

How important are entropic contributions to enzyme catalysis?

J. Villà, M. Štrajbl, T. M. Glennon, Y. Y. Sham, Z. T. Chu, and A. Warshel*

Department of Chemistry, University of Southern California, Los Angeles, CA 90089

Communicated by Thomas C. Bruice, University of California, Santa Barbara, CA, August 21, 2000 (received for review April 3, 2000)

The idea that enzymes accelerate their reactions by entropic effects has played a major role in many prominent proposals about the origin of enzyme catalysis. This idea implies that the binding to an enzyme active site freezes the motion of the reacting fragments and eliminates their entropic contributions, $(\Delta S_{\text{cat}}^{\ddagger})'$, to the activation energy. It is also implied that the binding entropy is equal to the activation entropy, $(\Delta S_{\text{w}}^{\ddagger})'$, of the corresponding solution reaction. It is, however, difficult to examine this idea by experimental approaches. The present paper defines the entropic proposal in a rigorous way and develops a computer simulation approach that determines $(\Delta S^{\ddagger})'$. This approach allows us to evaluate the differences between $(\Delta S^{\ddagger})'$ of an enzymatic reaction and of the corresponding reference reaction in solution. Our approach is used in a study of the entropic contribution to the catalytic reaction of subtilisin. It is found that this contribution is much smaller than previously thought. This result is due to the following: (i) Many of the motions that are free in the reactants state of the reference solution reaction are also free at the transition state. (ii) The binding to the enzyme does not completely freeze the motion of the reacting fragments so that $(\Delta S^{\ddagger})'$ in the enzymes is not zero. (iii) The binding entropy is not necessarily equal to $(\Delta S_{\text{w}}^{\ddagger})'$.

Many prominent proposals (e.g., see refs. 1 and 2) and textbooks that consider biochemical systems (e.g., refs. 3 and 4) invoke entropic contributions as major factors in enzyme catalysis. These proposals, which are intuitively very appealing (e.g., see ref. 5), have assumed that the large configurational space available for the reacting fragments in water would be drastically restricted in the enzyme active site. It has been thus deduced that this should lead to large entropic contributions to the difference between the activation barrier in the enzyme and in the reference solution reaction. However, the validity of these proposals is far from being obvious (6, 7). For example, the very influential proposal introduced by Page and Jencks (1) reflects the assumption that the formation of the transition state in a bimolecular reaction in solution involves complete loss of three translational and three rotational degrees of freedom. However, two or more of these degrees of freedom are usually almost free in the transition state (see below). More serious is the implicit assumption that the entropic contribution to catalysis is given approximately by the negative of the binding entropy (see below). Other problems with simple estimates of the entropic contribution will be mentioned in the next section.

The main stumbling block for determining the validity of the entropic proposal is the absence of direct experimental information about the corresponding contribution of the reacting fragments to the activation entropy in the enzyme and in solution. In this respect it is interesting to note the recent analysis of cytidine deaminase by Wolfenden and co-workers (8). This study found that the entropies of activation in the enzyme and in water are very similar and that the overall catalysis is due to enthalpic effects. Interestingly, it was found that the activation entropy in water and the binding entropy are significant. However, it is not known what parts of these entropies are associated with the restriction of the motion of

the reacting fragments and what parts are associated with solvation entropies.

Intramolecular cyclizations of model compounds were used as “proofs” that enzyme catalysis involves large entropic contributions (e.g., ref. 2). However, recent simulations of Bruice and co-workers (9) indicated that the rate acceleration associated with smaller ring size is not necessarily associated with entropic contributions. Furthermore, our studies (6) have challenged the direct relevance of these model compounds to enzyme catalysis.

The entropic contributions to enzyme catalysis can be determined in principle by computer simulation approaches, but the development of the proper method is far from trivial because of enormous convergence problems. Nevertheless, some encouraging progress has been made in studies of related problems. This includes the progress in evaluating solvation entropies (10) and in evaluating binding entropies (11). Yet, no attempt has been made in evaluating activation entropies by simulation approaches. Note in this respect that the recent interesting studies reported by Kollman and co-workers (12, 13) have not provided ways for consistent evaluation of activation entropies or binding entropies.[†] At any rate, it is clear that calculations of activation entropies in enzyme active sites and in the corresponding reference solution reactions are essential for a quantitative understanding of entropic contributions to enzyme catalysis (6).

The present work introduces an approach for evaluating the contribution from the motion of the reacting fragments to the activation entropies of enzymatic reactions and uses this approach in assessing the importance of entropic effects to enzyme catalysis.

Defining the Problem

To evaluate the importance of a specific contribution to catalysis it is essential to define the relevant thermodynamic cycle. This is particularly important in considering entropy contributions, whose definition and estimates involve in many cases incomplete thermodynamic cycles. The starting point of the present study is the free energy diagram of Fig. 1. This diagram compares the activation free energies $\Delta G_{\text{p}}^{\ddagger}$ and $\Delta G_{\text{w}}^{\ddagger}$ for a given reaction in a protein (p) and in water (w) and also compares the corresponding activation free energies $(\Delta G^{\ddagger})'$ when the reacting fragments are frozen (see below). As shown in Fig. 1, $\Delta G_{\text{p}}^{\ddagger}$ is given by $\Delta G_{\text{bind}} + \Delta G_{\text{cat}}^{\ddagger}$, and thus reflects the binding energy ΔG_{bind} ,

Abbreviations: EVB, empirical valence bond; MD, molecular dynamics; FEP, free energy perturbation; RS, reactant state; TS, transition state; RRA, restrain–release approach.

*To whom reprint requests should be addressed at: E-mail: warshel@usc.edu.

