Energetics of the Catalytic Reaction of Ribonuclease A: A Computational Study of Alternative Mechanisms

Timothy M. Glennon and Arieh Warshel*

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

Received May 7, 1998. Revised Manuscript Received July 30, 1998

Abstract: The energetics of the catalytic reaction of ribonuclease A have been explored with empirical valence bond (EVB) simulations to elucidate the origin of the enormous catalytic power of this enzyme and to examine different mechanistic alternatives. The two mechanisms analyzed were a general acid−general base mechanism with a dianionic transition state and a triester-like mechanism with a monoionic transition state. The first step of the analysis used experimental information to determine the activation energy of each assumed mechanism in a water cage. This has provided an experimentally based reference point for the catalytic effect of the enzyme. The next step of the analysis involved EVB simulations of the reaction in water and calibrated these simulations against the above-mentioned energetics of the reference reaction. The simulations were performed in the protein environment without changing any EVB parameters. The catalysis was measured as the difference in the overall activation free energy of the respective mechanism in water and in protein. In the mechanism with the dianionic transition state a catalytic effect of $\sim 18 \pm 6$ kcal/mol was established which is in good agreement with the experimentally derived estimate of ~ 21 kcal/mol. In the mechanism with the monoionic transition state, a value of $\sim 12 \pm 6$ kcal/mol was calculated which is 5 kcal/mol lower than its estimated value, making it slightly less likely to be the actual mechanism in the enzyme. The origin of the catalytic power is attributed to an electrostatic reduction of the activation barriers. This reduction is associated with the preorganized polar environment of the enzyme.

1. Introduction

The ribonucleases (RNases) are a class of enzymes that catalyze the cleavage of phosphodiester bonds in RNA.[1−5] These enzymes are very effective. RNase A, in particular, provides an estimated $\sim 10^{15}$ increase in reaction rate (see below). The structure of the enzyme[6] and the main residues which are thought to provide the catalytic effect are shown in Figure 1. It is unclear, however, as to what is the actual mechanism and how it is being accelerated.

The general mechanism of RNase A, as shown in Figure 2, consists of two steps: a transphosphorylation that cleaves the RNA strand and forms a cyclic phosphate and then a hydrolysis of the cyclic phosphate by a water molecule. Most of the debate about the mechanism is centered on the transphosphorylation reaction. The key issues, which we discuss below, include the changes in the charge distribution of the phosphate during the nucleophilic attack, the roles of the individual residues shown in Figure 1, and the route of the proton shuttle.

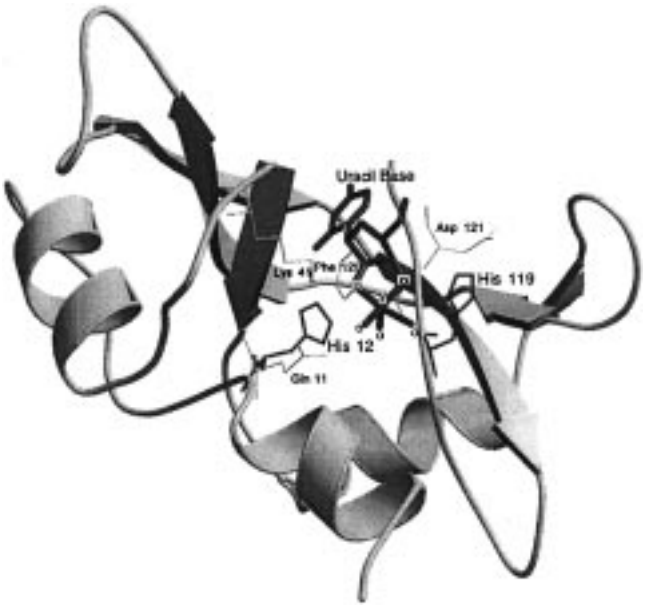


Figure 1. The structure of RNase A. The residues thought to be significant to catalysis are depicted in stick format.

To illustrate the various issues we have outlined the transphosphorylation steps of some examples of suggested mechanisms in Figures 3 and 4. Figure 3 provides an example of the type of mechanism referred to as the general acid−general base mechanism. In the first step a proton is transferred from O2′ to the general base, His 12. The next step is the nucleophilic attack of O2′ on the phosphorus which is followed by a proton transfer from the general acid, His 119, to O2P. In the last

* To whom correspondence should be addressed.

(1) Wlodawer, A. In *Biological Macromolecules and Assemblies*; Jurnak, F. A., McPherson, A., Eds.; Wiley: New York, 1985; Vol. II; p 393.
(2) Blackburn, P.; Moore, S. In *The Enzymes*; Boyer, P. D., Ed.; Academic Press: New York, 1982; Vol. 15; p 317.
(3) Richards, F. M.; Wyckoff, H. W. In *The Enzymes*; Boyer, P. D., Ed.; Academic Press: New York, 1971; Vol. 4; p 647.
(4) Perreault, D. M.; Anslyn, E. V. Unifying the Current Data on the Mechanism of Cleavage−Transesterification of RNA. *Angew. Chem., Int. Ed. Engl.* **1997**, *36*, 432.
(5) Raines, R. T. Ribonuclease A. *Chem. Rev.* **1998**, *98*, 1045.
(6) Borah, B.; Chen, C.-W.; Egan, W.; Miller, M.; Wlodawer, A.; Cohen, J. S. Nuclear Magnetic Resonance and Neutron Diffraction Studies of the Complex of Ribonuclease A with Uridine Vanadate, A Transition-State Analogue. *Biochemistry* **1985**, *24*, 2058.

S0002-7863(98)01594-7 CCC: $15.00
Published on Web 09/16/1998

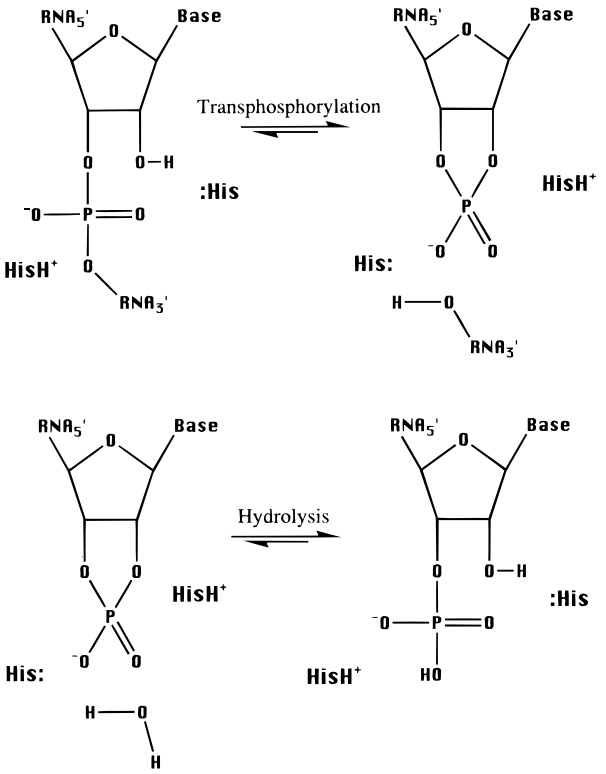


Figure 2. Two main steps in the reaction of RNase.

step, the proton is transferred to O−RNA3′, which is then cleaved from the phosphate. From a classification point of view, the most significant attribute of this class of mechanisms is not its general acid−general base nature but the *dianionic* nature of the phosphodiester intermediate (or transition state) of the nucleophilic attack step. We therefore refer to this type of mechanism throughout as the dianionic intermediate (DI) mechanism. However, this mechanism is frequently considered as a concerted mechanism where the dianionic intermediate is not described explicitly. This is the case with the mechanism of Gerlt and Gassman,[7] who also proposed that Lys 41 and Phe 120 or Gln 11 stabilize the transition state with low barrier hydrogen bonds (LBHB) where the proton forms a partially covalent hydrogen bond between the proton donor and a nonbridging phosphoryl oxygen (note, however, that the concept of LBHB is problematic[8]). Another version of the concerted model was considered by Herschlag.[9]

The other class of mechanisms involves a monoionic intermediate following the nucleophilic attack and is referred to here as the MI mechanisms. One version of the MI mechanism, which we refer to as the internal proton transfer mechanism, is shown in Figure 4. In the first step, a proton is transferred internally from O2′ to O2P. The second step is the nucleophilic attack and the last step is the cleavage of the leaving group. The other mechanism shown in Figure 4 is the so-called modified Breslow mechanism.[10] In this mechanism the nucleophilic attack is concerted with two proton transfers, one from O2′ to His 12 and another from His 119 to O2P. In the next two steps a proton is transferred from His 12 to a nonbridging oxygen and another is transferred from O2P to His 119. The last step is the cleavage of the leaving group. This class is also referred to as the triester mechanism because the nonbridging oxygen of the phosphate is protonated prior to the nucleophilic attack in the nonconcerted pathways of the mechanism and during it in some concerted ones. It is viewed by some that this protonated oxygen corresponds to a methylated oxygen which would make the phosphate a triester. However, this similarity cannot be invoked in quantitative analysis, as was demonstrated in a recent computational analysis of hydroxyl anion attack on trimethyl phosphate and dimethyl phosphate in the gas phase.[11]

