water of viscosity $\eta \approx 1$ cP, it evaluates to:

$$k_{coll} = \frac{8RT}{3\eta} \approx 6.6.10^9\ M^{-1}\ s^{-1} \qquad (1)$$

This equation is valid for spheres that interact isotropically and only at very short range. It predicts k_{coll} to be independent of their size, because the increased target size of a bigger sphere compensates for its slower diffusion. Real proteins do not behave that way. Orientation very much matters; targets are well-defined regions of the molecular surface, and their assembly can take place only when they are properly positioned. Making the molecule bigger slows diffusion without necessarily changing the target size, thus reducing the association rate. Moreover, the molecules bear electric charges and are subject to coulombic forces even before they collide. Thus, Equation 1 must be modified to account for a) the precise localization and orientation of the interface and b) long-range interactions between diffusing molecules. The first constraint is geometric and taken care of by a factor p_r. This is defined as the probability that, in the encounter pair, the situation shown in Figure 2 prevails: the two components of the putative interface face each other in the right orientation. Similar figures appear in studies by Berg and von Hippel[4] and Zhou[10] who also analyzed the rate of diffusion-controlled macromolecular reactions. However, their models do not restrict rotation of the molecules about the line of centers, and their angular parameters define the size of the patches on the molecular surface, not their orientation. In Figure 2, parameters locate the patches on the surface and orient them one relative to the other, leaving no degree of rotational/translational freedom to the two components of the precomplex.

The effect of interactions may now be considered. In our model, short-range electrostatic (hydrogen bonds), van der Waals, and hydrophobic (dehydra-

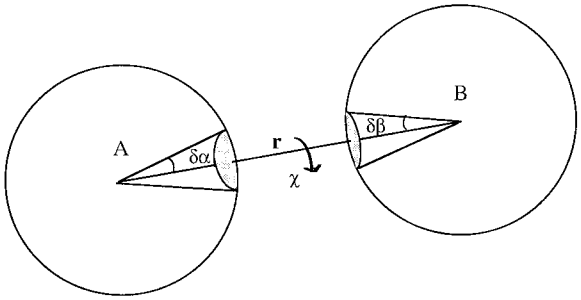

Fig. 2. Geometric constraints in the precomplex. The relative positions and orientations of A and B in the precomplex A:B are defined by the direction and length of the center-to-center vector **r** and by the torsion angle χ of rotation about **r**. In a reference frame attached to A (resp. B), the direction of **r** determines the position on the surface of A (resp. B), of the specific patch that will form the interface. The size of the patches is not considered here, only the probability of finding **r** within an angle $\delta\alpha$ (resp. $\delta\beta$) of the correct direction. This is simply proportional to $2\pi\ (1-\cos\delta\alpha)$, the solid angle of the cone. Then, the combined probability that **r** is correctly oriented in both reference frames and that χ is within $\delta\chi$ of the correct value is:

$$p_r = \frac{1}{2}(1-\cos\delta\alpha)\frac{1}{2}(1-\cos\delta\beta)\frac{\delta\chi}{\pi} \approx \frac{1}{16\pi}\delta\alpha^2\,\delta\beta^2\,\delta\chi$$

The value $p_r \approx 1.5 \cdot 10^{-5}$ observed for the barnase-barstar complex suggests that each parameter is fixed to within ~14 degrees in the precomplex. Note that, although this figure resembles Figure 1 of Ref. 4, the angular parameters play a different part and the formula yielding k_a comes out quite different.