[†]Ref. 12 used molecular dynamics (MD) runs to estimate the force constants necessary for keeping the system in the reactant conformation with a rather arbitrary standard deviation of 0.2 Å and 20° for distances and angular dependence, respectively. These force constants were used for deriving the entropy contributions by using equations derived by Hermans and Wang (11). However, this approach does not amount to the evaluation of the ground state entropy but rather to assuming the corresponding effective volume. Ref. 13 used gas phase entropy calculations to estimate the entropy cost of bringing the reactants in water to the corresponding configuration in the ES complex. However, the gas phase entropy may be drastically modified in solution. More importantly, the assumption (2) that the estimated entropy is equal to $-\Delta\Delta S_{\text{w}}^{\ddagger} \rightarrow_{\text{cat}}$ is not justified (see text).

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. §1734 solely to indicate this fact.

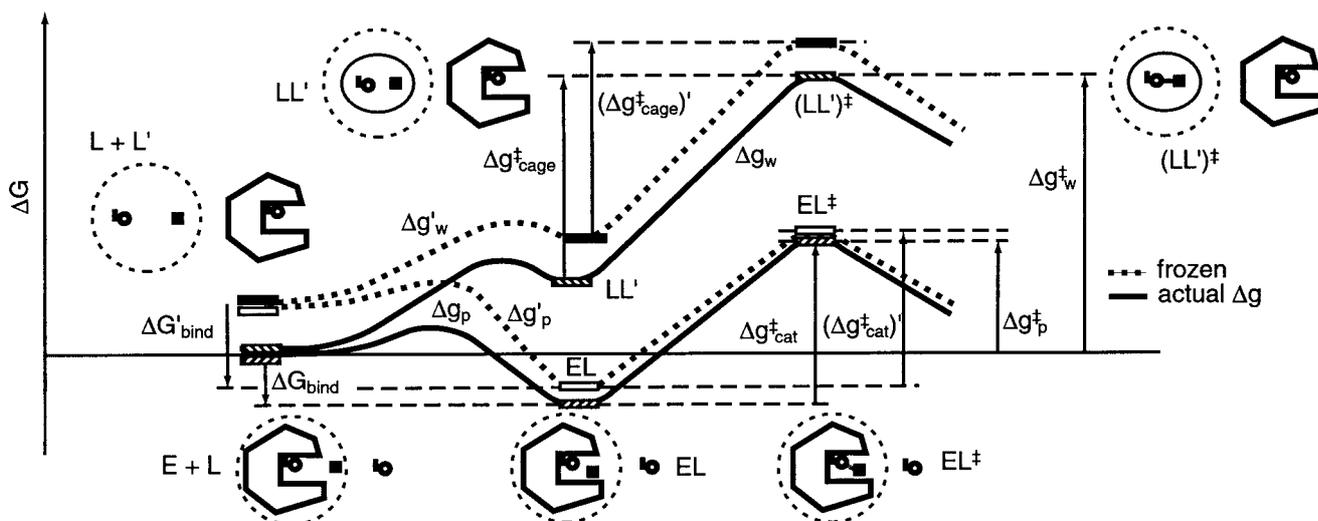


Fig. 1. Comparison of the free energy profiles Δg for a given reaction in a protein (Δg_p^\ddagger) and in water (Δg_w^\ddagger). The figure represents these free energy profiles in the absence (Δg) and presence ($\Delta g'$) of constraints for the movement of the reacting fragments. Note that Δg_p^\ddagger corresponds approximately to k_{cat}/K_m (it is given by $\Delta G_{bind} + \Delta g_{cat}^\ddagger$). The figure uses the notation of L for ligand rather than S for substrate to prevent confusion with the symbol S used for entropy.

whose nature is quite clear and is not involved in the catalytic puzzle (see also ref. 7). The main open question is the relationship between Δg_{cat}^\ddagger (which corresponds to k_{cat}) and the activation barrier of the reaction in water, Δg_w^\ddagger . It is also interesting, although not essential, to compare Δg_{cat}^\ddagger to the barrier Δg_{cage}^\ddagger obtained by correcting Δg_w^\ddagger for the free energy associated with bringing the reactants to a solvent cage, where each of the $(n - 1)$ reactants (from an overall n reactants) is at a relatively small distance from the central reactant (the n th reactant). At this distance the interaction with the central ligand is small and each reactant moves almost freely. In a case of small reactant molecules we can define the cage volume by requiring that the central reactant will have a standard 1 M concentration and the other reactants will have a concentration that is less than or equal to 55 M. In other words, taking into account the free energy contribution associated with confining a molecule that occupies a molar volume v_0 ($1,660 \text{ \AA}^3$ at 300 K) (6) to a volume v_{cage} we may write

$$\Delta g_w^\ddagger = \Delta g_{cage}^\ddagger - n_r RT \ln \left(\frac{v_{cage}}{v_0} \right), \quad [1]$$

where Δg_w^\ddagger is the activation barrier in solution at a standard state concentration, and v_{cage} is the volume of the solvent cage, which is taken to be larger than $1,660 \text{ \AA}^3/55$. Here n_r is the number of the ligands that become bound to the substrate at the transition state. This number is equal to $(n - 1)$ in a fully concerted reaction, whereas it reduces to 2 in a stepwise reaction. The seemingly arbitrary definition of Δg_{cage}^\ddagger (which will not change our final evaluation of Δg_w^\ddagger) reflects two considerations. First, early workers (e.g., refs. 1 and 2) implied or stated correctly that the free energy price associated with bringing the reacting fragments to a distance of a weak interaction is given by simple concentration considerations and is given approximately by $-RT \ln 55$. Thus the real puzzle was (and is) why Δg_{cat}^\ddagger is much smaller than the barrier obtained by considering the small volume correction to Δg_w^\ddagger . And second, comparing Δg_{cat}^\ddagger and Δg_{cage}^\ddagger provides conceptually and computationally closely related quantities, the activation barrier in the enzyme active site and the activation barrier in a solvent cage, where the reactants are in close proximity. It should also be stated that our cage system is very different than the system used by Jencks and others, where the reactants in solution are frozen in a space similar to that

available in the enzyme active site. This traditional definition makes it rather difficult to define a simple thermodynamics cycle for the comparison of Δg_{cat}^\ddagger and Δg_w^\ddagger .