The contributions of individual residues to the catalytic effect of RNase can provide valuable mechanistic information. The mechanisms in Figures 3 and 4 are consistent with the catalytic effect that is often attributed to His 12 and His 119 but do not provide unique information about the role of these crucial residues. The observed roles of other residues, particularly Lys 41, are even harder to rationalize.[7,9,12−14] This residue may be providing charge stabilization to the transition state or alternatively providing hydrogen bond stabilization.[7] It has also been proposed,[12] in a theoretical study based on minimal models, that Lys 41 acts as a general base and deprotonates the O2′ hydroxyl, but as will be argued below, this proposal is problematic. Other residues that have also been deemed important to catalysis include Gln 11 and Phe 120, which have been suggested as providing LBHB's.[7] These residues can clearly help by regular hydrogen bond stabilization rather than LBHB. Asp 121 has also been implicated as an important residue.[15]

The above-mentioned mechanisms are based on crystallographic analyses,[9] mutagenesis data,[5,9,13,16] and kinetic evidence,[17,18] as well as chemical experience. However, these approaches have been unable to determine in an unequivocal way the actual mechanism of RNase. Observations of the relative position of His 12 in various crystal structures are often in conflict with each other or lead to ambiguous conclusions.[4,9] Even if such observations were in agreement with each other, predicting a mechanism based solely on structure is not expected to be conclusive. To reliably deduce an enzymatic mechanism, the analyses should be based on energetics and use crystal structures as a starting point for obtaining the relevant structure−energy−function correlation. One means of accomplishing this has been through mutagenesis studies which have attempted to

(7) Gerlt, J. A.; Gassman, P. G. Understanding the Rates of Certain Enzyme-Catalyzed Reactions: Proton Abstraction from Carbon Acids, Acyl-Transfer Reactions, and Displacement Reactions of Phosphodiesters. *Biochemistry* **1993**, *32*, 11943.

(8) Warshel, A.; Papazyan, A. Energy Considerations Show that Low-Barrier Hydrogen Bonds Do Not Offer A Catalytic Advantage Over Ordinary Hydrogen Bonds. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 13665.

(9) Herschlag, D. Ribonuclease Revisited: Catalysis via the Classical General Acid−General Base Mechanism or a Triester-like Mechanism. *J. Am. Chem. Soc.* **1994**, *116*, 11631.

(10) Breslow, R.; Xu, R. Recognition and Catalysis in Nucleic Acid Chemistry. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 1201.

(11) Florián, J.; Warshel, A. A Fundamental Assumption About OH^- Attack in Phosphate Hydrolysis Is Not Fully Justified. *J. Am. Chem. Soc.* **1997**, *119*, 5473.

(12) Wladkowski, B. D.; Krauss, M.; Stevens, W. J. Transphosphorylation Catalyzed by Ribonuclease A: Computational Study Using ab Initio Effective Fragment Potentials. *J. Am. Chem. Soc.* **1995**, *117*, 10537.

(13) Messmore, J. M.; Fuchs, D. N.; Raines, R. T. Ribonuclease A: Revealing Structure: Function Relationships with Semisynthesis. *J. Am. Chem. Soc.* **1995**, *117*, 8057.

(14) Haydock, K.; Lim, C.; Brunger, A. T.; Karplus, M. Simulation Analysis of Structures on the Reaction Pathway of RNase A. *J. Am. Chem. Soc.* **1990**, *112*, 3826.

(15) Umeyama, H.; Nakayawa, S.; Fujii, T. *Chem. Pharm. Bull.* **1979**, *27*, 974.

(16) Thompson, J. E.; Raines, R. T. Value of General Acid−Base Catalysis to Ribonuclease A. *J. Am. Chem. Soc.* **1994**, *116*, 5467.

(17) Anslyn, E.; Breslow, R. On the Mechanism of Catalysis by Ribonuclease: Cleavage and Isomerization of the Dinucleotide UpU Catalyzed by Imidazole Buffers. *J. Am. Chem. Soc.* **1989**, *111*, 4473.

(18) Thompson, J. E.; Kutateladze, T. G.; Schuster, M. C.; Venegas, F. D.; Messmore, J. M.; Raines, R. T. Limits to Catalysis by Ribonuclease A. *Bioorg. Chem.* **1995**, 23, 471.

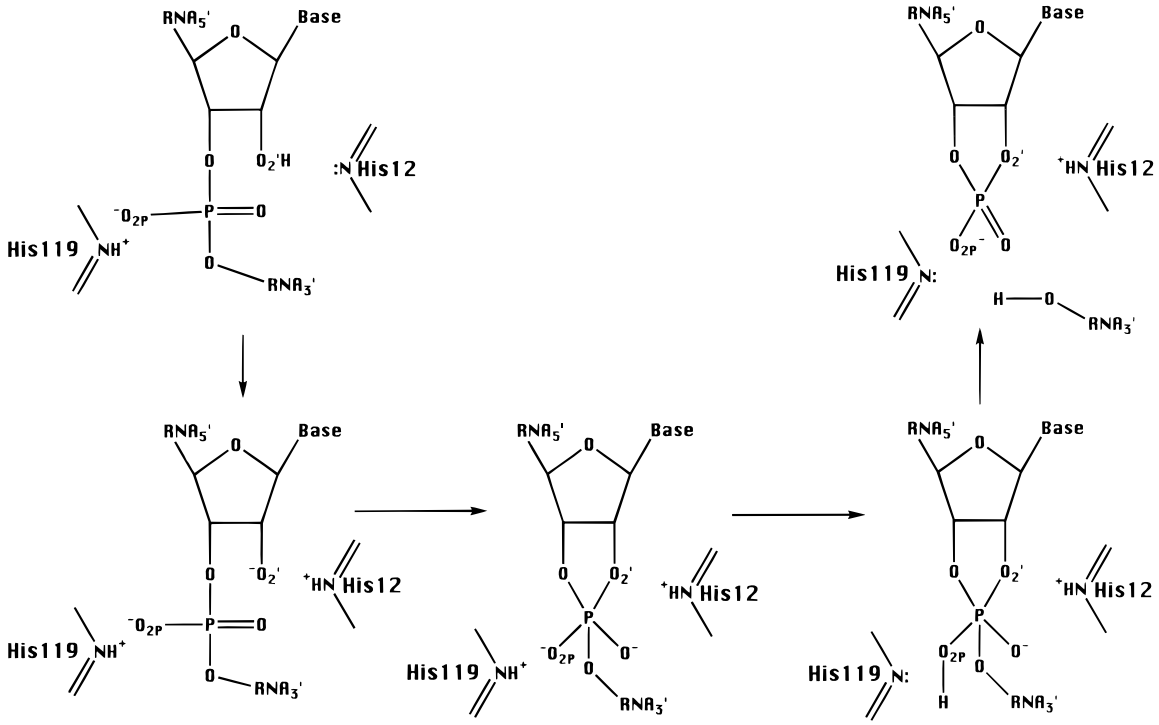


Figure 3. Schematic description of the general acid−general base mechanism for RNase A transphosphorylation. This mechanism is referred to in the paper as the dianionic intermediate (DI) mechanism.

pinpoint the involvement and significance of various residues in the enzymatic mechanism. For example, Raines and coworkers have studied the effect of mutations of Lys 41 and interpreted their findings as evidence that this residue donates an ordinary hydrogen bond.[13] They have also performed similar studies on His 12 and His 119 and established their contribution to the catalytic power of the enzyme.[5,16]

Computational investigations of RNase energetics have mostly focused on phosphate chemistry in the gas phase.[19,20] Some have attempted to model the active site by using models of a few catalytic residues in an otherwise gas-phase model,[21] and others have tried to examine related reactions in the presence of a few water molecules.[22] In a study which made progress toward a more complete model, the immediate surrounding environment of 10 amino acid residues and six solvent molecules was modeled by effective fragment potentials.[12] This study was interpreted that RNase acts through a protonated monoanionic pentacoordinated mechanism with lysine acting as a base.[12] However, even this theoretical attempt considered only a few explicit residues in the active site and could not account for the enormous effect of the surrounding solvent. For example, the Lys 41 is deprotonated in the model of ref 12, while with the proper representation of its environment, it would be protonated. Molecular dynamics (MD) simulations of the ground state of the enzyme−substrate system were also reported[23] in an attempt to examine mechanistic issues. However, although such simulations can help in refining the ground-state structure, they cannot tell us about the activation free energies of competing mechanisms and thus cannot be used in a conclusive way in mechanistic studies (e.g., see ref 24).