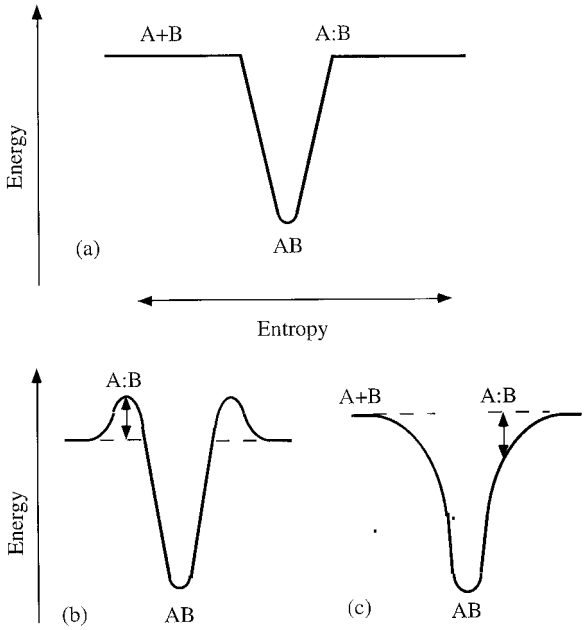

Fig. 3. Energy-entropy diagrams for rigid-body protein-protein recognition. Relative to the freely diffusing molecules A+B, both the complex AB and the precomplex A:B have a low entropy because of the loss of external (rotation-translation) degrees of freedom. In AB, additional entropy is lost as internal (vibrations and side-chain conformation) degrees of freedom freeze at the interface, but the energy is low because of favorable short-range interactions. In (**a**) there is no interaction and no energy change until AB forms; A:B is therefore at a free energy maximum. In (**b**), repulsive long-range interactions add an enthalpic activation barrier (double arrow) that makes association slower, to the entropic barrier already present in (a), making it more obvious that A:B is the transition state. In (**c**), attractive interactions lower the barrier and accelerate association as found in barnase-barstar.

tion) interactions have no effect on the rate constant for association, because they are not present in the transition state. They appear only in the final complex AB and determine its stability. In contrast, net electric charges and dipole moments create forces and torques that extend over distances larger than the average protein size. Such long-range electrostatic interactions introduce a correlation between the position and orientation of molecules in solution. The positional correlation depends on the net charges, collision being more probable between molecules carrying charges of opposite sign. The orientational correlation is due to the charge-dipole and dipole-dipole interactions. As a consequence, both the collision rate and the probability that a collision yields a precomplex, change. This is illustrated (Fig. 3) by energy-entropy diagrams drawn in three cases: a) in the absence of long-range interactions; b) when charge repulsion and/or unfavorable dipole orientations make association slower than predicted by Equation 1; and c) in the opposite case where it is faster. Although only b) displays an energy barrier, there is a free energy barrier in all three cases due to the low entropy of the A:B species relative to the free molecules, after which the system falls into a deep energy well. We now write the association rate constant as:

$$k_a = \kappa\, q_t\, q_r\, p_r\, k_{coll} \qquad (2)$$

where k_{coll} is the rate of 'free' collision given by Equation 1; κ, which we shall assume ≈ 1, is the probability that A:B yields AB, equivalent to the transmission coefficient of reaction rate theory; p_r is the geometric factor defined above; q_t and q_r are acceleration factors due (q_t) to the effect of long-range interactions on the collision rate and (q_r) to their orientational or steering effect. A value of q_t has been derived analytically in model systems. The simplest of all has a pair of point charges diffusing in a medium of uniform dielectric constant and in the absence of screening counter-ions [3]. Then, q_t depends only on the ratio of U_{elec}, the coulombic potential energy, to the thermal energy RT. U_{elec} is calculated at the distance of closest approach, i.e., in the encounter pair. Zhou[10] analyzed the collision of hard spheres bearing localized charges at non-zero ionic strength and found that q_t varies like the Boltzmann factor $\exp(-U_{elec}/RT)$. The analytical derivation was quite elaborate even though it assumed cylindrical symmetry, and a more realistic analytical treatment of the interaction of charges and electric dipoles needs to be performed.

These effects are more easily incorporated in computer simulations by Brownian dynamics. Northrup et al.[11] did such a simulation on cytochrome c peroxidase. The results suggest that long-range electrostatics can accelerate the reaction with cytochrome c by at least three orders of magnitude and that dipolar steering contributes more than coulombic forces between point charges, e.g., $q_r \gg q_t$. A model study of the dynamics of the O_2^- ion binding to superoxide dismutase[12,13] yields an acceleration factor of only one order of magnitude. This small ligand has a full charge but no dipole, and orientational effects are negligible.

The free enthalpy of activation

Multipliers of the free collision rate k_{coll} in Equation 2 are related to a component of the free enthalpy of activation, which we note $\Delta G^\dagger$:

$$\Delta G^\dagger = -RT \ln \frac{k_a}{k_{coll}} \qquad (3)$$

In Figure 1, this is the free enthalpy change on going from A . . . B to A:B. The actual activation barrier $\Delta G_a^\dagger$ encountered by associating molecules depends on their concentrations. In the 1 M standard state, $\Delta G_a^\dagger$ is equal to $\Delta G^\dagger$ plus the translational free enthalpy change ΔG_{trans}, implicitly present in k_{coll}. We associate ΔG_{rot} with the geometric orientation factor p_r and introduce a term noted ΔG_{Irange} to describe the overall effect of long-range interactions:

$$\Delta G_{rot} = -RT \ln p_r$$

$$\Delta G_{Irange} = -RT \ln q_t q_r$$

$$\Delta G^\dagger = \Delta G_{rot} + \Delta G_{Irange} \qquad (4)$$

In the reverse direction (dissociation), the activation barrier between AB and A+B is:

$$\Delta G_d^\dagger = \Delta G_{trans} + \Delta G_{rot} + \Delta G_{Irange} + \Delta G_d \qquad (5)$$

To satisfy microscopic reversibility, $\Delta G_d^\dagger$ must also be equal to the free enthalpy change between A:B and AB. This contains all the enthalpic end entropic terms associated with dehydration of the interface, short-range interactions, conformation changes, side-chain immobilization, and so on, which determine the equilibrium constant. All these terms contribute to the free enthalpy of dissociation ΔG_d, but not to $\Delta G^\dagger$ in our model of the association reaction.