Now, the major question one has to address is related to the entropy contribution associated with the restriction of the motions of the reacting fragments by the enzyme active site, relative to the corresponding contribution in the solvent cage. This effect does not include the contribution from the change in the entropy of the environment (e.g., solvent), since the early proposals were defined clearly in terms of the substrate entropy rather than the entropy of the surroundings. Thus, we are not addressing here the change in the so called cratic entropy (11) that includes the overall change in entropy upon binding, but rather the changes in the contribution, $\Delta S'$, from the configurational space of the reacting fragments. The factors involved in the catalytic effect of $\Delta S'$ are illustrated in Fig. 1. The figure compares the free energy profile Δg_w and the corresponding $(\Delta g_w)'$, obtained by freezing the degrees of freedom of the substrate that are perpendicular to the reaction coordinate. Similarly, the figure compares Δg_p and $(\Delta g_p)'$. As clarified above, we are interested in the contribution of $\Delta S'$ to $(\Delta g_{cat}^\ddagger - \Delta g_w^\ddagger)$ or to $(\Delta g_{cat}^\ddagger - \Delta g_{cage}^\ddagger)$. Using Fig. 1, we obtain:

$$\begin{aligned} -T(\Delta \Delta S_{cage \rightarrow cat}^\ddagger)' &= -T(\Delta S_{cat}^\ddagger - \Delta S_{cage}^\ddagger)' \\ &= -T[(S_{cat}^{TS} - S_{cage}^{TS})' - (S_{cat}^{RS} - S_{cage}^{RS})'], \end{aligned} \quad [2]$$

where RS and TS, respectively, designate the corresponding contributions from the reactants state and transition state in the protein or in the solvent cage. Of course we can also obtain easily $(\Delta \Delta S_{w \rightarrow cat}^\ddagger)'$.

The first point to note from Eq. 2 is that one must consider the entropy contributions from the RS and TS both in the enzyme and in the solvent cage. To the best of our knowledge all previous estimates, with the exception of ref. 6, have not considered both the RS and TS and have not defined the catalytic effect in a rigorous way. Probably the most systematic attempt was made by Jencks (page 720 of ref. 2), who defined the problem by considering the hypothetical case with identical ΔH^\ddagger in the enzyme and in water. He assumed that ΔS_{cat}^\ddagger is zero and estimated ΔS_w^\ddagger by considering the process of bringing the

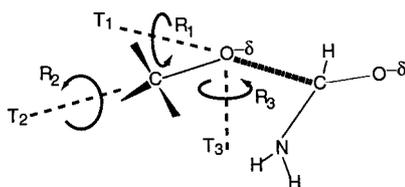


Fig. 2. Schematic representation of the translational and rotational degrees of freedom in the nucleophilic attack of a CH_3O^- group on an amide. The figure depicts only the transition state, but the corresponding RS picture can be obtained by stretching the $\text{C}\cdots\text{O}$ distance. For simplicity we place the rotation axes on the attacking oxygen. As illustrated, R_2 remains a free rotation in the TS, R_1 becomes a low-frequency torsional oscillator, while T_1 , T_2 , and T_3 become bending motions.

reactants in water to the same frozen position that they have in the enzyme. Furthermore, the TS in water was also assumed to be frozen in the same way. Basically, it was assumed implicitly that $S_w^{\text{TS}} = S_p^{\text{TS}} \approx 0$, and that the activation entropy in water can be approximated by the binding entropy of the protein. In other words, Jencks as well as most other workers implied that we can freeze the motion of the reactants in solution in the same restricted space available for these motions in the protein and that the free energy that should be invested in this configurational restriction is the entropic contribution to the difference between Δg_w^\ddagger and $\Delta g_{\text{cat}}^\ddagger$. This (apparently incorrect assumption) can be formulated by the more accurate terms of Fig. 1 as:

$$-T\Delta\Delta S_w^{\ddagger \rightarrow \text{cat}} \approx -T[(S_{\text{cage}}^{\text{RS}})'] - (S_{\text{EL}}^{\text{RS}})'] - n_r RT \ln \left(\frac{v_{\text{cage}}}{v_0} \right), \quad [3]$$

where $S_{\text{cage}}^{\text{RS}}$ is the hypothetical frozen RS of the solution reaction. Here we subtract $-T(S_{\text{EL}}^{\text{RS}})'$ from $-T(S_{\text{cage}}^{\text{RS}})'$, because the substrate is not completely frozen in the protein. However, as seen from Fig. 1, this is a problematic assumption in large part because of the fact that S' is not zero at the TS of the solution reaction. Or in other words because of the fact that many of the motions that are free in the RS in the solvent cage are also free in the TS in this cage.

