Calculations of Activation Entropies of Chemical Reactions in Solution

ARTICLE *in* THE JOURNAL OF PHYSICAL CHEMISTRY B · MAY 2000

Impact Factor: 3.3 · DOI: 10.1021/jp0003095

5 AUTHORS, INCLUDING:

Marek Strajbl
myinst
18 PUBLICATIONS **827** CITATIONS
SEE PROFILE

Yuk Y Sham
University of Minnesota Twin Cities
60 PUBLICATIONS **2,186** CITATIONS
SEE PROFILE

Jordi Villà-Freixa
University of Vic - Central University of Catal…
90 PUBLICATIONS **1,995** CITATIONS
SEE PROFILE

Zhen T Chu
University of Southern California
30 PUBLICATIONS **2,192** CITATIONS
SEE PROFILE

Calculations of Activation Entropies of Chemical Reactions in Solution

Marek Strajbl,[†] Yuk Y. Sham, Jordi Villà, Z.-T. Chu, and Arieh Warshel*

Department of Chemistry, University of Southern California, Los Angeles, California 90089-1062

Received: January 26, 2000; In Final Form: March 10, 2000

The elucidation of the role of entropic effects in enzyme catalysis is a problem of practical and fundamental interest. In order to address this problem it is essential to develop simulation methods capable of evaluating the entropic contribution, $(\Delta S^{\ddagger})'$, of the reacting fragments to the total activation entropy, $\Delta S^{\ddagger}$. In fact, the general ability to evaluate activation entropies of chemical reactions in solution has long been a challenge to computational chemists. The present work develops and examines a method for evaluation of $(\Delta S^{\ddagger})'$ and $\Delta S^{\ddagger}$. This method introduces a thermodynamic cycle that considers the transformation between the reactants state (RS) and the transition state (TS) in two paths. In the first path the reacting fragments are constrained to move along a single reaction coordinate while in the second path they are allowed to move also in the subspace perpendicular to the reaction coordinate. The difference between the activation barriers that correspond to the two paths provides the desired $-T(\Delta S^{\ddagger})'$. The cycle also involves two steps where a Cartesian restraint that fixes the reacting fragments in the RS and TS is released. The free energies, $\Delta G''$s, associated with these restraint release steps are used to complete the thermodynamic cycle and to provide the actual estimate of $(\Delta S^{\ddagger})'$. This estimate is optimized by using different initial conditions and by selecting the smallest value of $|\Delta G'|$. The solvent contributions to the activation entropy are evaluated using a newly developed version of the Langevin dipole model. The potential surfaces used in the present work are obtained by the empirical valence bond (EVB) method. This method provides analytical yet reliable potential surfaces that reflect properly the motions of the reacting fragments and the coupling between these motions and the solvent polarization. The analytical representation of the EVB surfaces allows us to perform the extensive sampling necessary in order to obtain converging results. Our method is examined by evaluating the activation entropy for the hydrolysis of formamide in water. It is found that the calculated activation entropies reach convergence in a reasonable computer time and that the agreement between the calculated and observed results is reasonable. This method can be easily implemented in studies of enzymatic reactions and helps in assessing the importance of entropic effects in enzyme catalysis.

I. Introduction

Many chemical reactions involve changes in the configurational space available for the reacting fragments and in the corresponding activation entropies. The evaluation of such entropic effects is of major interest in view of their assumed role in enzyme catalysis.[1] That is, many prominent proposals for the origin of enzyme catalysis involve the implicit or explicit assertion that the motions of the reacting fragments in active sites of enzymes are much more restricted than the corresponding motions in solution and that, as a result of this effect, the activation entropies for reactions in enzyme active sites are much smaller than in the corresponding barriers for the uncatalyzed reactions (e.g., refs 1−5).

Experimental estimates of the entropic contributions from the motion of the reacting fragments (which are referred to here as $(\Delta S^{\ddagger})'$) are complicated by the fact that the overall activation entropy involves contributions from the corresponding change in the entropy of the solvent. This challenges computer simulation approaches to evaluate $(\Delta S^{\ddagger})'$. Such an evaluation cannot be accomplished by standard ab initio gas phase calculations, since the entropic contributions in solution are expected to be very different than the corresponding gas phase contributions.[6] It is also not clear how to evaluate $(\Delta S^{\ddagger})'$ by current computer simulation approaches. Some advances have been made in calculations of solvation entropies[7] and in calculations of binding entropies,[8] as well as in the evaluation of ring closure entropies.[9] However, to the best of our knowledge, no advance has been made in calculating activation entropies of chemical reactions in solution. In fact, most of the current methods for simulation of chemical reactions in solution do not reflect the solute contribution to the activation entropy. This is true, for example, for approaches that constrain the substrate to move along the gas phase reaction coordinate. Such treatments do not reflect the decrease in the solute conformational freedom upon moving from the reactant state to the transition state. One of the few methods that reflect the solute entropic contribution to the calculated activation free energies is the empirical valence bond (EVB) method.[10] Yet, even with this method it is not clear how to extract the entropic component from the calculated activation free energy.

In the present work we combine the EVB method and a restraint-release approach and develop a method for calculations of the entropic contributions of the reacting fragments to the activation barriers of chemical reactions in condensed phases. Our method is outlined in section II and the results for a test

* Corresponding author.
[†] Permanent address: Institute of Physics, Charles University, Prague, 12116, Czechoslovakia.

10.1021/jp0003095 CCC: $19.00
Published on Web 04/12/2000

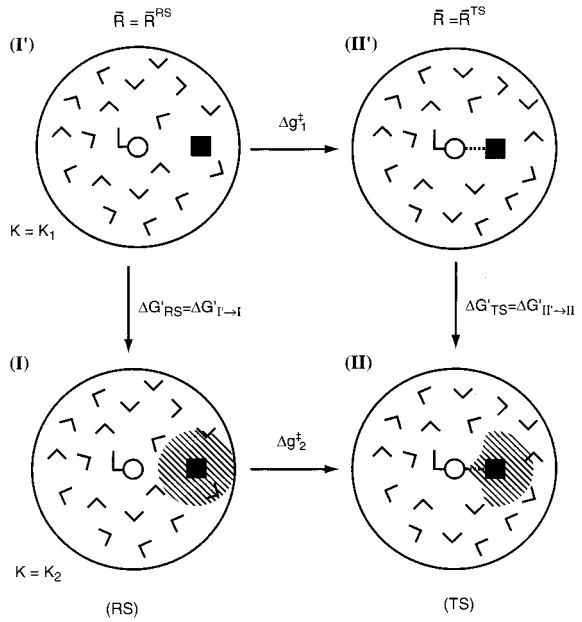


Figure 1. Thermodynamic cycle used in evaluating the entropic contributions from the motions of the reacting fragments to the activation free energies of chemical reactions in solution. The fragments are fixed in (I′) and (II′) and allowed to move (as indicated by the shaded area) in (I) and (II).

case of amide hydrolysis in solution are presented in section III. Finally, the potential of the present method is discussed in section IV.