Another fundamental problem that has not been addressed in a quantitative way is the origin of the catalytic power of RNase. As will be shown below, this enzyme catalyzes its reaction by more than 20 kcal/mol relative to the corresponding reaction in water. Although mutation experiments help to elucidate the role of key residues as discussed above, they cannot provide unique information about the *totality* of the catalytic effect. More specifically, the effect of mutation experiments is frequently interpreted with the implicit assumption that the sum of the effect of all the catalytic residues account for the total catalysis (e.g., ref 25). Yet this may not be the case, not only because mutation effects are not always additive but because catalysis reflects the effect of all the residues of the native protein relative to a state where the entire protein is replaced by water. This point can be realized by considering a hypothetical *nonpolar* protein with a single polar residue that forms a hydrogen bond with the transition state of the substrate. In a pure nonpolar environment, ∼15 kcal/mol is provided from this single residue. That is, a hydrogen bond to a negatively charged group gives around 30 kcal/mol[8] and about half of this value is expected in a nonpolar environment with a dielectric constant of ∼2. Some may interpret this experimental result as evidence that the overall catalytic effect of the protein is *more* than 15 kcal/mol. However, our hypothetical protein will be an *anti-catalyst* since a single hydrogen bond in a nonpolar environment would stabilize a charge transition state much *less* than water does. In other words, nonpolar environments

(19) Lim, C.; Tole, P. Endocyclic and Exocyclic Cleavage of Phosphorane Monoanion: A Detailed Mechanism of the RNase A Transphosphorylation Step. *J. Am. Chem. Soc.* **1992**, *114*, 7245.

(20) Dejaegere, A.; Liang, X.; Karplus, M. Phosphate Ester Hydrolysis: Calculation of Gas-Phase Reaction Paths and Solvation Effects. *J. Chem. Soc., Faraday Trans.* **1994**, 90, 1763.

(21) Wladkowski, B. D.; Krauss, M.; Stevens, W. J. Ribonuclease A Catalyzed Transphosphorylation: An *ab Initio* Theoretical Study. *J. Phys. Chem.* **1995**, *99*, 6273.

(22) Yliniemela, A.; Uchimaru, T.; Tanabe, K.; Taira, K. Do Pentacoordinate Oxyphosphorane Intermediates Always Exist. *J. Am. Chem. Soc.* **1993**, *115*, 3032.

(23) Haydock, K.; Allen, L. C. In *Progress in Clinical and Biological Research*; Rein, R., Liss, A. R., Eds.; 1985; Vol. 172A; p 87.

(24) Náray-Szabó, G.; Fuxreiter, M.; Warshel, A. In *Computational Approaches to Biochemical Reactivity*; Náray-Szabó, G., Warshel, A., Ed.; Kluwer Academic Publishers: Dordrecht, The Netherlands, 1997; p 237.

(25) Mildvan, A. S.; Weber, D. J.; Kuliopulos, A. Quantitative Interpretations of Double Mutations of Enzymes. *Arch. Biochem. Biophys.* **1992**, *294*, 327.

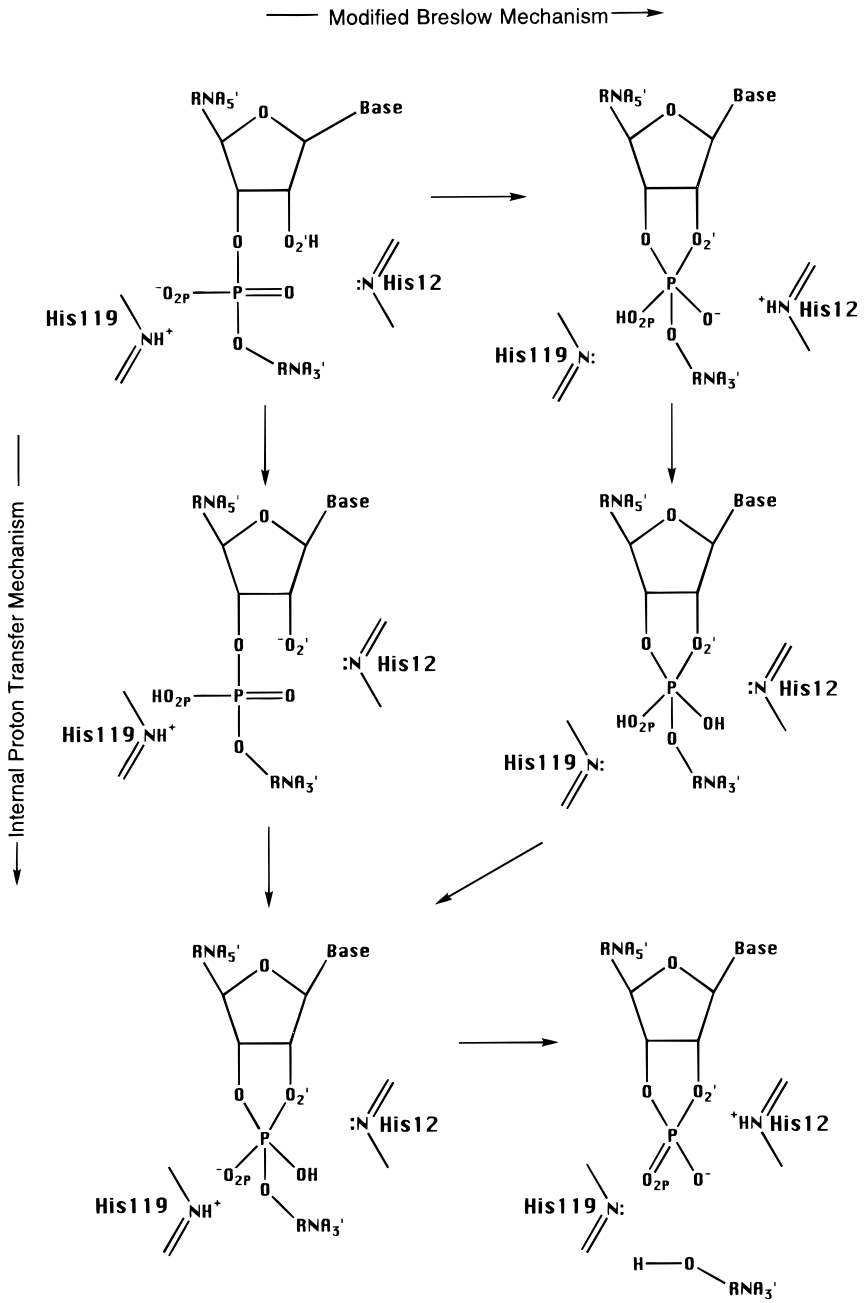


Figure 4. Schematic description of the triester-like mechanism for RNase A transphosphorylation. This class of mechanisms is referred to in the paper as the monoanionic intermediate (MI) mechanism.

destabilize charges enormously and a single hydrogen bond cannot compensate for this (see Figure 3 in ref 8).

To elucidate the relationship between the structure of RNase to its overall catalytic effect, it is important to have a computational model that provides the energetics of the assumed mechanism in the protein active site and in solution. Here, we will use such a model in analyzing the energetics of RNase in the DI mechanism of Figure 3 and the internal proton-transfer MI mechanism of Figure 4. This study will be based on a series of empirical valance bond (EVB)[26] simulations. We should perhaps emphasize that the EVB calculations in protein are calibrated against their respective reference reactions in water. This approach allows one to calculate not only the activation free energy of the reaction in the protein ($\Delta g^{\ddagger}_{pro}$) but also the difference between $\Delta g^{\ddagger}_{pro}$ and $\Delta g^{\ddagger}_{wat}$. It is this difference after all that determines the catalytic power of an enzyme. The values of $\Delta g^{\ddagger}_{wat}$ are evaluated using experimental values of pK_a's, reaction free energies, and activation energies of relevant processes in solution, and in some cases ab initio results. Using the experimentally based estimates of $\Delta g^{\ddagger}_{wat}$ and EVB simulations we calculate the catalytic effects of the enzyme in the above-mentioned mechanisms.

2. Analysis of Relevant Thermodynamics

To quantify the origin of the catalytic power of a given enzyme, it is crucial to have a clear idea about the overall effect of the protein and solvent system. Comparing the rate constant of the given enzymatic reaction to that of the uncatalyzed reaction in water is of course very useful but it might not be sufficient for this purpose. Such a comparison may involve different reactions (the fastest mechanism in solution might be different than the actual mechanism in the enzyme active site)

(26) Warshel, A. *Computer Modeling of Chemical Reactions in Enzymes and Solutions*; John Wiley & Sons: New York, 1991.