To further clarify this point we can consider the hypothetical case when all the motions are frozen in the enzyme in both the RS and the TS (so that $(\Delta S_{\text{cat}}^{\ddagger})' = 0$) and all the motions are free in solution in the RS and in the TS (so that $(\Delta S_{\text{cage}}^{\ddagger})' = 0$). In this case we have $(\Delta\Delta S_{\text{cage} \rightarrow \text{cat}}^{\ddagger})' = 0$ and $\Delta\Delta S_w^{\ddagger \rightarrow \text{cat}} \approx 8$ entropy units. On the other hand, if we only consider the RS in water and in the protein site we may obtain a large $(\Delta\Delta S_{\text{cage} \rightarrow \text{cat}}^{\ddagger})'$.

Finally, it might be useful to comment on the inherent assumption of Page and Jencks (1) that the formation of a bimolecular complex involves the loss of three translational and three rotational degrees of freedom. This model was based on considering two spherical fragments. However, in actual formation of the cage complex we still retain between one to three of these degrees of freedom. For example, in the attack of a CH_3O^- group on an amide we do not lose the rotation around the C–O axis, and two of the translational motions of the CH_3 group are almost free (see Fig. 2).

The above discussion should not be viewed as a list of problems with previous estimates but as a reminder that the problem of evaluating entropic contributions to enzyme catalysis has not been solved. It seems to us that this challenging problem cannot be resolved without some form of computer simulation approach.

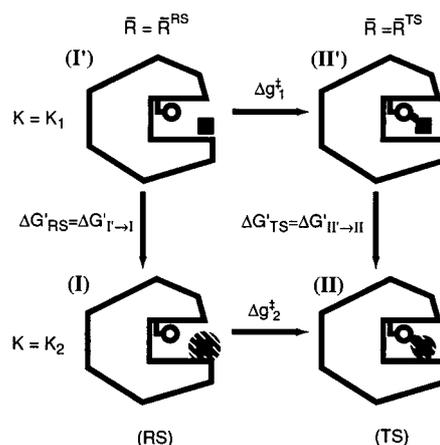


Fig. 3. Thermodynamic cycle used for the evaluation of the entropy contribution to the activation free energy of the reaction. The fragments are fixed in (I') and (II') and allowed to move (as indicated by the shaded area) in (I) and (II). The same circle is also used for the reaction in water (see figure 1 of ref. 14).

Methods

As clarified in the previous section, our task is to evaluate the entropic contribution of the reacting fragments to $\Delta\Delta g_w^{\ddagger \rightarrow \text{cat}}$. To do so we developed a restrain–release approach (RRA) which is related in some aspects to the approach developed by Hermans and co-workers (11) for studies of binding entropies. Our approach has been described in detail in ref. 14, where it was used for studies of the activation entropy of amide hydrolysis in water. Here we discuss only the main points of this approach, which is based in the thermodynamic cycle of Fig. 3. This cycle considers the activation free energy, Δg^\ddagger , for the given reaction in the given system (the enzyme or solvent cage) in two limiting conditions. In the first case (the upper part of the cycle) the reacting system is transformed from the RS to the TS along a (unspecified) reaction coordinate while a restraint is used to minimize the available configuration space in the direction perpendicular to this reaction coordinate. In this case the activation free energy, Δg_1^\ddagger , does not involve the entropic contributions of the solute, $(\Delta S_1^\ddagger)' = 0$, since the corresponding motions are frozen. In the second case (the lower part of the cycle) the reacting fragments are free to move so that the corresponding activation barrier, Δg_2^\ddagger , includes the entropic contributions of the solute. Thus, the difference between the two Δg^\ddagger values gives the desired $-T(\Delta S^\ddagger)'$. The possible enthalpic contribution can be minimized with the proper selection of the initial conditions for the simulation (see below). Using the thermodynamic cycle, we can obtain $(\Delta g_2^\ddagger - \Delta g_1^\ddagger)$ from $\Delta G'_{\text{RS}}$ and $\Delta G'_{\text{TS}}$. This can be done by imposing a strong position restraint in both states I' and II' and evaluating the corresponding free energies $\Delta G'_{\text{RS}} = \Delta G'_{\text{I}' \rightarrow \text{I}}$ and $\Delta G'_{\text{TS}} = \Delta G'_{\text{II}' \rightarrow \text{II}}$ associated with the release of these constraints (see below). In this way we can write

$$-T(\Delta S^\ddagger)' = -T(\Delta S_2^\ddagger)' \approx \Delta g_2^\ddagger - \Delta g_1^\ddagger \approx \Delta G'_{\text{TS}} - \Delta G'_{\text{RS}}. \quad [4]$$

The enthalpic contribution to $\Delta G'$ will be discussed below. Now, the practical evaluation of the $\Delta G'$ values involves the introduction of restrain potentials of the form

$$U_{\text{rest},j}^N = \frac{K_j}{2} \sum_i (R_i^N - \bar{R}_i^N)^2, \quad [5]$$

where i runs over the substrate coordinates and $\bar{\mathbf{R}}^N$ are reference coordinates that define the minimum of the restrain potential

(see below) at the given state ($N = I$, or $N = II$ for the RS and TS, respectively). The reference coordinates $\bar{\mathbf{R}}^N$ are evaluated by running molecular dynamics (MD) relaxation runs on the RS and the TS with $K = 0$ (see also below). Different conditions for the MD runs can generate different values of $\bar{\mathbf{R}}^N$, and the selection of the optimal coordinates as well as the implications of this selection will be discussed below.