II. Strategies for Calculations of Activation Entropies

As discussed in the Introduction, our goal is to develop a method capable of giving converging estimates of the entropic contributions of the motion of the reacting fragments to the activation energies of chemical reactions in condensed phases. In order to address this challenge we start with the thermodynamic cycle of Figure 1. This cycle considers the activation free energy, $\Delta g^{\ddagger}$, for the given reaction in two limiting conditions. In the first case (the upper part of the cycle) the reacting system is transformed from the ground to the transition state along a (unspecified) reaction coordinate while using a restraint to minimize the available configuration space in the direction perpendicular to the reaction coordinate. In this case the activation free energy, $\Delta g^{\ddagger}_1$, does not include the entropic contributions of the solute, $(\Delta S^{\ddagger}_1)' = 0$, 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, $\Delta g^{\ddagger}_2$, includes the entropic contributions of the solute. Thus, the difference between the two $\Delta g^{\ddagger}$'s gives the desired $-T(\Delta S^{\ddagger})'$, plus an enthalpic contribution that can be minimized with the proper selection of the restraint (see below). However, it is usually more convenient to use the properties of the thermodynamic cycle and obtain $\Delta g^{\ddagger}_2 - \Delta g^{\ddagger}_1$ from $\Delta G'_{RS}$ and $\Delta G'_{TS}$ (where RS and TS designate reactant state and transition state, respectively). That is, instead of defining a constrained motion from I′ to II′ we can start by imposing a strong position restraints in both states I′ and II″ and evaluate the corresponding free energies $\Delta G'_{RS} = \Delta G'_{I'\to I}$ and $\Delta G'_{TS} = \Delta G'_{II'\to II}$ associated with the release of the corresponding constraints (see below). In this way we can write

$$-T(\Delta S^{\ddagger})' = -T(\Delta S^{\ddagger}_2)' \cong \Delta g^{\ddagger}_2 - \Delta g^{\ddagger}_1 \cong \Delta G'_{TS} - \Delta G'_{RS} \quad (1)$$

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

$$U^N_{rest,j} = (K_j/2)\sum_i (R^N_i - \bar{R}^N_i)^2 \quad (2)$$

where i runs over the solute coordinates and $\bar{R}^N_i$ are reference coordinates that define the minimum of the restraint potential (see below) at the given state (N = I, or N = II for the RS and TS, respectively). The reference coordinates $\bar{R}^N_i$ are evaluated by running molecular dynamics (MD) relaxation runs on the ground and the transition states with $K = 0$. Different initial conditions can, of course, generate different $\bar{R}$'s and the selection of the optimal $\bar{R}$ as well as the implication of this selection will be discussed below.

The constraint release free energies ($\Delta G'$) are evaluated by a free energy perturbation (FEP) approach (e.g., ref 11) using

$$\Delta G'_N(K_1 \to K_2) = \sum_{m=0}^{n-1} \Delta\Delta G'_N(m \to m+1) \quad (3)$$

$$\Delta\Delta G'_N(m \to m+1) = -\beta^{-1} \ln\langle \exp\{-\beta(U^N_{m+1} - U^N_m)\}\rangle_{U^N_m} \quad (4)$$

where

$$U^N_m = (1-\lambda_m)U^N_{rest,1} + \lambda_m U^N_{rest,2} + E \quad (5)$$

and λ_m is changed from 0 to 1 in n increments. Here E designates the unconstrained potential surface of the system (see below) and $U_{rest,j}$ corresponds to the use of K_j. Generally, one is interested in the entropic contribution for a 1 M standard state. This contribution cannot be obtained by using $K_2 \to 0$ since this corresponds to infinite volume. This problem can be solved, in principle, by choosing a simulation sphere of a volume which is equal to the molar volume ($v_0 = 1660$ Å^3) while allowing K_2 to approach zero. However, such an approach is expected to encounter major convergence problems since the restraints are unlikely to sample the large simulation sphere in a reasonable simulation time. A faster convergence would be obtained by allowing the reactants to move in a smaller effective volume v_{cage}. Therefore, we must consider the free energy associated with changing v_{cage} to v_0. This free energy is simply given by[12]

$$\Delta G'' = -\beta^{-1} \ln(v_0/v_{cage}) \quad (6)$$

Thus, we have

$$-T(\Delta S^{\ddagger}_2)_0' \cong \Delta G'_{II'\to II} - \Delta G'_{I'\to I} - \beta^{-1}\ln(v_0/v_{cage}) \quad (7)$$

where $(\Delta S^{\ddagger}_2)'_0$ designates an entropic contribution in the standard state. The problem is, however, to introduce the proper volume restraint. That is, if our solute was an isolated atom then the restraint could have been easily introduced by $K_2 = K_{cage}$ where[8,12]

$$v_{cage} = \left(\frac{2\pi}{\beta K_{cage}}\right)^{3/2} \quad (8)$$

This equation is also valid for the center of gravity of a polyatomic molecule. In our case, however, the restraint is imposed on the Cartesian coordinates of each fragment atom and this involves a restriction of rotational and intramolecular entropy. Thus, we have to find an alternative way of confining the reactants to v_{cage}. Our solution of this problem involves the use of *two* constraints rather than just one; one for mapping

purpose and the other for restricting the system to the proper v_{cage}. That is, in addition to the mapping restraint we impose quadratic restraints on the distances between each reacting fragment to the central fragment. These restraints are not allowed to change during the FEP mapping process. Thus, we use a mapping potential of the form

$$U_m^N = (1-\lambda_m)U_{rest,1}^N + \lambda_m U_{rest,2}^N + \sum_{L=1}^{l-1}(K_{cage}/2)(R_c^L - \bar{R}_c^L) + E \quad (9)$$

where R_c^L is the distance between an atom of the central fragment and an atom of the Lth fragment. Using U_m^N leaves v_{cage} unaffected by the change of λ_m. Now we can let K_2 approach zero without a divergence in $\Delta S'$ since the volume of the system is restricted by the K_{cage} term. The values of R_c^L and K_{cage} are chosen in a way that the Lth fragment can move freely around the central ligand while still being in a close distance to this ligand.