The Kinetics of Barnase-Barstar Association

We now turn to the study of Schreiber and Fersht[6,7] of how electrostatic interactions affect the rate of barnase-barstar association. A comprehensive set of experimental data provides a test of our model and fixes the value of some of its parameters, which we can compare with numerical estimates from computer simulations. Barnase is an extracellular bacterial ribonuclease. Inside the cell, it has a natural protein inhibitor called barstar.[8] Very tight binding is needed because any intracellular nuclease activity is lethal,[14] and K_d is in the femtomolar range. Both proteins are globular and small with M_r near 12 and 10 kDa. Three-dimensional structures are available for the complex[15,16] and for each of the two components.[17,18] They show that the proteins behave like rigid bodies to a good approximation: side chains rearrange, but the main chain hardly moves. Association removes 1590 Å^2 of protein surface from contact with the solvent; 46% of this surface is polar. Buried polar groups form a total of ~10 hydrogen bonds and salt bridges. These numbers are typical of proteins that form stable, specific complexes.[1,19] Thus, barnase-barstar is a representative example of protein-protein recognition.

Electrostatic steering and facilitated association

Electrostatic interactions make an essential contribution to the thermodynamic stability of the barnase-barstar complex.[5] They are equally important in determining the kinetics of association. Binding is extremely fast at low ionic strength; k_a is 10^9 $M^{-1} \cdot s^{-1}$ at $I = 10^{-2}$ M and extrapolates to $\approx 10^{10}$ $M^{-1} \cdot s^{-1}$ at $I = 0$. It slows down considerably when electrostatic interactions are shielded or mutated away. Extrapolation to infinite ionic strength yields $k_a \approx 10^5$ $M^{-1} \cdot s^{-1}$ both on the wild-type proteins and on a series of point mutants bearing different net charges. These effects are largely unspecific because they disappear in the presence of a polyanion, such as heparin, and amino acid substitutions that change the rate of association are not necessarily limited to the interface.[7]

At neutral pH, the net electrostatic charges of wild-type barnase and barstar are approximately +2 and −6. Point mutations on barnase that increase its net positive charge by one or two units accelerate association by a factor of 3–10 at low ionic strength; mutations that decrease it by one unit slow it by a similar factor. Point mutations on barstar that make the charge less negative also make association slower. Moreover, a double barnase mutant that has approximately zero net charge, associates at the minimum rate $k_a \approx 10^5$ $M^{-1} \cdot s^{-1}$ at all but the lowest ionic strengths.[7] This is what we expect from simple point charge electrostatics, because this particular mutant protein has $U_{elec} \approx 0$ and $q_t \approx 1$ at whatever ionic strength.

At infinitely high ionic strength, k_a depends on the solvent viscosity as expected from Equation 1, but it is independent of the net charge on the proteins. We may therefore assume that under these conditions both q_t and q_r are near unity leading to:

$$p_r \approx \frac{k_a}{k_{coll}} \approx 1.5.10^{-5}$$

Because $\Delta G_{lrange} \approx 0$, the value of $\Delta G^{\dagger}$ given by Equation 3 is also equal to ΔG_{rot}:

$$\Delta G^{\dagger} \approx \Delta G_{rot} \approx 6.6 \text{ kcal.mol}^{-1}$$

At the other extreme where $I = 0$, $k_a \approx 10^{10}\ M^{-1} \cdot s^{-1}$ is approximately equal to k_{coll} for the wild-type proteins. Thus, $\Delta G^{\dagger} \approx 0$ as ΔG_{lrange} fully compensates ΔG_{rot}. Electrostatic acceleration is even larger in mutants where the net charge of barnase is +3 or +4 instead of +2. Then, $\Delta G^{\dagger}$ becomes negative, and, at low ionic strength, k_a exceeds the diffusion limit given by Equation 1. For the wild-type proteins at $I = 0$, we have:

$$q_t q_r p_r = \frac{k_a}{k_{coll}} \approx 1.5$$

By comparison with the value of p_r derived above, we obtain:

$$q_t q_r \approx 10^5$$

Available kinetic data do not distinguish between the two terms. In addition to having net charges of opposite signs, barnase and barstar have very asymmetric charge distributions and a large dipole moment. The barstar dipole points straight away from the interface, where barnase accumulates positively charged residues (Fig. 4). The charge-dipole component of the electrostatic interaction must therefore help orient the inhibitor in a way that depends on the net charge of barnase as suggested by the mutant data. However, the point substitutions described above have a much lesser effect on dipoles than on the charges, and other mutations must be designed to specifically test the dipole contribution.