The constraint release free energies ($\Delta G'$) are evaluated by a free energy perturbation (FEP) approach, where we use a mapping potential of the form

$$U_m^N = (1 - \lambda_m)U_{\text{rest},1}^N + \lambda_m U_{\text{rest},2}^N + \sum_{F=1}^{f-1} \frac{1}{2} K_{\text{cage}} (R_c^F - \bar{R}_c^F) + E, \quad [6]$$

where λ_m is changed from 0 to 1 in n increments and E designates the unconstrained potential surface of the system. R_c^F is the distance between an atom of the central fragment $F = 1$ and an atom of the F th fragment. The K_{cage} term is needed to prevent divergence when $K_2 \rightarrow 0$. The value of R_c^F is chosen in a way that the F th fragment can move freely around the central ligand while still being close to this ligand. K_{cage} is chosen so that v_{cage} will correspond to 55 M concentration.

Now we can write

$$-T(\Delta S^\ddagger)'_0 = \Delta G'_{\text{TS}} - \Delta G'_{\text{RS}} - n_r RT \ln \left(\frac{v_{\text{cage}}}{v_0} \right), \quad [7]$$

where $(\Delta S^\ddagger)'_0$ designates the entropic contribution in the 1 M standard state and the last term in Eq. 7 gives the free energy associated with the change of the cage volume from v_{cage} to v_0 .

The results of the thermodynamic cycle of Fig. 3 depend in principle on the chosen $\bar{\mathbf{R}}$ and may involve, in addition to $\Delta S'$, contributions from the enthalpy of the reacting fragments and from their solvation by the surrounding environment. To extract $\Delta S'$ from the total $\Delta G'$ we have to formulate the cycle of Fig. 3 in a more rigorous way. This is done by the approach presented in the appendix of ref. 14, which considers the free energy $\Delta G'(\bar{\mathbf{R}})$ of the system as a function of the restraint coordinates of the solute. Using the quadratic expansion of this potential of mean force (PMF) and evaluating the corresponding contributions has established that $\Delta G'(\bar{\mathbf{R}}_0)$ (where $\bar{\mathbf{R}} = \bar{\mathbf{R}}_0$ is the $\bar{\mathbf{R}}$ that minimizes the PMF) gives the desired $\Delta S'$ without additional enthalpic contributions. Although the minimization of this PMF is impractical at present, we can estimate the exact result for $\Delta S'$ by performing several random simulations with different $\bar{\mathbf{R}}^N$ and selecting the one that gives the minimum value of $|\Delta G'|$. In other words, we can use the lower bound (14)

$$|-T\Delta S'_N(\bar{\mathbf{R}}_0^N)| = |\Delta G'_N(\bar{\mathbf{R}}_0^N)| \leq |\Delta G'_N(\bar{\mathbf{R}}^N)| \quad [8]$$

and estimate the value of $|-T\Delta S'_N(\bar{\mathbf{R}}_0^N)|$ by finding the smallest $|\Delta G'_N(\bar{\mathbf{R}}^N)|$. With $\Delta S'_N(\bar{\mathbf{R}}_0^N)$ we can obtain $(\Delta S^\ddagger)'$ by using (see appendix of ref. 14)

$$-T(\Delta S^\ddagger)' = \Delta G'_{\text{TS}}(\bar{\mathbf{R}}_0^N) - \Delta G'_{\text{RS}}(\bar{\mathbf{R}}_0^N). \quad [9]$$

Eqs. 8 and 9 and the treatment in ref. 14 outline the formal requirements for rigorous evaluation of $(\Delta S^\ddagger)'$. However, because the evaluation of $\Delta S'$ with random $\bar{\mathbf{R}}$ is very expensive we confine the present study to the evaluation of the values of $\Delta S'_N$ at several values of $\bar{\mathbf{R}}$, obtained by running MD simulations on the corresponding potential surface with zero restrain potential. We then use the $\bar{\mathbf{R}}$ values that give the lowest $|\Delta G'|$ in evaluating $\Delta S'$.

To evaluate activation entropies it is essential to perform long MD or Monte Carlo (MC) samplings of the corresponding potential surfaces. Such potential surfaces should provide a reliable representation of the relevant reactions and also allow for sufficient sampling at a reasonable computer time. At present, it is impossible to satisfy this requirement with a high-level *ab initio* potential surface because this would require an enormous amount of computer time. Approaches that obtain the solute potential surface by fitting it to the corresponding *ab initio* gas phase surface and then consider the interaction of the gas phase charges with the solvent can provide reasonable estimates of the solvent contribution to the activation entropies in reactions that involve relatively small charge separation. However, it is hard to use such approaches for evaluation of the solute contribution to the activation entropy. This is particularly true when the solute is constrained to move along the gas phase reaction coordinate (e.g. ref. 15) and the substrate fluctuations in the directions perpendicular to the reaction coordinate are not taken into account. In our opinion the optimal strategy is to use empirical valence bond (EVB) potential surfaces (6). These surfaces can be fitted to reliable *ab initio* surfaces and to relevant experimental information and then to provide a consistent and reliable description of the solution reaction. This description includes a consistent incorporation of the solvent in the solute Hamiltonian, consistent description of the motion of the reacting fragments, and an analytical representation of the given surface that allows one to perform very extensive MD simulations. The EVB method was described extensively elsewhere (6) and has been used by many other groups (see ref. 14 and references therein). Thus we will mention here only several crucial points.