The results of the thermodynamic cycle of Figure 1 depend in principle on the chosen $\bar{R}$ and may involve, in addition to $\Delta S'$, contributions from the enthalpy of the reacting fragments and from their solvation by the surrounding environment. In order to extract $\Delta S'$ from the total $\Delta G'$ we have to formulate the cycle of Figure 1 in a more rigorous way. This is done in the Appendix where we define the solute effective free energy by integrating over the solvent coordinate at each value of the solute coordinate $\bar{R}$. Using the corresponding potential of mean force (PMF) we can show that

$$-T\Delta S'_{N'\rightarrow N}(\bar{R}^N) = \Delta G'_{N'\rightarrow N}(\bar{R}^N) - (W_N(\bar{R}^N) - W_N(\bar{R}_0^N)) \quad (10)$$

where $W_N(\bar{R}^N)$ is the PMF at $\bar{R}^N$ and $\bar{R}_0^N$ gives the minimum of W_N. With this relationship it is clear that in order to obtain the exact value of $\Delta S'$ we have to find the $\bar{R}$'s that minimizes the PMF (see Appendix). Although this is impractical at present, we can exploit eq 10 and select from several simulations with different $\bar{R}^N$ the one that gives the minimal value of $|\Delta G'|$. In other words, we can try to use the lower bound

$$|-T\Delta S'_{N'\rightarrow N}(\bar{R}_0^N)| = |\Delta G'_{N'\rightarrow N}(\bar{R}_0^N)| \le |\Delta G'(\bar{R})| \quad (11)$$

and estimate $\Delta G'(\bar{R}_0^N)$ and $\Delta S'(\bar{R}_0^N)$ by finding the smallest $|\Delta G'(\bar{R})|$. With $\Delta S'_{N'\rightarrow N}(\bar{R}_0^N)$ we can obtain $(\Delta S^{\ddagger})'$ by using (see Appendix)

$$-T(\Delta S^{\ddagger})' = -T(\Delta S'_{TS}(\bar{R}_0^{II}) - \Delta S'_{RS}(\bar{R}_0^{I})) = \Delta G'_{TS}(\bar{R}_0^{II}) - \Delta G'_{RS}(\bar{R}_0^{I}) \quad (12)$$

In order to relate the calculated $(\Delta S^{\ddagger})'$ to experimental entropies it is essential to evaluate the total activation entropy, $\Delta S^{\ddagger}$, that includes both the solute and the solvent contribution. Of course, the decomposition of $\Delta S^{\ddagger}$ to solute and solvent contributions is not unique and depends on $\bar{R}$. Fortunately, we can obtain the relevant $\Delta S_{sol}^{\ddagger}$ by using the corresponding $\bar{R}_0^N$ that optimizes $|\Delta G'|$. Thus, we have (see Appendix)

$$-T\Delta S^{\ddagger} = -T(\Delta S^{\ddagger})' - T\Delta S_{sol}^{\ddagger} = \Delta G'_{TS}(\bar{R}_0^{II}) - \Delta G'_{RS}(\bar{R}_0^{I}) - T\Delta S_{sol}^{TS}(\bar{R}_0^{II}) + T\Delta S_{sol}^{RS}(\bar{R}_0^{I}) \quad (13)$$

In order to evaluate $\Delta S_{sol}^{\ddagger}$ we used here the recently developed version of the Langevin dipole (LD) model[13] implemented in the program ChemSol 2.0.[14] This method evaluates the solvation entropy by estimating the restriction of the solvent motion due to the field from the solute.

Equations 10−13 outline the formal requirements for rigorous evaluation of $(\Delta S^{\ddagger})'$ and $\Delta S^{\ddagger}$. However, since the evaluation of $\Delta S'$ with random $\bar{R}$ vectors is very expensive, we confine the present study to evaluation of the values of $\Delta S'_N$ at several $\bar{R}$'s, obtained by running MD simulations on the corresponding potential surface with zero restraint potential. We then use the $\bar{R}$'s that gives the lowest $|\Delta G'|$ in evaluating $\Delta S'$ and ΔS_{sol}.

The constraint release approach used in evaluating $\Delta G'$ is similar in some respects to the approach used by Hermans and Wang[8] in calculating binding entropies. One of the differences, however, is that Hermans' approach restraints the translations and rotations of the ligand while our approach uses a more general Cartesian restraint. This is particularly important when one deals with intramolecular motions and/or the distances between reacting fragments. Perhaps more importantly, the present study identifies the enthalpic and solvation contributions to $\Delta G'(\bar{R})$ and shows how to optimize them (see Appendix).

In order 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 since 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 estimate 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 (so that $\Delta S'$ is ignored). In our opinion, the optimal strategy is to use EVB potential surfaces[10] that can be fitted to reliable ab initio surfaces[16,17] and to relevant experimental information.[17] The refined surfaces are expected to provide consistent and reliable description of the solution reaction. This 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 allow one to perform very extensive MD simulations. The EVB method was described extensively elsewhere and is now used by other research groups.[18−26] Thus, we will mention here only several crucial points.