Geometric Constraints and Rotational-Translational Entropy

Docking evaluations of the geometric factor p_r

The magnitude of the geometric p_r factor can also be estimated by simulating random collision on a computer and counting how many encounter pairs have the two molecules in the right position and orientation. Northrup and Erickson[20] did such a simulation with a pair of hard spheres carrying three contact points (I call them hooks for simplicity) each. A bond was made if hooks were within 2 Å of each other. The success rate was $5 \cdot 10^{-4}$ for spheres making one bond and $2 \cdot 10^{-7}$ with three bonds, the minimum number needed to uniquely determine the mode of association. The authors found the latter ratio much too low to account for observed rates and concluded that two bonds must be sufficient to commit the pair to form a stable complex. This is equivalent to leaving one degree of freedom to the components in the transition state.

We performed a similar calculation on the barnase-barstar system by using a rigid-body docking algorithm[21,22] that uses a protein model with one sphere per residue and does not take long-range interaction into account. A total of $5.6 \cdot 10^7$ dockings were needed to cover all orientations by systematically exploring angular parameters in steps of 7.5 degrees. Sixteen produced docked complexes with angular parameters within one step of the correct solution and an interface burying at least 1000 Å^2. If we take these 16 dockings as having generated precomplexes, the success rate is $3 \cdot 10^{-7}$, close to that for making three bonds in,[20] and a factor of 50 below the estimate of p_r derived above from experimental kinetics. In our simulation, the factor of 50 just means that 1000 Å^2 is too stringent a success criterion. Taking 800 Å^2, half the final interface area as a cutoff, raised the success rate by a factor of 15, reasonably close to the experimental value. The low fraction of three-bonded pairs in[20] may be due to the stringent condition that three bonds be made with only three hooks present on each partner. A real protein has many potential binding features, and not all of them need be active in a precomplex. Except in a completely rigid system, adding hooks would increase the rate of three-bond success. Overall, both types of simulations are compatible with the observed value of $p_r \approx 1.5 \cdot 10^{-5}$ and with the idea that, in the transition state, all six rigid-body parameters fixing the relative position and orientation of the two molecules are already determined.

Rotational-translational entropy

The value of p_r yields $\Delta G_{rot} = 6.6$ kcal $\cdot$ mol^{-1} $\cdot$ ΔG_{rot} is a component of the activation barrier representing the cost of selecting the particular orientation that makes a precomplex of the encounter pair. Because ΔG_{rot} is essentially entropic, the value derived from the experimental data on barnase-barstar is equivalent to 22 e.u. (cal $\cdot$ mol^{-1} $\cdot$ K^{-1}). This entropy change is associated with the loss of three degrees of rotational freedom when two molecules become one in the complex, which we note ΔS_{rot}. At the same time, three degrees of translational freedom are also lost, and the corresponding entropy change is noted ΔS_{trans}. It should be pointed out that the six degrees of freedom do not vanish but become internal degrees of freedom or vibrations, which carry less entropy.

This value of ΔS_{rot} is at the low end of the range Finkelstein and Janin[23] derived for the BPTI pancreatic inhibitor in the trypsin-BPTI complex (24–30 e.u.). This derivation involved no kinetic argument; rather, it was based on the assumption that the relative motion of the components of a stable complex is comparable with that of whole molecules in a protein crystal. For lysozyme, the RMS amplitude of