The EVB approach describes the reacting system by mixing several diabatic states that represent the reactants, products, and crucial intermediates. The energy of each diabatic state i (ε_i) is described by a force-field like potential function and the off-diagonal terms (H_{ij}) are described by simple analytical functions. The parameters in the ε_i and H_{ij} values are determined by using both experimental information and *ab initio* calculations. To perform our RRA calculations we have to confine the reactants

Table 1. $\Delta G'$ (in kcal/mol) for the hydrolysis of the Tyr-Gly dipeptide in water and in the active site of subtilisin

System	$\Delta G'_{\text{RS}}$			$\Delta G'_{\text{TS}}$			$-T(\Delta S^\ddagger)'_0$
Water	-91.9	-103.7	-89.6	-95.7	-87.0	-99.3	2.6 + 2.4
	-90.0	-98.6	-93.8	-134.1	-91.7	-101.7	
	-98.9	-96.0	-100.2	-105.0	-99.2	-102.7	
Protein	-102.3	-103.8	-113.0	-103.5	-98.6	-96.6	2.5
	-100.6	-100.1	-92.9	-90.4	-95.9	-100.0	
	-103.0	-106.6	-103.4	-95.4	-98.4	-105.3	

The calculations were done with $r_a = 14 \text{ \AA}$, and the results reported were evaluated with different $\bar{\mathbf{R}}$ values, obtained by the nine relaxation runs described in the text. The underlined values in boldface are the lowest absolute values obtained in each case and thus the optimal estimate of the corresponding $-T\Delta S'$ (see text). $-T(\Delta S^\ddagger)'_0$ is obtained from Eq. 7 with the optimal $\Delta G'$ and a 2.4 kcal/mol volume correction for the water case.

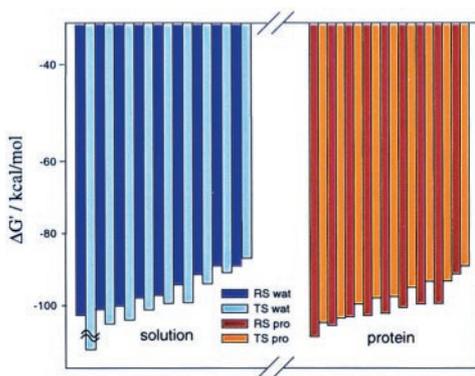


Fig. 4. $\Delta G'$ values at different initial conditions: in water (Left), and in the protein site (Right). Black and gray bars correspond to $\Delta G'$ in the reactant state and in the transition state, respectively. The figure represents the results of Table 1 for the simulations with region II radius of 14 Å, and the calculated values have been plotted in decreasing order of $|\Delta G'|$.

to the specified region of the potential surface (e.g., RS and TS). Here we limit our discussion to the common case where the system is represented by three diabatic states. In the case of a stepwise mechanism it is reasonable to use (14)

$$\begin{aligned} E_{RS} &\approx \varepsilon_1 \\ E_{TS} &\approx \lambda_1^{\text{TS}} \varepsilon_1 + \lambda_2^{\text{TS}} \varepsilon_2 + \lambda_3^{\text{TS}} \varepsilon_3, \end{aligned} \quad [10]$$

where the λ s are the values of the FEP mapping parameters that bring the system to the TS region.

Results and Discussion

The present study considered as a test case the catalytic reaction of subtilisin. The potential surface of this reaction was evaluated by the previously developed EVB potential surface (16). The three resonance structures that roughly correspond to the three states in the mechanism of action of subtilisin (reactants, intermediate, and products) are described in figure 6 of ref. 16. The parameters for this EVB potential energy surface are similar to those used in ref. 16. The rate-limiting step of the reaction is the attack of the O_γ of Ser-221 on the carbonyl carbon of the substrate. Thus, the entropy calculations have been done considering the RS and the TS, whose ground state energies E_g^{RS} and E_g^{TS} were obtained, respectively, by using Eq. 10 and $\lambda_n^{\text{TS}} = (0.0, 0.4, 0.6)$, as described above. Within this model the system is divided into several regions (see ref. 17 for a detailed discussion). The reaction region (region I) is surrounded by a protein + solvent sphere (region II, with a radius r_a) that is surrounded by external regions with special polarization and position boundary conditions (17). The protein atoms in these external regions are held by strong constraints. All the calculations were done with the ENZYMIK package (17).

The simulations involved the release of the position restraints in 5 FEP stages, changing K_j from 100 to 30, from 30 to 3, from 3 to 0.3, from 0.3 to 0.03, and from 0.03 to 0.003, where all the values of K are given in kcal/(mol·Å²). Each of these stages involved 11 mapping steps, and each of these steps followed an initial relaxation of the given system. The total simulation time for these five stages was 1,100 ps for both the water and protein runs. K_{cage} was taken as 0.4 kcal/mol (which corresponds to v_{cage} of 1,660/55 Å³) and R_c^f were taken as 3.0 and 3.1 for the O_γ -C distance and the O_γ -N_{e2} distance, respectively. To generate the different $\bar{\mathbf{R}}$ s needed for obtaining the optimal $\Delta G'$ we use different relaxation runs with different constraints on region I and II. Overall, we generated nine different $\bar{\mathbf{R}}$ s for the RS and nine for the TS. Next we performed RRA calculations for the different $\bar{\mathbf{R}}$ s. Table 1 gives the results of the simulations with $r_a =$

14 Å. Additionally, we depict in Fig. 4 the values of the $\Delta G'$ in an increasing order. As is apparent from the figure, we obtain a reasonable convergence in the selection of minimal values of $|\Delta G'|$. The differences between $\Delta G'_{\text{TS}}$ and $\Delta G'_{\text{RS}}$ (plus the volume correction of Eq. 7) yield the final values of $-T(\Delta S^\ddagger)_0'$ that are given in the last column of Table 1. As seen from the table, the entropic contribution of the reacting fragments is not so different in the enzyme and in the solvent cage. More specifically, the simulations give a difference of $(2.5 - (2.6 + 2.4)) = -2.5$ for $-T((\Delta S_{\text{cat}}^\ddagger)' - (\Delta S_{\text{w}}^\ddagger)')$. This estimate already includes the effect of K_{cage} (≈ -2.4 kcal/mol). Note that in this volume correction n_r is taken as 1, because the solution reaction has been found to involve a stepwise mechanism (18). The present estimate of $-T(\Delta \Delta S^\ddagger)'$ involves a few kcal/mol error and is likely to change with a more complete convergence (see footnote §). The main point is, however, that $|T(\Delta \Delta S^\ddagger)_0'|$ is much smaller than previously thought. The reason for this exciting finding is rationalized in Fig. 5, which shows snapshots of the last FEP frame of the simulation where K changes from 0.003 to 0.0003 kcal/(mol·Å²). As seen in the figure the degree of movement for each system (cage and protein) is similar in RS and in TS. If this conclusion will hold in other systems, which is likely (see below), then entropic effects do not contribute in a major way to the reduction of $\Delta g_{\text{cat}}^\ddagger$.