The EVB method describes the reaction system by mixing several diabatic states using the Hamiltonian

$$H_{ii} = \epsilon_i = \sum_k \Delta M_k^{(i)}(b_k^{(i)}) + \sum_l \gamma_l^{(i)} k_l^{(i)}(\theta_l^{(i)} - q_{0,l}^{(i)})^2 + U_{nb}^{(i)} + U_{Ss}^{(i)} + \alpha^{(i)}$$

$$H_{ij} = A_{ij}\exp\{-\mu_{ij} r_{ij}\} \quad (14)$$

where $\Delta M_k^{(i)}$ denotes the Morse potential relative to its minimum value for the kth bond in the ith valence bond (VB) structure. The second and third terms denote the bond angle bending contribution and the nonbonded electrostatic and van der Waals interactions between the reacting groups. The factor

TABLE 1: Parameters Used in the EVB Calculation[a]

potential term (system)	value (or source)
bond and bond angle	
H_2O, OH^-, H_3O^+	ref 31
$HCONH_2$, $HC(O^-)(OH)NH_2$	refs 10, 27[b]
nonbonded	
$H_2O...H_2O$	ref 31
$O^-...H$ ($V = Ae^{-\mu r}$)	$A = -130.0$ $\mu = 1.0$
$O^+...O^-$ ($V = Ae^{-\mu r}$)	$A = -130.0$ $\mu = 1.0$
$O^-(2)...C(8)$ ($V = Ar^{-12} - Br^{-6}$)	$A = 50000.0$ $B = 30.0$
off-diagonal elements	
H_{12} (taken as $C\,e^{-\mu r}$)	$C = 40.0$ $\mu = 1.0$
H_{23} (taken as $C\,e^{-\mu r}$)	$C = 30.0$ $\mu = 1.0$
gas phase shifts (kcal/mol)	$\Delta\alpha_{12} = 115.0$ $\Delta\alpha_{23} = 39.0$

Atomic Charges[d]

atom	Ψ_1	Ψ_2	Ψ_3
H(1)	0.20	0.10	0.04
O(2)	−0.40	−1.10	−0.40
H(3)	0.20	0.40	0.40
H(4)	0.20	0.40	0.40
O(5)	−0.40	−0.20	−0.20
H(6)	0.20	0.40	0.40
H(7)	0.00	0.00	−0.20
C(8)	0.90	0.50	0.60
O(9)	−0.70	−0.70	−1.00
N(10)	−1.10	−0.70	−0.70
H(11)	0.45	0.45	0.33
H(12)	0.45	0.45	0.33

[a] Energies in kcal/mol, distances in Å. [b] The bond angle parameters for the HXY angles are taken as the corresponding parameters for CXY. [c] The solute−solvent parameters are the standard ENZYMIX parameters.[32] [d] The charges were obtained by starting with ab initio PCM-HF/6-31G*//HF/6-31G* charges and then optimizing the charges in the process of fitting the EVB surface to the ab initio surface.

$\gamma_l^{(i)}$ in the second term is a coupling between bonds that are being broken or formed and the angles that depend on these bonds. The term $U_{Ss}^{(i)}$ designates the interaction between the ith state of the solute (S) and the solvent molecules (s). The parameter $\alpha^{(i)}$ determines the energy of the ith state at infinite separation between the reacting fragments. The off-diagonal elements (the H_{ij}) depend on the r_{ij} distance between the heavy atoms that define the bonds that are being broken and formed in the transition from state i to state j.

The present study involves the hydrolysis of formamide in water and the relevant EVB parameters are summarized in Table 1. The actual ground state potential surface of our system (E_g) is obtained by diagonalization of the EVB Hamiltonian:

$$\mathbf{HC}_g = \mathbf{E}_g\mathbf{C}_g \tag{15}$$

The EVB free energy surface and activation energy are determined by a FEP/umbrella sampling approach.[10,27] The FEP step in this approach uses a mapping potential of the form

$$\epsilon_m = \sum_i \lambda_m^{(i)}\epsilon_i \tag{16}$$

where the parametric change of λ_m moves the system gradually from the reactant to product states. The free energy that corresponds to the actual ground state potential surface is obtained by an umbrella sampling approach from the configurations generated in the FEP process.

In order to perform our restraint release calculations we have to confine the reactants 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 (Figure 2). In the case of a stepwise mechanism we can express the E_g of eq 15 for the RS and TS by

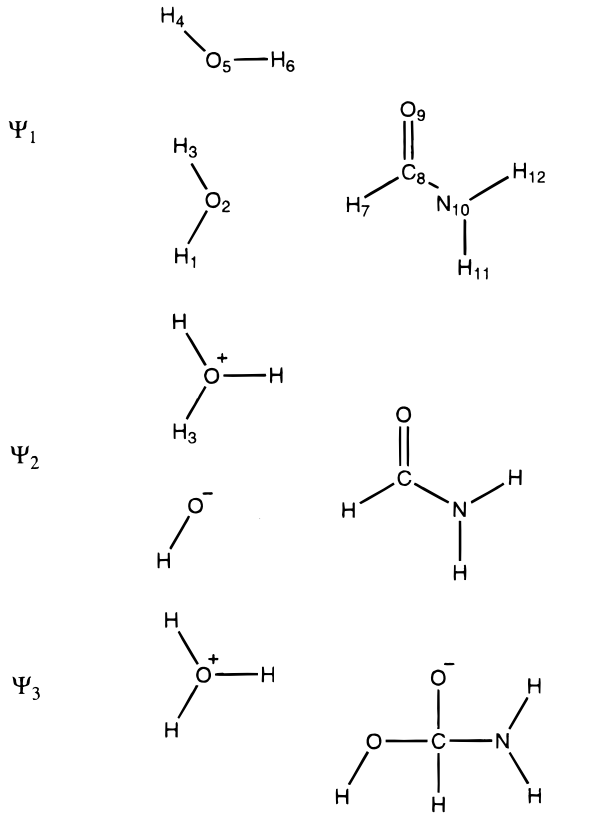


Figure 2. Resonance structures used in representing the hydrolysis of formamide in water.

$$E_g^{RS}(x) \cong \epsilon_1(x)$$

$$E_g^{TS}(x) \cong {}^1/_2[(\epsilon_2(x) + \epsilon_3(x)) - ((\epsilon_2(x) - \epsilon_3(x))^2 + 4H_{23}(x)^2)^{1/2}] \tag{17}$$

where x is the reaction coordinate which is usually taken as $\epsilon_2 - \epsilon_1$ for the first step and $\epsilon_3 - \epsilon_2$ for the second step.[10] The expression used for E_g^{RS} can be viewed as using $\lambda = (1.0, 0.0, 0.0)$ in eq 16. In the case of E_g^{TS} we should, in principle, constrain x to the transition state region ($x = x^{\ddagger}$). However, the same result can be basically obtained using

$$E \cong \lambda_2\epsilon_2 + \lambda_3\epsilon_3 \tag{18}$$

where λ_2 and λ_3 are the values used to bring the system to the TS region in the FEP procedure (usually the values of λ_2 and λ_3 are approximately 0.5 but the exact value should be determined by the proper mapping procedure). The use of eq 18 instead of $E^{TS}(x^{\ddagger})$ is reasonable since the E of eq 18 has a minimum along the reaction coordinate at $x^{\ddagger}$ and its dependence on the subspace perpendicular to x is very similar to that of E^{TS}. Similarly, in cases of concerted mechanisms we can use

$$E = \lambda_1\epsilon_1 + \lambda_2\epsilon_2 + \lambda_3\epsilon_3 \tag{19}$$

for the corresponding transition state.