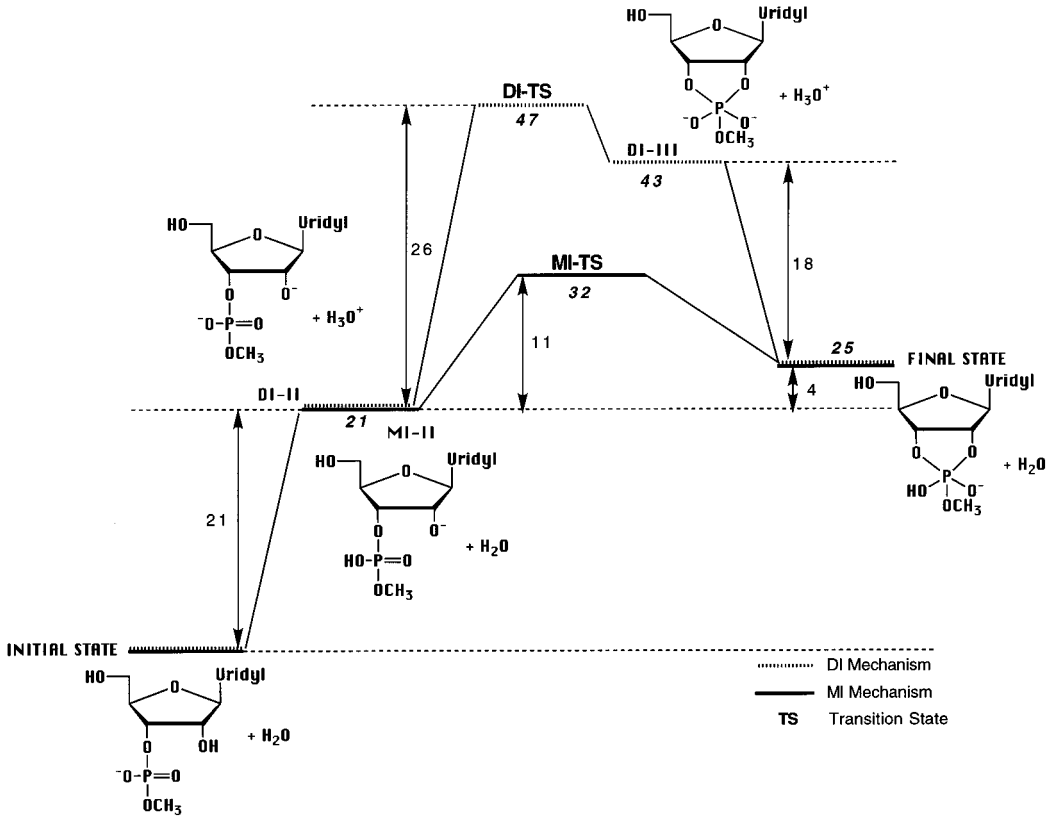


Figure 5. Free energy profiles deduced from experimental and experimentally derived data for the DI and MI mechanisms in water. Energies are in kcal/mol.

and therefore will not be related directly to the effect of the active site environment. Thus, it is crucial to define a reference state that is directly relevant to the actual catalytic effect of the enzyme. As much as computational analysis is concerned, it is our view that the most effective reference state is provided by placing the substrate in a solvent cage and forcing it to undergo the *same* chemical reaction as the one assumed in the given enzymatic reaction.[26] This reference state allows one to determine the actual environmental effect of the enzyme by comparing the activation energy in the enzyme site ($\Delta g^{\ddagger}_{cat}$) to the corresponding barrier in a solvent cage ($\Delta g^{\ddagger}_{cage}$). In this way we eliminate somewhat trivial and unnecessarily confusing concentration effects. Having the fragments constrained to the same solvent cage allows one to calculate by computer simulation the relevant entropic effects and to examine their catalytic contributions. We therefore can determine the overall effect of the enzyme active site on any assumed mechanism rather than trying to reproduce the difference between the uncatalyzed and catalyzed reaction. Thus, for example, the replacement of water by histidine as a general base is considered by us as a change in the "chemistry" which is fully correlated with corresponding change in pK_a's and not as an open puzzle about the origin of enzyme catalysis. Therefore this effect is not included in the difference between $\Delta g^{\ddagger}_{cat}$ and $\Delta g^{\ddagger}_{cage}$ (both the enzyme and the reference reaction include the His residue). Our reference state is not yet widely used perhaps because of the fact that not many quantitative calculations of enzyme catalysis are performed or because the corresponding energetics cannot be determined by direct experiments. However, once we accept the use of thermodynamic cycles, it is almost always possible to obtain a quantitative estimate of the relevant energetics using experimental information about elementary steps of the given reaction in water. This point will be illustrated below by analyzing two possible mechanisms of RNase.

We start our analyses by considering the energetics of the DI and MI nonconcerted mechanisms in water. The corresponding free energy profiles are summarized schematically in Figure 5. In the first step of the MI mechanism, a proton is transferred internally to one of the nonbridging phosphate oxygens. The route that this proton takes is irrelevant. Whether it is transferred externally through a water molecule or directly from O2′H will not matter as long as it does not involve the rate-limiting step. The calculated standard free energy of this proton transfer (PT) step is determined using the equation[26]

$$(\Delta G_0)_{PT} = 1.38(\mathrm{p}K_a(\mathrm{O2'H}) - \mathrm{p}K_a(\mathrm{O2P})) + \Delta G_{QQ} \quad (1)$$

$$= (\Delta G_0)^{\infty}_{PT} + \Delta G_{QQ}$$

where ΔG_{QQ} is the change in the electrostatic interaction between the donor and acceptor during the proton-transfer process. $(\Delta G_0)^{\infty}_{PT}$ is the proton-transfer energy when the donor and acceptor are at infinite distance. With a pK_a of ~16 for O2′ and a pK_a of 0.76 for O2P, as determined by Guthrie,[27] we obtain $(\Delta G_0)_{PT} \cong 21$ kcal/mol. This ΔG_0 corresponds to a standard state of 1 M for the donor and acceptor. Now, in principle, we should invest free energy in bringing the donor and acceptor (O2′ and O2P) to the same solvent cage, but in our specific case, they are confined to such a cage since they are attached to the same molecule. Thus $\Delta G_{PT} \cong \Delta G_0$.

In the DI mechanism in water, the proton is transferred to a water molecule, keeping the phosphate a diester. Using an equation similar to the one above and a pK_a value of −1.7 for

(27) Guthrie, J. P. Hydration and Dehydration of Phosphoric Acid Derivatives: Free Energies of Formation of the Pentacoordinate Intermediates of Phosphate Ester Hydrolysis and of Monomeric Metaphosphate. *J. Am. Chem. Soc.* **1977**, *99*, 3991.

H_3O^+,[28] the proton-transfer standard free energy is estimated to be 24 kcal/mol. Now we have to consider the fact that the proton acceptor is a water molecule (whose concentration is 55 M). This will increase ΔG by RT ln 55. However, we also have to consider the system where H_3O^+ is in the same solvent cage as the O2′$^-$. This requires 55 M H_3O^+, bringing ΔG back to 24 kcal/mol. We also have to consider the electrostatic attraction, ΔG_{QQ}, between O2′$^-$ and H_3O^+. This interaction can be estimated to a reasonable approximation using a dielectric constant of ∼30[29] ($\Delta G_{QQ} \cong -2.5$ kcal/mol). Note that the dipole−charge interactions (e.g., the interaction between H_3O^+ with the residual charges of the ring atoms) have a very small effect in solution, as is evidenced by the fact that different polar substituents do not change significantly the pK_a's of ionizable molecules.[29] Thus they are neglected in the calculations of $\Delta\Delta G_{QQ}$. The above considerations give $\Delta G_{PT} \cong 21$ (see related considerations in ref 30).

In the next step of the DI and MI mechanisms, as shown in Figure 5, the O2′$^-$ attacks the phosphorus. The activation energy of this step, $\Delta g^{\ddagger}_{NA}$, is estimated from experimentally derived reaction rates. Because data for the exact reactions are not available, we must piece together information from several different reactions in aqueous solution until we capture all the necessary chemistry. We begin with reactions which have a leaving group similar to that in our RNase model. For the DI case we turn our attention to the work of Guthrie, who estimated that the nucleophilic attack of a hydroxide ion on a dimethyl phosphate diester has a rate constant of 6.8×10^{-12} M^{-1} s^{-1}.[27] This rate can be converted to an activation free energy using

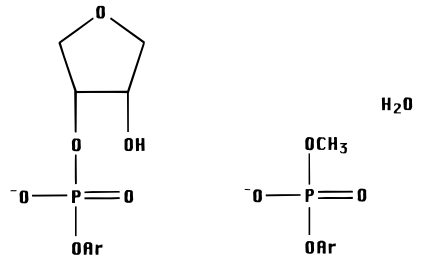
$$\Delta g^{\ddagger} = RT(\ln A - \ln k) \qquad (2)$$

where R is the gas constant, T is the temperature, and A is the pre-exponential factor (RT ln A is ∼18 kcal/mol). Thus, we obtain around 33 kcal/mol for the corresponding activation free energy. This value is based however on somewhat indirect considerations and might represent an overestimate. In particular, we estimate from our ab initio calculations[31] that the corresponding barrier is around 30 kcal/mol. This value was obtained by first performing a direct calculation that gave around 33 kcal/mol and then taking into account the fact that the same type of calculation for OH^- attack on monoester monoanion overestimates the observed value by ∼4 kcal/mol. Here, we take the value of 30 kcal/mol as our best estimate.