Concluding Remarks

This work examined the importance of the entropic contribution of the reacting fragments in enzyme catalysis. This was done by developing a computer simulation method capable of evaluating $(\Delta S^\ddagger)'$ in enzyme and in solution, and using this method in calculations of $(\Delta \Delta S^\ddagger)'$ for the catalytic reaction of subtilisin. It was found that $(\Delta \Delta S^\ddagger)'$ is much smaller than commonly assumed and should not be approximated by $T(\Delta S_{\text{bind}}^\ddagger)'$.

Our finding of a rather small $(\Delta \Delta S^\ddagger)'$ is consistent with a study (19) of nucleophilic displacements on phenyl esters in water. It was shown that dividing $-T\Delta S_{\text{w}}^\ddagger$ by the kinetic order of the reaction (first, second, and third) gives 4 to 5 kcal/mol. This finding was interpreted to indicate that bringing each species to the TS does not involve more than 5 kcal/mol entropic contribution (19, 20). Since the solvent contributions (14) to both $-T\Delta S_{\text{w}}^\ddagger$ and $-T(\Delta S_{\text{cat}}^\ddagger)'$ are positive, $|-T(\Delta \Delta S^\ddagger)'|$ should be quite small.

Obviously one may wonder what is the origin for the difference between the present results and previous estimates. One of the reasons that were alluded to in the Introduction is the fact that many of the motions that are free at the RS are also free at the TS. Another important factor is associated with the fact that $(\Delta S_{\text{cat}}^\ddagger)'$ is not zero in the protein. Finally, as was noted in the Introduction (see Fig. 2), the early estimate of Page and Jencks

§ To examine the convergence we performed additional calculations with r^a of 13 Å and 16 Å. For $r_a = 13$ Å we obtained ΔG_{TS} (in kcal/mol) -86 and -85 for the water case and -89 and -87 in the protein case. These results are similar to those obtained with $r_a = 15$ Å. On the other hand, in the 16-Å case we obtained for $-T(\Delta \Delta S^\ddagger)'$ more positive value (by about 9 kcal/mol) than for the smaller radius. This is not a major problem, because it represents a well-understood convergence problem and because it still leads to the same conclusion as obtained from the small r_a . That is, while the natural tendency is to assume that the results obtained with larger r_a are more reliable, we are limited here by enormous convergence problems that are reduced significantly for smaller r_a . Moreover, the incomplete convergence does not present a serious problem, because our main point is that $|\Delta \Delta g_{\text{w} \rightarrow \text{cat}}^\ddagger|$ is small. By reducing the motions in the protein we obtain a lower limit for $-T(\Delta S_{\text{cage}}^\ddagger)'$ and an upper limit for the entropic contribution to catalysis. Thus, we can focus on the results obtained with $r_a \leq 14$ Å. We also examined the convergence of the simulations (rather than the search of the optimal $\bar{\mathbf{R}}$). This was done by examining the change in the results in the last 550 ps. It was found that the results become more negative for both the RS and TS. However, $(\Delta S_{\text{cat}}^\ddagger)'$ and $(\Delta S_{\text{cat}}^\ddagger)'$ changed by only 0.5 and 0.2 kcal/mol, respectively. In general, we expect that with more complete convergence $\Delta S_{\text{w}}^\ddagger$ will become negative by a few kcal/mol.