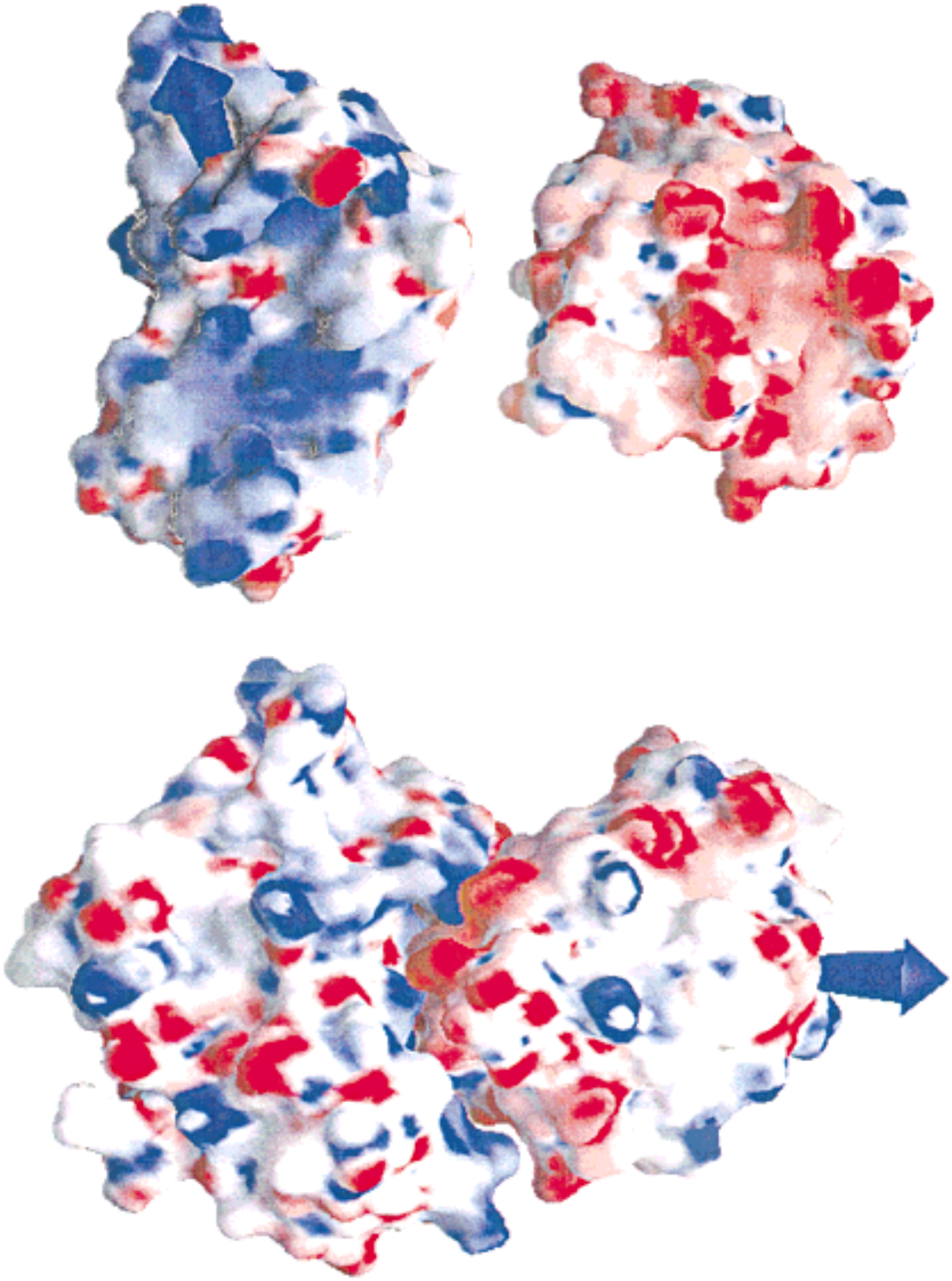

Fig. 4.

translational movements within the crystal[24] is known to be $\delta x \approx 0.25$ Å near 300 K. The translational entropy change ΔS_{trans} (23 e.u. in this calculation) was directly derived from δx; ΔS_{rot} = 24–30 e.u. was obtained in a second step by considering oscillations that would generate 0.25 Å movements at the surface of a globular molecule of radius 6–15 Å, a range that covers the size of the proteins considered both then and in the present study. Erickson[25] and other authors leave more residual freedom to components in the complexes, typically δx = 1–2 Å instead of 0.25 Å; consequently, they obtain lower values of ΔS_{trans} and ΔS_{rot}. In contrast, a simulation of insulin dimerization[26] argues in favor of a larger entropy change. The kinetic data on barnase-barstar provide a fully independent estimate that is compatible with the higher value of ΔS_{rot}, not the lower. They validate the hypothesis that motions within a specific protein-protein complex in solution resemble collective movements observed in lysozyme crystals and therefore, indirectly validate the estimate of ΔS_{trans} based on this hypothesis.