For the MI case we cannot use the Guthrie estimate of OH^- attack on trimethyl phosphate since the corresponding rate for attack on neutral methyl phosphate is expected to be quite different (see Discussion in refs 11 and 32). Thus, we consider two estimates. First, we consider the energetics of the attack of OH^- on neutral phosphate monoester as a reasonable estimate for the corresponding energetics for the attack on neutral phosphate diester. This reflects the fact that the main difference between having neutral phosphate monoester and triester is the replacement of CH_3 by H (see ref 11), and the same effect exists in the case of neutral diester. The relevant estimated activation barrier for the attack on neutral phophodiester is ∼13 kcal/mol. Our second estimate is based on a recent study of Wolfenden and co-workers,[33] who studied H_2O attack on phosphate diester dianion and obtained a barrier of ∼36 kcal/mol at 25 °C. We now consider that the energetics of a direct attack by water are very similar to the energetics of a stepwise reaction where the proton is transferred to the phosphate followed by the OH^- attack. Since the free energy of the proton-transfer step are ∼21 kcal/mol (as found from pK_a considerations), the attack by OH^- should have a barrier of about 15 kcal/mol (15 + 21 = 36). The estimate of 15 kcal/mol is taken here as the most reasonable estimate for OH^- attack on a neutral phosphate diester.

The two rate constants should be multiplied by 55 M to allow for the nucleophile to be in the same solvent cage as the phosphate ester. This gives 13 and 28 kcal/mol for the corresponding activation barriers for the MI and DI mechanisms, respectively.

Next we need to study a reaction which contains a more RNA-like system. That is, we should account for the fact that the nucleophilic attack is internal and the reacting molecule contains the five-membered ribose sugar ring. The effect of the ring on the nucleophilic attack step is probably quite small. One way of assessing it is to compare the estimated rate of this step for molecules A and B.



The rate constant for the attack on A is 4.5×10^{-7} s^{-1} at 50 °C (k_5 in ref 34). This is translated to $\Delta g^{\ddagger} \cong 28.6$ kcal/mol. The activation barrier for the nucleophilic attack on B can be estimated from the linear free energy relationship of Kirby and Varvoglis[35] and is ∼28 kcal/mol. Another instructive estimate can be obtained from the work of Taylor and Kluger,[36] who determined that the activation enthalpy of the formation of a cyclic transition state of methyl ethylene phosphate is about 2 kcal/mol lower than the activation enthalpy of the corresponding noncyclic species (see Figure 2 in ref 36). Assuming that in this case the differences in activation enthalpy and activation free energy are similar, one obtains a ring catalytic effect of around 2 kcal/mol. Here, we take this value as a rough estimate of the effect of forming a five-member ring in the nucleophilic attack step.

The adjusted $\Delta g^{\ddagger}_{NA}$ is now ∼11 kcal/mol for the MI mechanism and ∼26 kcal/mol for the DI mechanism. Note that

(28) Guthrie, J. P. Hydrolysis of Esters of Oxy Acids; pK_a Values for Strong Acids; Brønsted Relationship for Attack of Water at Methyl; Free Energies of Hydrolysis of Esters of Oxy Acids; and a Linear Relationship between Free Energy of Hydrolysis and pK_a Holding over a Range of 20 pK Units. *Can. J. Chem.* **1978**, *56*, 2342.

(29) Warshel, A.; Russell, S. T. Calculations of Electrostatic Interactions in Biological Systems and in Solutions. *Q. Rev. Biophys.* **1984**, *17*, 283.

(30) Fuxreiter, M.; Warshel, A. Origin of the Catalytic Power of Acetylcholinesterase: Computer Simulation Studies. *J. Am. Chem. Soc.* **1998**, *120*, 183.

(31) Florián, J.; Warshel, A. Unpublished results.

(32) Florián, J.; Warshel, A. Phosphate Ester Hydrolysis in Aqueous Solution: Associative Versus Dissociative Mechanisms. *J. Phys. Chem. B* **1998**, *102*, 719.

(33) Wolfenden, R.; Ridgway, C.; Young, G. Spontaneous Hydrolysis of Ionized Phosphate Monoesters and Diesters and the Proficiencies of Phosphatases and Phosphodiesterases as Catalysts. *J. Am. Chem. Soc.* **1998**, *120*, 833.

(34) Usher, D. A.; Richardson, D. I., Jr.; Oakenfull, D. G. Models of Ribonuclease Action. II. Specific Acid, Specific Base, and Neutral Pathways for Hydrolysis of a Nucleotide Diester Analog. *J. Am. Chem. Soc.* **1970**, *92*, 4699.

(35) Kirby, A. J.; Varvoglis, A. G. The Reactivity of Phosphate Esters. Monoester Hydrolysis. *J. Am. Chem. Soc.* **1967**, *89*, 415.

(36) Taylor, S. D.; Kluger, R. *J. Am. Chem. Soc.* **1992**, *114*, 3067.

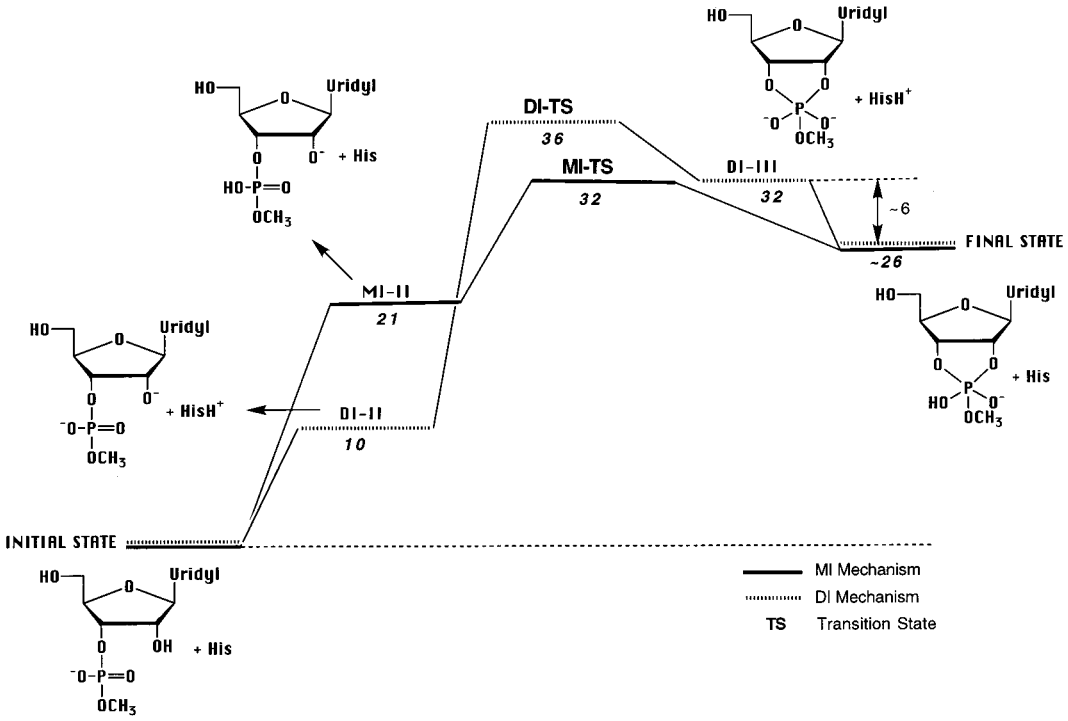


Figure 6. Free energy profiles deduced from experimental and experimentally derived data for the DI and MI mechanisms in water with histidine acting as the general acid−general base. Energies are in kcal/mol.

the value of $\Delta g^{\ddagger}_{NA}$ for the transesterification reference reaction cannot be obtained from the well-known experimental studies of the hydrolysis of cyclic phosphate[37] (the hydrolysis step involves a cleavage of a cyclic system rather than a *formation* of a cyclic system). It is also important to realize that the difference between the rate of hydrolysis of DNA and RNA in water cannot be translated directly to the ring effect in the transphosphorylation step. This difference ($\sim 3 \times 10^4$ according to ref 38) might reflect the hydrolysis step or some other factors (i.e., hydrolysis by OH^- versus hydrolysis by H_2O), and although its actual origin is of great interest, it is out of the scope of the present work.

The reaction free energies, ΔG's, can be also estimated from the corresponding experimental information. The free energy for the nucleophilic attack step, ΔG_{NA}, for the dimethyl phosphate ester monoanion has been estimated by Guthrie to be around 22 kcal/mol.[27] The ΔG_{NA} for the nucleophilic attack on the neutral monoester was estimated by Guthrie to be around 4 kcal/mol. Because the final states in Figure 5 are identical for both mechanisms, they have the same free energies. This energy is obtained by adding the above 4 kcal/mol to the 21 kcal/mol for the proton-transfer step in the MI mechanism. This gives 25 kcal/mol.