III. Results and Discussion

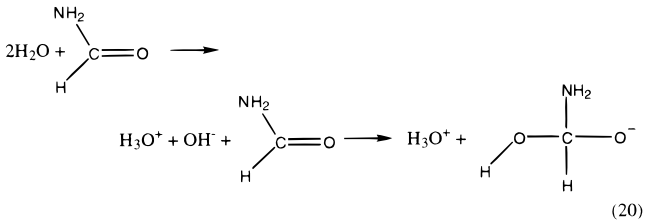
The present work considered as a test case the hydrolysis of formamide in water. The entropy of activation of this reaction

TABLE 2: Entropic Contributions to the Activation Free Energy of the Hydrolysis of Formamide in Water[a]

	$\Delta G'$						
λ used[b]	(i)	(ii)	(iii)	(iv)	(v)	(vi)	$\Delta G'$ at optimized[c] $\bar{R}$
(1.0, 0.0, 0.0) ≈ *RS*	**−42.9 (8.8)**	−43.2 (10.4)	−43.7	−43.2 (9.2)	−44.8 (9.2)	−46.3 (9.7)	−42.9 (8.8)
(0.0, 1.0, 0.0)	−41.6 (13.3)	**−39.2 (12.3)**	−40.5	−42.7 (12.3)	−43.0 (12.5)	−41.7 (12.9)	−39.2 (12.3)
(0.0, 0.5, 0.5) ≈ *TS*	−37.0 (20.2)	−34.8 (17.4)	−35.3	−38.3 (21.3)	−35.2 (13.1)	**−35.1 (12.7)**	−35.1 (12.7)
(0.5, 0.5, 0.5)		−28.9	(15.2)				
(0.3, 0.5, 0.5)					−38.5 (13.2)	−35.4 (14.6)	
(0.0, 0.0, 1.0)	−35.5 (21.8)	−35.5 (13.9)	−36.4	**−32.2 (18.7)**	−35.0 (14.1)	−36.2 (14.4)	−32.2 (18.7)
$-T(\Delta S^{\ddagger})'$							7.8
$-T(\Delta S^{\ddagger})_{sol}$							3.9
$-T(\Delta S^{\ddagger})_{calc}$							11.7
$-T(\Delta S^{\ddagger})_{obs}$[d]							8.5

[a] Energy values are given in kcal/mol. The table includes the results of six sets ((i)−(vi)) of simulations. The constraint release simulation for each λ involve a $\bar{R}$ vector generated by running trajectories with different initial conditions (i.e., different starting coordinates) over the potential surface $E = \lambda_1\epsilon_1 + \lambda_2\epsilon_2 + \lambda_3\epsilon_3$. The simulation times for the given constraint release were 100 ps except for (iii) where the initial $\bar{R}$ are the same as in (ii) but the simulation time was 400 ps. The value of $-T(\Delta S)_{sol}$ for each case is given in parentheses. The optimal values of each case are in bold. [b] The relative weights of the energies of different resonance structures. [c] As indicated in the text the results for the $\bar{R}$ that give the smallest $|\Delta G'|$ (at each λ) give the best approximation for the given $-T\Delta S'$. [d] From Radzicka and Wolfenden.[28]

at room temperature has been estimated to be ∼28 eu ($-T\Delta S^{\ddagger} \approx 8.5$ kcal/mol) using the related results for hydrolysis of peptides.[28] The rate-determining step of this reaction can be described schematically as



(20)

where the actual path might also be concerted. The potential surface for the solution reaction has been studied by a combined Langevin dipole (LD) ab initio approach[29] and found to have a flat transition state region which includes an oxyanion-like character, where the concerted and stepwise mechanisms have similar free energy. However, the calculated free energy does not include the solute entropic contribution, which is expected to increase the free energy of the concerted path. Thus, we focused on the stepwise mechanism.

In order to be able to simulate the entropic contribution we used the EVB model. The reaction was described by the three resonance forms of Figure 2. The parameters used are summarized in Table 1.

The constraint release calculations were performed for the reactant state (which was approximated using $\lambda = (1.0, 0.0, 0.0)$), the state where the proton has been transferred (simulated using $\lambda = (0.0, 1.0, 0.0)$) and the transition state for the nucleophilic attack step (simulated using $\lambda = (0.0, 0.5, 0.5)$), and also for the structures representing transition state of the concerted pathway (simulated using $\lambda = (0.5, 0.5, 0.5)$ and $\lambda = (0.3, 0.5, 0.5)$). The simulation of each state involved the change of K from 100 to 0.0003 kcal mol^{-1} Å^{-2} and K_{cage} was taken as 0.5 kcal mol^{-1} Å^{-2}. The simulations were done at 300 K with 1 fs time steps. The $\bar{R}$'s were evaluated by 20 ps relaxation run and the FEP calculations involved 100−400 ps simulation time. The results are summarized in Table 2.