Together, ΔS_{trans} and ΔS_{rot} determine the value of ΔG_{rt}, the free enthalpy change associated with the loss of six rotational-translational degrees of freedom. Its significance, emphasized by Page and Jencks,[27] can be stated in this way: for two molecules to achieve a dissociation constant of 1 M (and ΔG_d = 0), interactions that favor association, whether hydrophobic or polar, must provide at least the equivalent of ΔG_{rt}. To achieve the much lower values of the dissociation constant observed in specific complexes like barnase-barstar or trypsin-BPTI, they must provide $\Delta G_d + \Delta G_{rt}$, not just ΔG_d as too often assumed. Thus, ΔG_{rt} is an important quantity when relating model binding energies to experimental free enthalpy changes,[28,29] and it is not a small correction as sometimes assumed.[30,31] A value of $\Delta G_{rt} \approx 15$ kcal $\cdot$ mol^{-1} applies to the trypsin-BPTI complex according to Finkelstein and Janin.[23] The present analysis suggests that it also applies to the barnase-barstar complex.

DISCUSSION

The model of rigid-body association developed here differs from previous ones in requiring that, in the transition state, the relative positions and orientations of the two molecules are fixed, although perhaps not as accurately as in the final product. This implies stringent geometric requirements, the loss of six degrees of freedom, and a significant entropy of activation, which, together with the Einstein-Smoluchowski equation, lead to a predicted value $k_a \approx 10^5$ M$^{-1} \cdot$ s^{-1} for rigid-body association. This is valid in homogeneous solution and in the absence of electrostatic assistance, but not necessarily for immobilized molecules, heterogeneous solutions, or two-dimensional diffusion in membranes where fewer than six degrees of freedom are lost. Such situations require an appropriate treatment, yet we expect that the general model should still be applicable within the limits of the rigid-body approximation.

As in our previous study of affinity,[29] the rigid-body approximation is essential to our analysis of kinetics. Without it, protein-protein recognition becomes as complex a problem as protein folding. Of course, both phenomena have the same physicochemical origin. Both make use of the energy of the newly formed interactions between protein atoms (compensated in part by a loss of interactions with the solvent), to fight a loss of entropy (conformational and rotational-translational, compensated in part by an increased solvent entropy). The precomplex of our model is analogous from that point of view to the molten globule state, which is a proposed intermediate in folding,[32] and its funnel-like evolution toward the final product (Fig. 3) recalls current models of protein folding.[33] Yet, the rigid-body approximation makes recognition simple compared with folding. The approximation covers many systems, including protease-inhibitor and antibody-antigen complexes, and the barnase-barstar complex developed here. In these systems, the number of degrees of freedom is small, and we can make rate predictions that would be impossible if large conformation changes occurred.

Except at very high ionic strength, barnase-barstar association is much faster than the expected 10^5 M$^{-1} \cdot$ s^{-1}. Fast association, assisted by long-range electrostatic interactions, is probably as essential to the physiological function of barstar as its tight binding to barnase. With in vivo barstar concentrations in the micromolar range and $k_a = 10^5$ M$^{-1} \cdot$ s^{-1}, inhibition would require seconds or minutes, a time during which barnase would exert its RNAse activity inside the cell. Mutations of barstar that raise the dissociation constant above 10^{-9} M are lethal upon barnase overexpression.[14] A loss of affin-

Fig. 4. Electrostatic interactions and barnase-barstar association. The region of the barnase surface (**top left**) that surrounds the active site is positively charged and complementary to the negatively charged region of the barstar surface surrounding its active site (**top right**). The two surfaces form the interface of the complex (**bottom**). Note that, in making the figures on the top, barnase has been rotated 90° to the left and barstar 90° to the right, relative to their position in the complex. Molecular surfaces and electrostatic potentials are rendered with GRASP[38] by using atomic coordinates of the X-ray structure of the complex.[15] Negative potential is in red; positive is in blue. The dipole moment of barstar (blue arrow on bottom right) points away from the interface, whereas the dipole moment of barnase (blue arrow on top left) is approximately parallel to it. The intensity of either dipole evaluates to ~280 Debye units. In the calculation, Lys, Arg, Glu, and Asp side chains bear full charges, His is neutral, and peptide atoms bear partial charges. In solution, electrostatic interactions take place between charges and dipoles well before the two surfaces are in contact. In addition to stabilizing the complex, these interactions explain the very high rate of association observed at low and medium ionic strength.[7]

ity generally reflects an increased rate of dissociation and the loss of short-range interactions due to mutations at the interface. However, 10^{-9} M is only about one molecule per bacterial cell and barstar concentration is well above that. Thus, dissociation of the complex is not necessarily a sufficient explanation of the lethal effect, and there may well be cases in which slow association is the real cause.