The final states in Figure 5 now contain the same chemical structures for both MI and DI mechanisms and are poised for the release of the leaving group. However, the activation free energy for the cleavage step is most likely very similar to that of the nucleophilic attack (see ref 32). Thus, we can obtain the lower limit of the catalytic effect of the protein by studying the mechanisms only up through the nucleophilic attack step. If the protein reduces the barrier for the nucleophilic attack less than the barrier for the cleavage of the leaving group, then the nucleophilic attack is rate limiting and it is sufficient to focus on it. If the barrier of the cleavage is reduced less than the barrier for the nucleophilic attack, then the cleavage step is rate limiting, so that our experimental estimate (see below) will give the lower limit for the effect of the protein on the overall transphosphorylation step. More systematic analysis will be done in the future by studying effects of mutations on the two barriers.

In view of the above analysis, the overall activation barrier in water is ~32 kcal/mol for the MI nonconcerted mechanism and ~47 kcal/mol for the DI mechanism (see Figure 5). It should be noted, however, that these numbers correspond to the mechanisms in a water environment when a water molecule is the general base. Quite a different picture arises when we construct the free energy profiles using histidine as the general base instead of a water molecule, while still retaining the aqueous environment (see Figure 6). The free energy profile of the MI mechanism remains unchanged. The proton-transfer step in the DI mechanism now involves a histidine, with pK_a of ~7, rather than a water molecule. Thus $(\Delta G_0) = 1.38(16 - 7) = 12.4$ kcal/mol. The free energies of moving both His and $HisH^+$ to the same solvent cage with a concentration of 55 M are identical, and therefore, we should not include this concentration effect in our cycle. We should include though the electrostatic interaction between $HisH^+$ and O2′$^-$ (approximately −2 kcal/mol). Thus, we obtain $\Delta G_{PT} \cong 10$ kcal/mol. Of course, since the mechanism is still carried out in aqueous solution, the rate constants for the nucleophilic attack remain the same; the effect of the histidine on the internal nucleophilic attack is small in aqueous solution. Thus the overall activation barrier for the DI mechanism is ~36 kcal/mol.

By comparing the two profiles, it becomes evident that the overall reaction rates of the two mechanisms are now more similar, differing by only a factor of ~900 (the overall $\Delta g^{\ddagger}$'s are ~36 and ~32 kcal/mol for the transphosphorylation step of the DI and MI mechanisms, respectively). To the best of our knowledge, the overall activation barriers for the DI and MI mechanisms in water have not been estimated in other studies.

(37) Kumamoto, J.; Cox, J. R., Jr.; Westheimer, F. H. Barium Ethylene Phosphates. *J. Am. Chem. Soc.* **1956**, *78*, 4858.

(38) Radzicka, A.; Wolfenden, R. A Proficent Enzyme. *Science* **1995**, 267, 90.

It seems to us that the frequently invoked fact that triester hydrolyzes 10^3-10^6 times faster than diester (e.g. refs 4 and 9) cannot be translated to the difference in activation barriers in the mechanisms of Figure 6 without considering the energetics of the relevant proton-transfer steps. Furthermore, the assumption that the phosphotriester provides a good model for the MI system is not fully justified (see refs 11 and 32). These points are not brought here as a criticism of previous studies but to emphasize the importance of a systematic energy-based analysis.

The rate constant, k_{cat}, for the tranesterification step in the reaction of RNase is $\sim$500 s^{-1} depending on the substrate. This is equivalent to a total $\Delta g^{\ddagger}$ of $\sim$15 kcal/mol. Thus, the enzyme would have to provide $\sim$17 and $\sim$21 kcal/mol of catalysis to the MI and DI mechanisms, respectively. Note that Raines obtained an estimate of around 16 kcal/mol for the catalytic effect of RNase (using a standard state of 0.1 mM).[5] The difference between the two estimates reflect in part the use of different reference states. This gives around 21 kcal/mol for 1 M standard state. However, it is quite possible that the actual reaction in water involves the MI mechanism, while in protein, it may involve the DI mechanism. In this case, the agreement between our estimate and that of Raines is coincidental. That is, the DI mechanism in water (that includes His in the reaction system) has a barrier similar to that of the MI in water (that does not include His).

Obviously an enzyme that provides more than 20 kcal/mol catalysis can make one of these mechanisms more favorable than the other. The origin of this enormous catalytic effect is the main subject of the present work. The effect is so large that it presents a major challenge to various proposals for the source of enzyme catalysis. Much work on this subject has shown that, in general, proteins provide catalysis through electrostatic stabilization of the transition state. However, these findings are not yet generally accepted. This might reflect some skepticism of computer modeling studies and/or the electrostatic hypothesis and perhaps the availability of alternative suggestions such as ground-state destabilization.[18] At any rate, accounting consistently for a catalytic effect of around 20 kcal/mol presents an opportunity to discriminate between different catalytic proposals. This challenge is addressed in the next section where, with the use of EVB free energy perturbation (FEP) calculations, we investigate the feasibility of the DI and MI mechanisms and the source of catalysis for RNase A in particular and for enzymes in general.

3. EVB Simulations

3.1. The EVB-FEP Method. Simulations that determine the free energy along a reaction path, as opposed to only structures or dynamics, offer the best opportunity to accurately describe enzyme catalysis. This study uses the EVB method coupled with an umbrella sampling/FEP technique.[26] The EVB method has been described numerous times (e.g., ref 26), and thus, here, we will cover only the major points of relevance.

The enzyme/substrate system is divided into "quantum" and "classical" parts. The quantum part includes the portion of the substrate and the enzyme where bonds are being broken or formed. The atoms of this part are represented by a quantum mechanical Hamiltonian and are referred to as the EVB atoms. The rest of the system is described by a classical force field. The effect of the classical part on the quantum Hamiltonian is obtained through electrostatics, van der Waals, and bonding terms. The EVB Hamiltonian matrix elements are given by

$$\epsilon_i = H_{ii} = \sum_j \Delta M_j^{(i)}(b_j^{(i)}) + {}^1/_2 \sum_m \gamma_m^{(i)} k_m^{(i)} (\theta_m^{(i)} - \theta_{0,m}^{(i)})^2 + \sum_l K_l^{(i)}[1 + \cos(n_l^{(i)}\phi_l^{(i)} - \delta_l^{(i)})] + V_{\mathrm{nb,rr}}^{(i)} + \alpha^{(i)} + V_{\mathrm{nb,rs}}^{(i)} + V_{\mathrm{s}} \quad (3)$$

$$H_{ij} = A_{ij} \exp\{-\mu r\}$$

In the first equation, ΔM_j^i denotes the Morse potential relative to its minimum value for the jth bond in the ith resonance. The second and third terms, respectively, represent the bond angle bending contribution and the torsional angle twisting contribution. The variables b, θ, and ϕ are bond lengths, bond angles, and torsional angles, respectively. The factor γ_l^i in the second term is a coupling between bonds that are being broken or formed and those angles that depend on these bonds, and k_m is the force constant for the angle bending term. The periodicity of the potential is determined by n_l, and the variable δ is the phase shift term. The fourth term denotes nonbonded electrostatic and van der Waals interactions between the reacting groups (denoted by subscript r). The term α^i is the gas-phase energy of the ith state when all fragments are at infinite separation. The nonbonded interaction with the surrounding protein and solvent, denoted by subscript s, is given by $V_{\mathrm{nb,rs}}^i$. The last term represents the internal potential energy of the protein/solvent system. The r term in the H_{ij} expression is the distance between atoms whose bonding is changed upon transfer from state i to state j. The parameters A_{ij} and μ can be adjusted to reproduce the observed barrier for the reaction in solution or to reproduce gas-phase ab initio calculations. The actual ground-state potential surface, E_g, is obtained by diagonalizing the EVB Hamiltonian (see ref 26 for detailed discussion).

The EVB free energy is evaluated by driving the system from the reactant to the product (or intermediate) using a mapping potential of the form

$$E_{\mathrm{m}} = \sum_i \epsilon_i \eta_i^{\mathrm{m}} \quad (4)$$

where $\eta^{\mathrm{m}} = (\eta_1^{\mathrm{m}}, \eta_2^{\mathrm{m}}, ..., \eta_n^{\mathrm{m}})$ is a mapping vector whose components are changed by fixed increments during the mapping process (here, we use η rather than λ to prevent confusion with the reorganization energy). The data generated during this simulation is used to evaluate the reaction free energy, $\Delta g(X)$, using a combination of the FEP and umbrella sampling approaches (see ref 26 for details):

$$\exp[-\beta \Delta g(X)] = \exp[-\beta \Delta G_{\mathrm{map}}(\eta^0 \rightarrow \eta^{\mathrm{m}})] \times \langle \delta(X' - X) \exp[-\beta(E_g - E_{\mathrm{m}})] \rangle_{E_{\mathrm{m}}} \quad (5)$$

where X is the reaction coordinate, $\beta = 1/k_B T$, k_B is the Boltzmann constant, η^{m} is the value of η that keeps the system closest to X, E_g is the ground-state energy of the EVB Hamiltonian, and $\langle\ \rangle_{E_{\mathrm{m}}}$ designates an average over the trajectory with the given E_{m}. $\Delta G(\eta^0 \rightarrow \eta^{\mathrm{m}})$ is the free energy associated with changing the mapping potential from E_0 to E_{m} (this ΔG is obtained by a standard FEP approach).