The calculated $\Delta G'$ obtained in different sets of simulations is between 5 and 10 kcal/mol. However, as clarified in the text and in the Appendix, the best estimate of $-T(\Delta S^{\ddagger})'$ is obtained for each λ by taking the corresponding values from the run with $\bar{R}$ that gives the smallest $|\Delta G'|$. This approach gives $-T(\Delta S^{\ddagger})'$ of 7.8 kcal/mol. In order to compare this value to the experimental estimate of $-T\Delta S^{\ddagger}$ we have to add the corresponding solvent contribution (see eq 13). This should be done at the $\bar{R}$ values that minimize $|\Delta G'|$. The evaluation of ΔS_{sol} is done by using the recently developed version of the Langevin dipole (LD) model[13] as implemented in the program ChemSol 2.0 (see Methods section). The resulted contribution is ∼4 kcal/mol. Thus, the total calculated entropic contribution to the activation barrier is around 12 kcal/mol as compared to the experimental estimate of 8.5 kcal/mol. Note that in the case of the TS structure we used data from column (vi) although the data in column (ii) give smaller $|\Delta G'|$. However, the solvent entropy contribution $T\Delta S_{sol}$ for (ii) is highly unfavorable compared to (vi). Thus, the calculated results agree qualitatively with the corresponding observed result. Note also that the results obtained by 100 and 400 ps simulation run for the same $\bar{R}$ (sets (ii) and (iii)) give similar results. This indicates that the simulations reached a reasonable convergence.

In addition to the study of the stepwise mechanism we also examined in a preliminary way the activation entropy for more concerted pathways. This was done with $\lambda = (0.3, 0.5, 0.5)$ and $\lambda = (0.5, 0.5, 0.5)$. As can be seen from Table 2 we obtained somewhat a higher value of $-T\Delta S^{\ddagger} = 13.3$ kcal/mol for $\lambda = (0.3, 0.5, 0.5)$ compared to the $-T\Delta S^{\ddagger}$ of the stepwise reaction (see above). Since the activation free energies without the solute entropic contributions are similar for the stepwise and concerted paths[29] we may conclude that the stepwise path is more likely than the concerted path.

In addition to the constraint release approach we also examined the option of evaluating activation entropies by finite difference approach. That is, assuming that $\Delta S^{\ddagger}$ is temperature independent we tried to use the relationship

$$-\Delta S^{\ddagger} = \Delta\Delta g^{\ddagger}/\Delta T \tag{21}$$

The relevant $\Delta g^{\ddagger}$'s were evaluated by the standard FEP/EVB umbrella sampling approach.[10] This was done for different temperatures from 200 to 400 K. Different ΔT's were examined, following the prescription of Levy and co-workers.[7,30] Unfortunately, we could not obtain stable $\Delta S^{\ddagger}$ even with simulation times of 200 ps (e.g., we obtained $\Delta g^{\ddagger}_{\Psi_2\rightarrow\Psi_3}$'s of (60.6, 62.8, 57.2, 55.5, 57.5) for temperatures of (200, 250, 300, 320, 400 K). Further studies with improved sampling and longer simulation time are needed in order to examine the performance of eq 21. However, at present it appears that the constraint release approach provides significantly better convergence and stability.

IV. Concluding Remarks

This work develops and examines a method for calculations of activation entropies of chemical reactions in solution. The method developed focused on the entropic contribution of the reacting fragments. Reliable evaluation of this contribution is essential for further progress in the understanding of the role of entropic effects in enzyme catalysis. Our approach involves the thermodynamic cycle of Figure 1, where the activation barriers are considered for two paths; in one path the reacting fragments are restrained to a single reaction coordinated and in the second path they are allowed to move in the subspace perpendicular to the reaction coordinate. The difference between these two activation barriers can provide our $-T(\Delta S^{\ddagger})'$. However, instead of performing a direct calculation along the above paths we use the free energies $\Delta G'_{RS}$ and $\Delta G'_{TS}$ obtained by starting from the states where the fragments in the RS and TS, respectively, are forced to stay at a given $\bar{R}$ (by a strong Cartesian restraint) and then allowing the fragments to move by releasing the restraints. The difference between $\Delta G'_{RS}$ and $\Delta G'_{TS}$ gives us the desired $-T(\Delta S^{\ddagger})'$ and a residual contribution from the enthalpy of the system. This residual contribution was minimized by finding restraint coordinates that minimized $|\Delta G'_{RS}|$ and $|\Delta G'_{TS}|$.

In order to obtain converging results, it is essential to have ability to perform an extensive sampling of the available potential surface and this cannot be done at present with high-level ab initio approaches. Thus, we use here the EVB potential surfaces. These surfaces can provide a good approximation for the corresponding ab initio potential surfaces and describe in a consistent way the effect of the solvent on the solute Hamiltonian.

The value of $-T(\Delta S^{\ddagger})'$ found in this work is smaller than the value predicted by traditional estimates that involve the loss of three translations and three rotations in the transition state of a bimolecular reactions,[1−3] and amount to ∼15 kcal/mol. The reason for this becomes clear upon examination of the present system. It appears that many degrees of freedom have similar motions at the ground state and the transition state. This is the case, for example, for the second water molecule that became H_3O^+ at the transition state but still retained a large configurational freedom. Thus, a large part of the entropic contribution in the RS remains also in the TS. In view of the present result, it is quite likely that previous consideration of the entropic contributions to enzyme catalysis (e.g., refs 1−3) reflect an overestimate, since the entropic contributions of the reference solution reaction were probably overestimated. However, a more conclusive study of this important issue requires one to use the present approach in studies of the contribution of the substrates motion to the activation free energies of enzyme catalysis. Such studies are now in progress in our laboratory.

Acknowledgment. This work was supported by the NIH Grant GM24492. J.V. acknowledges EMBO fellowship ALTF 509-1998. We are grateful to Dr. J. Florian and Mr. C. F. Jen for insightful discussion.

Appendix

In order to obtain the solute entropic contribution to a transfer between two potential surfaces ($U_I \rightarrow U_{II}$), we have to separate the corresponding partition functions to solute and solvent contributions by writing

$$Q_N = \int dR \int dr\, e^{-U_N(R,r)\beta} = \int d\bar{R}\, [\int dR \int dr\, \delta(R - \bar{R}) e^{-U_N(R,r)\beta}] = \int d\bar{R}\, p_N(\bar{R}) = \int d\bar{R}\, e^{-W_N(\bar{R})\beta} \quad \text{(A1)}$$