In other systems, slow binding kinetics may be of lesser physiological consequence. Indeed, k_a values above those predicted by the Einstein-Smoluchowski equation are not common in protein-protein recognition. Association rates near 10^6 $M^{-1} \cdot s^{-1}$ are observed for BPTI binding to trypsin[34,35] and lysozyme binding to the monoclonal antibody D1.3.[36] In these two systems, the interface and geometric constraints are similar to barnase-barstar,[29] yet association is 10^3 times slower. Unlike barnase-barstar, the component proteins of trypsin-BPTI have net charges of the same positive sign, which implies electrostatic repulsion and $q_t < 1$ at medium ionic strength. Since k_a is still one order of magnitude above our reference value of 10^5 $M^{-1} \cdot s^{-1}$, q_r must be $\gg 1$. The substitution of Lys 15 at the active site of BPTI with a neutral alanine residue reduces k_a by a factor of 250, even though it makes the net charge of the molecule less positive.[35] Our interpretation is that dipole steering plays an important part in accelerating trypsin-BPTI association and that even though K15A BPTI binds trypsin 5 or 6 orders of magnitude slower than barstar does to barnase, the mechanism of association is the same. The effect of the K15A substitution on the dipole is unfavorable and large enough to overcome the favorable effect on the net charge. Because $q_t q_r$ is larger than 1 in the wild-type and less than 1 in the mutant, raising the ionic strength should slow wild-type BPTI binding to trypsin just as it does for barstar binding to barnase, but it might accelerate it for the K15A mutant.

One domain where fast association is often encountered is protein-DNA recognition. The case of the *lac* repressor-*lac* operator system has been analyzed in great detail by von Hippel and collaborators (see von Hippel and Berg[37] for a review). Electrostatic effects do not fully account for the observed k_a values, which can reach $5 \cdot 10^{10}$ $M^{-1} \cdot s^{-1}$, nor do they explain the rate dependence on the size of the DNA molecule. Proposed mechanisms have the protein sliding along the DNA or hopping from one segment of the double helix to another. In addition, large conformation change and disorder-order transitions affect both the protein and the DNA partner. These features have no equivalent in the systems analyzed here, which are much simpler and for which the minimum model that we propose appears to suffice.

ACKNOWLEDGMENTS

I thank Dr. J. Hermans (Chapel Hill, NC), M. Karplus (Cambridge, MA, and Strasbourg, France), and C. Robert of my lab for discussion and advice. This work was supported by the Université Paris-Sud and by Association pour la Recherche contre le Cancer.

REFERENCES

1. Janin, J., Chothia, C. The structure of protein-protein recognition sites. J. Biol. Chem. 265:16027–16030, 1990.
2. Janin, J. Protein-protein recognition. Prog. Biophys. Mol. Biol. 64:145–165, 1996.
3. Eigen, M., Hammes, G.G. Elementary steps in enzyme reactions. Adv. Enzymol. 25:1–39, 1963.
4. Berg, G.B., von Hippel, P.H. Diffusion-controlled macromolecular interactions. Annu. Rev. Biophys. Biophys. Chem. 14:131–160, 1985.
5. Schreiber, G., Fersht, A.R. Interaction of barnase with its polypeptide inhibitor barstar studied by protein engineering. Biochemistry 32:5145–5150, 1993.
6. Schreiber, G., Fersht, A.R. Energetics of protein-protein interactions: Analysis of the barnase-barstar interface by single mutations and double mutant cycles. J. Mol. Biol. 248:478–486, 1995.
7. Schreiber, G., Fersht, A.R. Rapid, electrostatically assisted association of proteins. Nat. Struct. Biol. 3:427–431, 1996.
8. Hartley, R.W. Barnase and barstar: Two small proteins to fold and fit together. Trends Biochem. Sci. 14:450–454, 1989.
9. Atkins, P.W. 'Physical Chemistry.' 4th edit. Oxford, UK: Oxford University Press, 1990.
10. Zhou, H.X. Brownian dynamics study of the influence of electrostatic interaction and diffusion on protein-protein association kinetics. Biophys. J. 64:1711–1716, 1993.
11. Northrup, S.H., Reynolds, J.C.L., Miller, C.M., Forrest, K.J., Boles, J.O. Diffusion-controlled association rate of cytochrome c and cytochrome c peroxidase in a simple electrostatic model. J. Am. Chem. Soc. 108:8162–8170, 1986.
12. Sines, J.J., Allison, S.S., McCammon, J.A. Point charge distributions and electrostatic steering in enzyme/substrate encounter: Brownian dynamics of modified copper/zinc superoxide dismutases. Biochemistry 29:9403–9412, 1990.
13. Getzoff, E.D., Cabelli, D.E., Fisher, C.L., Parge, H.E., Viezzoli, M.S., Banci, L., Hallewell, R.A. Faster superoxide dismutase mutants designed by enhancing electrostatic guidance. Nature 358:347–351, 1992.
14. Jucovic, M., Hartley, R.W. Protein-protein interaction: A Genetic selection for compensating mutations at the barnase-barstar interface. Proc. Natl. Acad. Sci. USA 93:2343–2347, 1996.
15. Guillet, V., Lapthorn, A., Hartley, R.W., Mauguen, Y. Recognition between a bacterial ribonuclease, barnase, and its natural inhibitor, barstar. Structure (Lond.) 1:165–177, 1993.
16. Buckle, M., Schreiber, G., Fersht, A.R. Crystal structural analysis of a barnase-barstar complex at 2.0 Å resolution. Biochemistry 33:8878–8889, 1994.
17. Mauguen, Y., Hartley, R.W., Dodson, E.J., Dodson, G.G., Bricogne, G., Chothia, C., Jack, A. Molecular structure of a new family of ribonucleases. Nature 29:162–164, 1982.
18. Lubienski, M., Bycroft, M., Freund, S.M.V., Fersht, A.R. ^{13}C assignment and three-dimensional structure of barstar using NMR spectroscopy. Biochemistry 33:8866–8877, 1994.
19. Janin, J. Principles of protein-protein recognition from structure to thermodynamics. Biochimie 7:497–505, 1995.
20. Northrup, S.H., Erickson, H.P. Kinetics of protein-protein association explained by Brownian dynamics computer simulations. Proc. Natl. Acad. Sci. USA 89:3338–3342, 1992.
21. Janin, J., Wodak, S.J. A reaction pathway for the quaternary structure change in hemoglobin. Biopolymers 24:509–526, 1985.