The reaction coordinate is taken as the difference between the values of the two mapping potentials that change during the given mapping step (e.g., $X = \epsilon_2 - \epsilon_1$ when we change η_1 and η_2). The relevant activation barrier, $\Delta g^{\ddagger}$, is given by $\Delta g^{\ddagger} = \Delta g(X^{\ddagger}) - \Delta g(X_0)$, where X_0 and $X^{\ddagger}$ are the values of the reaction coordinate at the ground state and transition state, respectively.

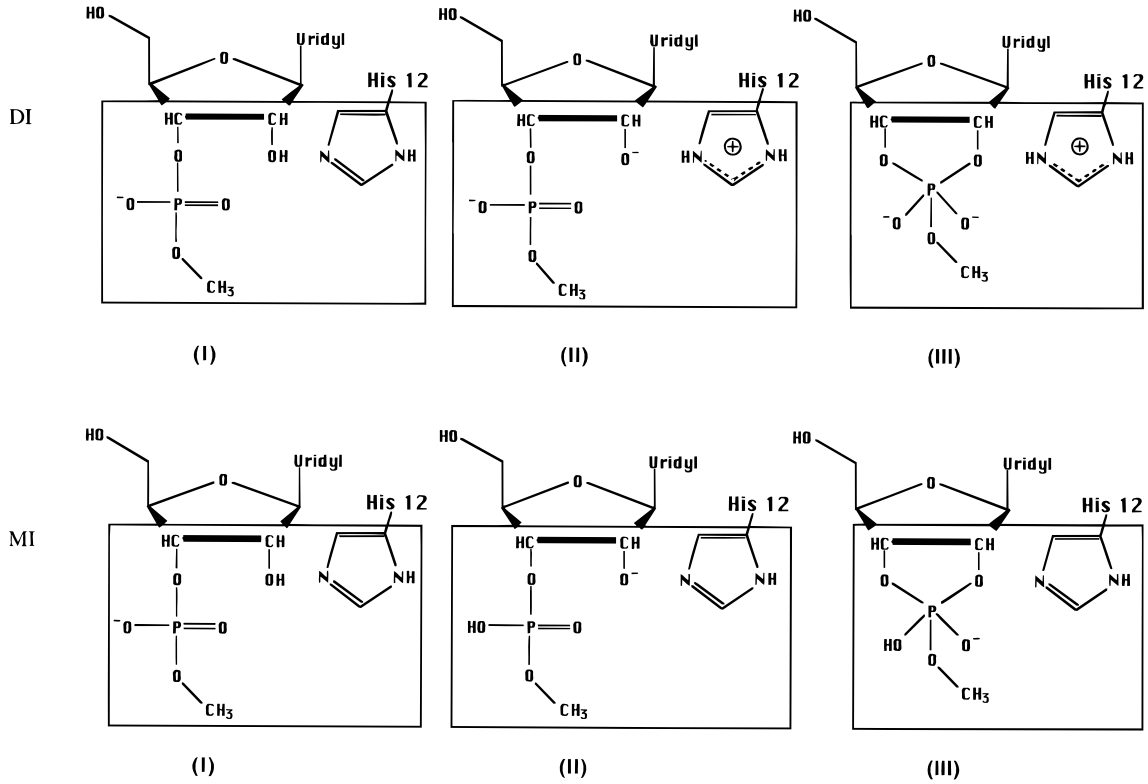


Figure 7. Valence bond structures for the DI mechanism (top) and the MI mechanism (bottom). Atoms enclosed in the marked boxes are defined as EVB atoms.

The effectiveness of the EVB method is evident from the fact that several research groups have recently adopted it for studies of chemical reactions in solutions and enzymes.[39−43]

3.2. EVB-FEP Setup for RNase. The EVB resonance states for the DI and MI mechanisms were selected on the basis of the information given in section 2. These resonance structures are enclosed in the corresponding boxes in Figure 7, where the atoms that are italicized as well as those of the imidazole ring of His 12 are defined as EVB atoms. We also performed other simulations of the DI mechanism in which we used a 2-hydroxyethyl methyl phosphate as the substrate.

The calculations considered involved the resonance structures of Figure 7 and two separate mapping procedures, I → II and II → III. The starting point of our simulations is a structure of RNase A complexed with uridine vanadate based on a joint study of X-ray and neutron diffractions.[6] Water molecules that are part of the structure were removed and replaced by the explicit water molecules generated by the ENZYMIX program.[44] The protein/solvent structure was successively relaxed for 1 ps at 100, 200, and 300 K, with a 1-fs time step. The η_1, η_2, and η_3 parameters were changed in increments of 0.1 starting from (1.0, 0.0, 0.0) to ((0.0, 1.0, 0.0) and then to (0.0, 0.0, 1.0), thus driving the system from the first to the last resonance structure. Here, we used a standard simulation length of 4 ps with a 1-fs stepsize for each of the 11 mapping steps. Relative to some molecular mechanical methods this seems like a short simulation length. However, our approach involves the special surface constrained all atom solvent (SCAAS) boundary conditions[45] and the local reaction field (LRF) for long-range treatment that give faster convergence.[26,44]

The EVB simulations were first performed in water to calibrate the gas-phase shift, α (see eq 3). The same value of α was then used in the protein calculations. The calibration of EVB parameters involves setting the values of α for the individual steps so that the calculated ΔG would reproduce the experimental ΔG. Values of 95 and 220 were used for the proton transfer and nucleophilic steps of the DI mechanism and 45 and 130 for the MI mechanism.

To analyze the mechanisms in water, only atoms of the substrate and His 12 were included in the solute region. The surrounding was modeled by a SCAAS water sphere whose radius was taken as 20 Å, and the resulting spherical system was surrounded by a surface of Langevin dipoles which were then surrounded by a bulk system (see ref 44). Using the same protocol and parameters of the water simulations, both mechanisms were then simulated in the protein environment.

The energetics from the two mapping steps (I → II and II → III) were pieced together to form free energy profiles for the two mechanisms in water and in protein. The amount of catalysis is calculated as $\Delta\Delta g^{\ddagger}_{w\rightarrow p}$, the difference between the overall $\Delta g^{\ddagger}$'s of the mechanism in water and in protein.

The present work has examined the stability of the calculations paying attention to the effect of different restraints. More specifically, since the model used in our simulation may be

(39) Kim, H. J.; Hynes, J. T. A Theoretical Model for SN1 Ionic Dissociation in Solution. *J. Am. Chem. Soc.* **1992**, *114*, 10508.

(40) Bursulaya, B. D.; Zichi, D. A.; Kim, H. J. Molecular Dynamics Simulation Study of Polarizable Solute Solvation in Water. 1. Equilibrium Solvent Structure and Solute Rotational Dynamics. *J. Phys. Chem.* **1996**, *100*, 1392.

(41) Bala, P.; Grochowski, P.; Lesyng, B.; McCammon, J. A. Quantum-Classical Molecular Dynamics Simulations of Proton-Transfer Processes in Molecular Complexes and in Enzymes. *J. Phys. Chem.* **1996**, *100*, 2535.

(42) Lobaugh, J.; Voth, G. A. The Quantum Dynamics of an Excess Proton in Water. *J. Chem. Phys.* **1996**, *104*, 2056.

(43) Neria, E.; Karplus, M. Molecular Dynamics of an Enzyme Reaction: Proton Transfer in TIM. *Chem. Phys. Lett.* **1997**, *267*, 23.

(44) Lee, F. S.; Chu, Z. T.; Warshel, A. Microscopic and Semimicroscopic Calculations of Electrostatic Energies in Proteins by the POLARIS and ENZYMIX Programs. *J. Comput. Chem.* **1993**, *14*, 161.

(45) King, G.; Warshel, A. A Surface Constrained All-Atom Solvent Model for Effective Simulations of Polar Solutions. *J. Chem. Phys.* **1989**, *91*, 3647.

Table 1. Polarizable EVB Results for the Dianion Intermediate Mechanism with Uridine Methyl Phosphate as the Substrate[a]

		constraints (kcal mol^{-1} Å^{-2})				ΔG(PT)[b] (kcal mol^{-1})		ΔG(NA)[c]		$\Delta g^{\ddagger}$(NA) (kcal mol^{-1})		total $\Delta g^{\ddagger}$ = ΔG(PT) + $\Delta g^{\ddagger}$(NA)		
	run	uracil	center	O1P−N	O2P−N	water	protein	water	protein	water	protein	water	protein	$\Delta\Delta g^{\ddagger}_{w\rightarrow p}$
obsd						10		∼22		26		36	15	21
calcd	1	3	3	3	3	14	3	17	13	26	19	40	22	18
	2	0	3	3	3	14	8	19	15	24	22	38	30	8
	3	0	1	1	1	13	8	21	15	29	21	42	29	13

[a] The model includes induced dipoles. The results of each run are an average of two simulations which differ slightly by the position of the center atom. Each simulation involved the following conditions: 44 ps of overall simulation time; 1-fs timestep; $\lambda = 0.1$; 0.03 kcal mol^{-1} Å^{-2} protein constraint. [b] PT denotes the proton transfer step. [c] NA denotes the nucleophilic attack.