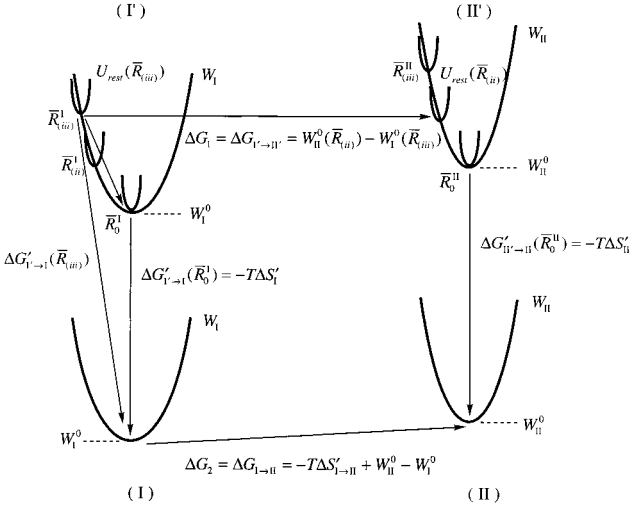


Figure 3. A schematic illustration of the effect of $\bar{R}$ and the enthalpic contribution to $\Delta G'$. The figure considers one-dimensional potentials of mean force (W) for two systems and the effect of confining the system to small segments of R by a strong constraint (K_1). The figure is essentially a rigorous equivalent of Figure 1 where the effect of restraining the system in different $\bar{R}$'s can be formulated. As explained in the Appendix and illustrated here, the value of $\Delta G'$ depends on $\bar{R}$ and the $\Delta G'$ obtained with the minimum of the corresponding W (i.e., $\bar{R}_0$) has the smallest absolute value. This $\Delta G'$ is our $-T\Delta S'$.

where U_N is the potential surface for state N and $\delta(R - \bar{R})$ indicates that the corresponding function would be collected at $\pm\Delta R/2$. Here R and r are the solute and solvent coordinate, respectively, $p_N(\bar{R})$ is the solute unnormalized probability distribution evaluated at the given $\bar{R}$ averaged over all the solvent coordinates and $W_N(\bar{R})$ is the corresponding potential of mean force (PMF) which includes, of course, the solute and solvent contributions. Expanding this potential around its minimum gives

$$W_N(\bar{R}) = U_N(\bar{R}) + \Delta g^N_{sol}(\bar{R}) = W^0_N(\bar{R}^N_0) + (K^N_W/2)(\bar{R} - \bar{R}^N_0)^2 \quad \text{(A2)}$$

where we could, of course, use R but we took $\bar{R}$ as our variable to indicate that this is our restraint coordinate. Here, $\Delta g_{sol}(\bar{R})$ is the solvation free energy at $\bar{R}$, K_W is the force constant of the quadratic term of our expansion and $\bar{R}^N_0$ is the value of $\bar{R}$ at the minimum of W_N.

Now, for the sake of simplicity, we continue by examining a case where the PMF corresponds to a one-dimensional solute coordinate, although extension to many dimensions is straightforward. Our one-dimensional case is described by the thermodynamic cycle of Figure 3. In this cycle, which is the equivalent of Figure 1, we divide R to segments that correspond to our q_N's. Each of this segments can be defined by the delta function of eq A1 or by introducing a strong quadratic constraint ($(K_1/2)(R - \bar{R})^2$) as is done in the present work. In this case, we have

$$W'_N(\bar{R}) = W^0_N(\bar{R}^N_0) + (K_W/2)(\bar{R} - \bar{R}^N_0)^2 + (K_1/2)(\bar{R} - \bar{R}^N_0)^2 = W_N(\bar{R}) + (K_1/2)(\bar{R} - \bar{R}^N_0)^2 \quad \text{(A3)}$$

Next, we use eq A2 and our quadratic constraint and evaluate the relevant partition functions.

$$Q_N = e^{-W_N{}^0(\bar{R})\beta} \int_{-\infty}^{\infty} d\bar{R}\, e^{-(K_W{}^N/2)(\bar{R}-\bar{R}_0{}^N)^2\beta} = e^{-W_N{}^0(\bar{R})\beta} \sqrt{2\pi/\beta K^N_W}$$

$$p'_N(\bar{R}) = e^{-W_{N'}(\bar{R})\beta} = e^{-W_N(\bar{R})\beta}\int_{-\infty}^{\infty} d\bar{R}' e^{-(K_1/2)(\bar{R}'-\bar{R}_0^N)^2\beta} = e^{-W_N(\bar{R})\beta}\sqrt{2\pi/\beta K_1}$$

$$p_N(\bar{R}) = e^{-W_N(\bar{R})\beta} = p'_N/\sqrt{2\pi/\beta K_1} \quad \text{(A4)}$$

where we use p' for the partition function evaluated with W_N and where $W_N^0 = W_N(\bar{R}_0^N)$. This gives

$$\Delta G_{I\to II} = -\beta^{-1}\ln(Q_{II}/Q_I) = \Delta W^0_{I\to II} - {}^1/_2\beta^{-1}\ln(K_W^I/K_W^{II})$$

$$\Delta G'_{I'\to I}(\bar{R}) = G_I - g'_I(\bar{R}) = -\beta^{-1}\ln(Q_I/p'_I(\bar{R})) = -\frac{1}{2}\beta^{-1}\ln(K_1/K_W^I) + (W_I^0 - W_I(\bar{R}))$$

$$\bar{E}_I = k_B T^2(\partial \ln Q_I/\partial T)_V = W_I^0 + (k_B/2)T - (T\partial W_I^0/\partial T) = W_I^0 + (k_B/2)T + T\Delta S^I_{sol}(\bar{R}_0{}^I) \quad \text{(A5)}$$

where $\bar{E}$ is the average energy, $\Delta W^0_{I\to II} = W^0_{II} - W^0_I$ and g' is defined by eq A5. We also have

$$S_I = k_B \ln Q_I + \bar{E}_I/T = (k_B/2)[1 + \ln(2\pi/\beta K_W^I)] + \Delta S^I_{sol}(\bar{R}_0^I)$$

$$S_{I'}(\bar{R}) = (k_B/2)[1 + \ln(2\pi/\beta K_1)] + \Delta S^I_{sol}(\bar{R}) \quad \text{(A6)}$$

Now we can write

$$\Delta S_{I\to II} = (k_B/2)\ln(K_W^I/K_W^{II}) + \Delta S^{II}_{sol}(\bar{R}_0^{II}) - \Delta S^I_{sol}(\bar{R}_0^I) = (k_B/2)\ln(K_W^I/K_W^{II}) + \Delta S^{I\to II}_{sol} = \Delta S'_{I\to II} + \Delta S^{I\to II}_{sol} \quad \text{(A7)}$$

$$\Delta S_{I'\to I}(\bar{R}) = (k_B/2)\ln(K_1/K_W^I) + \Delta S^I_{sol}(\bar{R}_0^I) - \Delta S^I_{sol}(\bar{R}) \quad \text{(A8)}$$