22. Cherfils, J., Duquerroy, S., Janin, J. Protein-protein recognition analyzed by docking simulation. Proteins 11:271–280, 1991.
23. Finkelstein, A.V., Janin, J. The price of lost freedom: Entropy of bimolecular complex formation. Protein Eng. 3:1–3, 1989.
24. Doucet, J., Benoît, J.P. Molecular dynamics studied by analysis of X-ray diffuse scattering by lysozyme crystals. Nature 325:643–647, 1987.
25. Erickson, H.P. Cooperativity in protein-protein association. The structure and stability of the actin filament. J. Mol. Biol. 206:465–474, 1989.
26. Tidor, B., Karplus, M. The contribution of vibrational entropy to molecular association. The dimerization of insulin. J. Mol. Biol. 238:405–414, 1994.
27. Page, M.I., Jenks, W.P. Entropic contribution to rate acceleration in enzymic and intramolecular reactions and the chelate effect. Proc. Natl. Acad. Sci. USA 68:1678–1683, 1971.
28. Novotny, J., Bruccoleri, R.E., Saul, F.A. On the attribution of binding energy in the antigen-antibody complexes McPC603, D1.3 and HyHEL-5. Biochemistry 28:4735–4748, 1989.
29. Janin, J. Elusive affinities. Proteins 21:30–39, 1995.
30. Miyamoto, S., Kollman, P.A. Absolute and relative binding free energy calculations of the interaction of biotin and its analogs with streptavidin using molecular dynamics/free energy perturbation approaches. Proteins 16:226–245, 1993.
31. Murphy, K.P., Xie, D., Thompson, K.S., Amzel, L.M., Freire, E. Entropy in biological binding processes: Estimation of translational entropy loss. Proteins 18:63–67, 1994.
32. Ptitsyn, O.B., Paine, R.H., Semitsonov, G.V., Zerovnik, E., Razgulyaev, O.I. Evidence for a molten globule state as a general intermediate in protein folding. FEBS Lett. 262:20–24, 1990.
33. Bryngelson, J.D., Onuchic, J.N., Socci, N.D., Wolynes, P.G. Funnels, pathways and the energy landscape in protein folding: A Synthesis. Proteins 21:167–195, 1995.
34. Vincent, J.P., Lazdunski, M. Trypsin-pancreatic trypsin inhibitor association. Dynamics of the interaction and the role of disulfide bridges. Biochemistry 11:2967–2977, 1972.
35. Castro, M.J., Anderson, S. Alanine point-mutations in the reactive region of bovine pancreatic trypsin inhibitor: Effects on the kinetics and thermodynamics of binding to β-trypsin and a-chymotrypsin. Biochemistry 35:11435–11446, 1996.
36. Hawkins, R.E., Russell, S.J., Baier, M., Winter, G. The contribution of contact and non-contact residues of antibody in the affinity of binding to antigen. The interaction of mutant D1.3 antibodies with lysozyme. J. Mol. Biol. 234: 958–964, 1993.
37. von Hippel, P.H., Berg, O.G. Facilitated target location in biological systems. J. Biol. Chem. 264:675–678, 1989.
38. Nicholls, A. GRASP: Graphical representation and analysis of surface properties. New York: Columbia University.