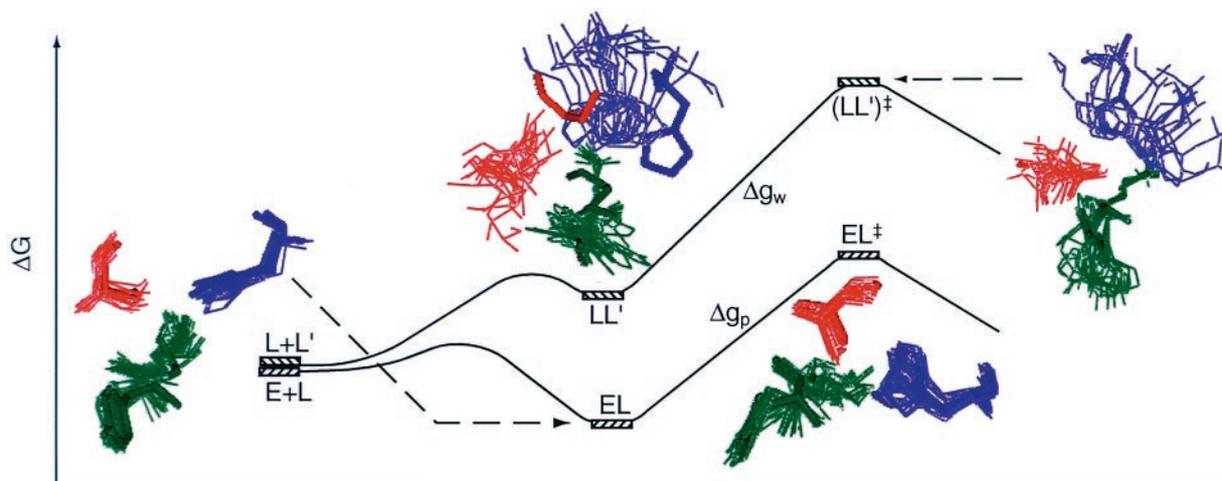


Fig. 5. Superposition of snapshots of the trajectories propagated at the last point of the FEP protocol where K changes from 0.003 to 0.0003 kcal/mol for the RS and the TS of the water surface (solid curve) and the protein surface (dashed curve). The free energy profiles are given in a schematic way. The 20 snapshots for each system were taken at equal time space during 50-ps runs. L designates ligand as in Fig. 1.

(1) provides an overestimate of the entropy loss upon binding, where some rotational and translational motions remain free even in the solvent cage.

The importance of entropic effects in enzyme catalysis has been frequently deduced from the trend in the rates of intramolecular cyclization reaction, where the formation of smaller rings is associated with larger rate constants. Recent studies (9) have suggested, however, that this trend may be due to RS destabilization enthalpic effects. This does not mean that such enthalpic effects play a significant role in enzyme catalysis. In fact, it is not clear that intramolecular cyclization reactions are directly relevant to most enzymatic reactions (6). At any rate, the present approach may provide a powerful way of exploring RS enthalpic contributions. That is, if instead of using the optimal \mathbf{R} for both the enzyme and the solution reaction we use the $\bar{\mathbf{R}}$ of the enzyme reaction for the water reaction we should be able to evaluate the enthalpic contribution associated with the distortion of the structure of the substrate by the enzyme.

The present work is an attempt to obtain a quantitative estimate of the entropic contribution of the reacting fragments to enzyme catalysis. The finding that these contributions are much smaller than previous estimates as well as simulations of

many enzymatic reactions is consistent with the view that electrostatic effects are the major factor in enzyme catalysis. As argued repeatedly before (*e.g.*, refs. 6 and 7), the electrostatic effects are due to the preorganized polar environment of the enzyme active site and to the relatively small reorganization free energy of this environment. The ability of the enzyme to provide smaller reorganization free energy than water does involves, of course, some entropic effects. These effects are *not*, however, the entropic effects introduced in all early proposals of entropic contributions to enzyme catalysis. That is, previous proposals were based on the assumption that the enzyme decreases the entropy of the reacting fragments. On the other hand, our electrostatic proposal involves the reduction of the reorganization free energy (which includes some entropic contributions) by folding the enzyme to its catalytic configuration. This contribution of the enzyme–enzyme rather than the enzyme–substrate interaction is the most important factor in enzyme catalysis.

We thank Dr. Chienyu Jen and Dr. Jan Florián for insightful discussions. J.V. acknowledges European Molecular Biology Organization Fellowship ALTF 509-1998. This work was supported by National Institutes of Health Grant GM24492.

- Page, M. I. & Jencks, W. P. (1971) *Proc. Natl. Acad. Sci. USA* **68**, 1678–1683.
- Jencks, W. P. (1986) *Catalysis in Chemistry and Enzymology* (Dover, New York).
- Kyte, J. (1995) *Mechanism in Protein Chemistry* (Garland, Hamden, CT).
- McMurry, J. & Castellion, M. E. (1999) *Fundamentals of Organic and Biological Chemistry* (Prentice–Hall, Englewood Cliffs, NJ).
- Blow, D. (2000) *Structure* **8**, R77–R81.
- Warshel, A. (1991) *Computer Modeling of Chemical Reactions in Enzymes and Solutions* (Wiley, New York).
- Warshel, A. (1998) *J. Biol. Chem.* **273**, 27035–27038.
- Snider, M. J., Gaunitz, S., Ridgway, C., Short, S. A. & Wolfenden, R. (2000) *Biochemistry* **39**, 9746–9753.
- Bruice, T. C. & Lightstone, F. C. (1999) *Acc. Chem. Res.* **32**, 127–136.
- Levy, R. M. & Gallicchio, E. (1998) *Annu. Rev. Phys. Chem.* **49**, 531–567.
- Hermans, J. & Wang, L. (1997) *J. Am. Chem. Soc.* **119**, 2707–2714.
- Stanton, R., Peräkylä, M., Bakowies, D. & Kollman, P. A. (1998) *J. Am. Chem. Soc.* **120**, 3448–3457.
- Kuhn, B. & Kollman, P. A. (2000) *J. Am. Chem. Soc.* **122**, 2586–2596.
- Štrajbl, M., Sham, Y. Y., Villà, J., Chu, Z. T. & Warshel, A. (2000) *J. Phys. Chem. B* **104**, 4578–4584.
- Chandrasekhar, J. & Jorgensen, W. (1985) *J. Am. Chem. Soc.* **107**, 2974–2975.
- Warshel, A., Sussman, F. & Hwang, J.-K. (1988) *J. Mol. Biol.* **201**, 139–159.
- Lee, F. S., Chu, Z. T. & Warshel, A. (1993) *J. Comput. Chem.* **14**, 161–185.
- Štrajbl, M., Florián, J. & Warshel, A. (2000) *J. Am. Chem. Soc.* **122**, 5354.
- Bruice, T. C. & Benkovic, S. J. (1964) *J. Am. Chem. Soc.* **86**, 418–426.
- Bruice, T. C. (1970) *The Enzymes*, ed. Boyer, P. D. (Academic, New York), 3rd Ed., Vol. 2, pp. 217–279.