Thus

$$\Delta S_{I'\to I}(\bar{R}_0^I) = (k_B/2)\ln(K_1/K_W^I) = \Delta S'_{I'\to I}(\bar{R}_0^I) = -\Delta G'_{I'\to I}(\bar{R}_0^I)/T \quad \text{(A9)}$$

Finally we have from eq A7 and eq A5

$$\Delta G'_{II'\to II}(\bar{R}_0^{II}) - \Delta G'_{I'\to I}(\bar{R}_0^I) = -(k_B T/2)\ln(K_W^I/K_W^{II}) = -T\Delta S'_{I\to II} \quad \text{(A10)}$$

The usage of this relationship requires, however, knowing the values of $\bar{R}_0^I$ and $\bar{R}_0^{II}$. This can be done, in principle, by minimizing the corresponding PMF but at present such calculations are not practical for a case when R involves more than a few coordinates. Fortunately, using eq A5 we can write

$$|\Delta G'(\bar{R}_0)| \le |\Delta G'(\bar{R})| \quad \text{(A11)}$$

Thus, we estimate $\Delta G'(R_0)$ by trying different $\bar{R}$'s and taking the $\Delta G'(\bar{R})$ with the smallest absolute value. Now we can use eqs A9 and A10 to evaluate $\Delta S'_{I\to II}$ and $\Delta S'_{I'\to I}(\bar{R}_0{}^I)$. The $\bar{R}$ that minimizes $|\Delta G'|$ is also used in evaluating the $\Delta S^{I\to II}_{sol}$ of eq A9 and the total entropy $\Delta S_{I\to II}$.

References and Notes

(1) Jencks, W. P. *Catalysis in Chemistry and Enzymology*; Dover: New York, 1987.
(2) Page, M. I.; Jencks, W. P. *Proc. Natl. Acad. Sci. U.S.A.* **1971**, *68*, 1678.
(3) Page, M. I. *Angew. Chem., Int. Ed. Engl.* **1977**, *16*, 449.
(4) McMurry, J.; Castellion, M. E. In *Fundamentals of Organic and Biological Chemistry*; Prentice Hall: New York, 1999; p 243.
(5) Jencks, W. P. *Adv. Enzymol.* **1975**, *43*, 219.
(6) Florián, J.; Warshel, A. *J. Phys. Chem. B* **1998**, *102*, 719.
(7) Levy, R. M.; Gallicchio, E. *Annu. Rev. Phys. Chem.* **1998**, *49*, 531.
(8) Hermans, J.; Wang, L. *J. Am. Chem. Soc.* **1997**, *119*, 2707.
(9) Lightstone, F. C.; Bruice, T. C. *J. Am. Chem. Soc.* **1996**, *118*, 2595.
(10) Warshel, A. *Computer Modeling of Chemical Reactions in Enzymes and Solutions*; John Wiley & Sons: New York, 1991.
(11) Warshel, A. In *Encyclopedia of Molecular Biology*; Creighton, T. E., Ed.; John Wiley & Sons: 1999; p 939.
(12) Sham, Y. Y.; Chu, Z. T.; Tao, H.; Warshel, A. *PROTEINS*, in press,.
(13) Florián, J.; Warshel, A. *J. Phys. Chem. B* **1999**, *103*, 10282.
(14) Florián, J.; Warshel, A. *ChemSol, Version 2.0*; University of Southern California: Los Angeles, 1999. The program can be downloaded from the anonymous ftp server ftp.usc.edu, directory/pub/warshel/cs2, or from the web page http://laetro.usc.edu.
(15) Chandrasekhar, J.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1985**, *107*, 2974.
(16) Muller, R. P.; Warshel, A. *J. Phys. Chem.* **1995**, *99*, 17516.
(17) Villà, J.; Bentzien, J.; González-Lafont, A.; Lluch, J. M.; Bertran, J.; Warshel, A. *J. Comput. Chem.*, in press.
(18) Kim, H. J.; Hynes, J. T. *Int. J. Quantum Chem. Quantum Chem. Symp.* **1990**, *24*, 821.
(19) Kim, H. J.; Hynes, J. T. *J. Am. Chem. Soc.* **1992**, *114*, 10528.
(20) Bala, P.; Grochowski, P.; Lesyng, B.; McCammon, J. A. In *Quantum Mechanical Simulation Methods for Studying Biological Systems*; Bicout, D., Field, M., Eds.; Springer: Berlin, 1995; p 119.
(21) Grochowski, P.; Lesyng, B.; Bala, P.; McCammon, J. A. *Int. J. Quantum Chem.* **1996**, *60*, 1143.
(22) Hinsen, K.; Roux, B. *J. Comput. Chem.* **1997**, *18*, 368−380.
(23) Lobaugh, J.; Voth, G. A. *J. Chem. Phys.* **1994**, *100*, 3039.
(24) Vuilleumier, R.; Borgis, D. *Chem. Phys. Lett.* **1998**, *284*, 71.
(25) Kim, Y.; Corchado, J. C.; Villà, J.; Xing, J.; Truhlar, D. G. *J. Chem. Phys.* **2000**, *112*, 2718.
(26) Schmitt, U. W.; Voth, G. A. *J. Phys. Chem.* **1998**, *102*, 5547.
(27) Warshel, A.; Sussman, F.; Hwang, J.-K. *J. Mol. Biol.* **1988**, *201*, 139.
(28) Radzicka, A.; Wolfenden, R. *J. Am. Chem. Soc.* **1996**, *118*, 6105.
(29) Strajbl, M.; Florián, J.; Warshel, A. *J. Am. Chem. Soc.*, in press.
(30) Kubo, M. M.; Gallicchio, E.; Levy, R. M. *J. Phys. Chem. B* **1997**, *101*, 10527.
(31) Åqvist, J. *J. Phys. Chem.* **1991**, *95*, 4587.
(32) Lee, F. S.; Chu, Z. T.; Warshel, A. *J. Comput. Chem.* **1993**, *14*, 161.