Table 2. Polarizable EVB Results for the Monoanion Intermediate Mechanism with Uridine Methyl Phosphate as the Substrate[a]

		constraints (kcal mol^{-1} Å^{-2})				ΔG(PT)[b] (kcal mol^{-1})		ΔG(NA)[c]		$\Delta g^{\ddagger}$(NA) (kcal mol^{-1})		total $\Delta g^{\ddagger}$ = ΔG(PT) + $\Delta g^{\ddagger}$(NA)		
	run	uracil	center	O1P−N	O2P−N	water	protein	water	protein	water	protein	water	protein	$\Delta\Delta g^{\ddagger}_{w\rightarrow p}$
obsd						21		∼4		11		32	15	17
calcd	1	3	3	3	3	22	20	10	6	12	2	34	22	12
	2	0	3	3	3	18	18	5	4	11	3	29	21	8

[a] The model includes induced dipoles. The results of each run are an average of two simulations which differ slightly by the position of the center atom. Each simulation involved the following conditions: 44 ps of overall simulation time; 1-fs timestep; $\lambda = 0.1$; 0.03 kcal mol^{-1} Å^{-2} protein constraint. [b] PT denotes the proton transfer step. [c] NA denotes the nucleophilic attack.

incomplete, we can try to represent the difference between the complete and incomplete model by a correction potential. The restraints introduced by such correction potentials are frequently introduced by harmonic constraints (e.g., the surface constraints used in our solvent models).[45,46] Thus, we will refer to the restraints as "constraints". In modeling enzymatic reactions we use (in addition to the standard surface constraints) three "constraints": (i) position constraints; (ii) distance constraints; and, (iii) protein constraints. Position and/or distance constraints are sometimes necessary to replace the forces which are eliminated due to an incomplete model. For example, the substrate is modeled here as an uridine methyl phosphate, not an RNA strand. Thus, since some important contacts that help hold the substrate in position may be missing, it might be important to add some constraints on the distance between the substrate atoms and protein atoms to keep the substrate from moving away from catalytic residues. The magnitude of the constraints will therefore reflect to some degree the structural incompleteness of the model. In the present study, we have added distance constraints between two substrate oxygens (O1P and O2P) and the reacting nitrogen of His 12. Protein constraints are designed to keep the simulated protein structure similar to the corresponding crystal structure. These constraint allow one to use less than a perfect force field and to retain correct protein structure even in cases of highly charged proteins. For example, we frequently prefer to include the effect of distant ionized residues on a macroscopic level while performing the microscopic simulations with these residues in their neutral state. Having a weak constraint of 0.03 kcal/(mol·Å^2) on all protein atoms allows one to accomplish this task. Of course, the sensitivity of the calculations to the constraint should be examined. This is done here by gradually reducing the strength of the constraints and examining the corresponding effect on the calculated $\Delta\Delta g^{\ddagger}_{w\rightarrow p}$.

3.3. Analysis of the EVB-FEP Results. The results of the simulation are summarized in Tables 1 and 2. These tables report the calculated ΔG's for the DI and MI mechanisms in the optimal simulation conditions. These simulations included the substrate with its uracil side chain and used a polarizable force field and the constraints listed in the tables (the nature of these constraints will be addressed below). The results listed for most of the simulations involve an average of the energetics of two simulations with somewhat different initial conditions. As seen from Table 1 we were able to obtain a $\Delta\Delta g^{\ddagger}_{w\rightarrow p}$ of 18 kcal/mol for the DI mechanism while using the constraints of set 1. This value is only 3 kcal/mol below the experimentally based estimate of ∼21 kcal/mol given in section 2. Upon removing the constraint on the position of the uracil base a significant amount of the catalytic effect is lost. The significance of this result will be discussed in section 4.

Table 2 summarizes the results for the MI mechanism. Here we obtain only 12 kcal/mol for the $\Delta\Delta g^{\ddagger}_{w\rightarrow p}$, which is 5 kcal/mol lower than the corresponding experimentally derived estimate. As in the DI case, some of the catalytic effect is lost upon releasing the uracil position constraint.

The calculated energetics of the DI and MI mechanisms are depicted schematically in Figure 8. The DI results are for run 1 of Table 1 while the MI results are for run 1 of Table 2. As seen from comparison of the two figures we reproduce a large catalytic effect. The effect appears to reflect the catalysis of both the proton transfer and nucleophilic attack steps.

Before discussing the significance of our results it might be useful to explore their sensitivity to different factors. Table 3 examines the effect of simulation time. In these simulations we used a simplified model without induced dipoles to decrease the computation time. Note that in this case we are only concerned about stability rather than the actual value of $\Delta\Delta g^{\ddagger}_{w\rightarrow p}$, which of course depends on the model used. The table reports the results for the dianionic mechanism in water and protein for simulations of 44 ps (4000 steps/mapping interval) and 88 ps (8000 steps/mapping interval). The results for $\Delta\Delta g^{\ddagger}_{w\rightarrow p}$ differ by 2 kcal/mol between the two sets of simulations. Therefore, 44-ps simulations are reasonably long. Furthermore, error ranges associated with different initial conditions, such as starting from different initial protein coordinates, are frequently larger. We thus place more emphasis on averaging over simulations of shorter time with different initial conditions, rather than on performing very long simula-

(46) Warshel, A. Calculations of Chemical Processes in Solutions. *J. Phys. Chem.* **1979**, *83*, 1640.

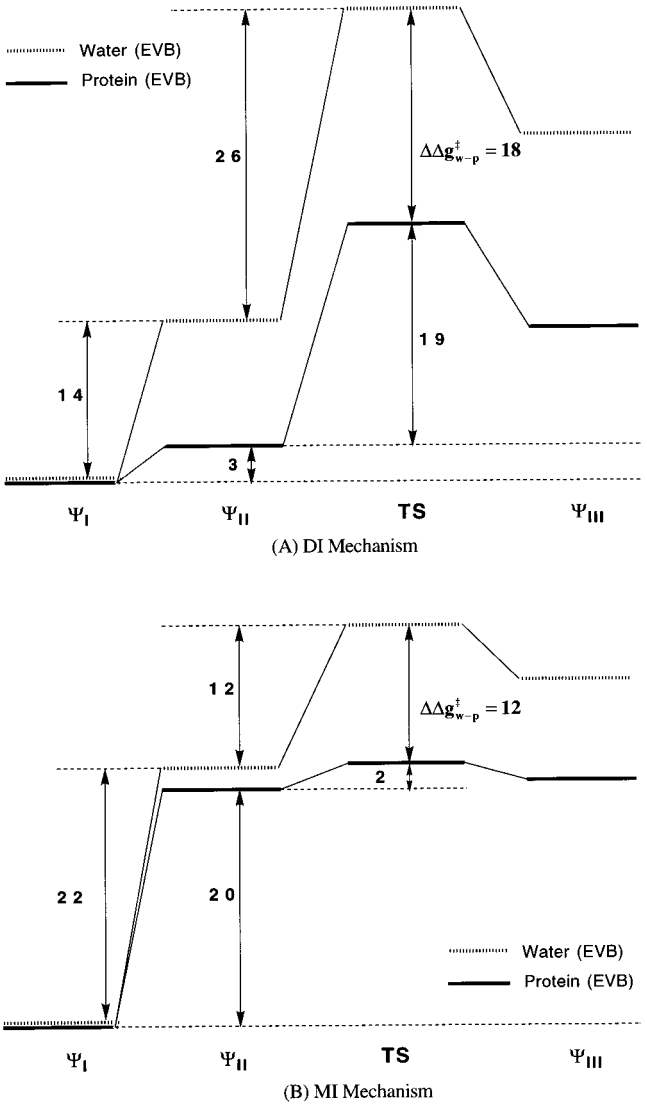


Figure 8. Schematic description of the EVB/FEP free energy profiles in water and protein for the DI (A) and MI (B) mechanisms. The Roman numerals refer to the valence bond structures of Figure 7, and TS refers to the transition state. Energies are in kcal/mol. The profiles are based on the data given in Tables 1 and 2. Note that the EVB parameters are calibrated against several simulations and not against these individual simulations. Thus the free energies for water do not equal the values given in Figure 6.

tions. The fact that averaging on several short trajectories converges faster than one long trajectory has been demonstrated before.[47]

The effect of the constraints (restraints) is examined in Table 4. In considering the effect of the constraints, it is important to examine the total $\Delta g^{\ddagger}$ rather than its individual contributions (ΔG(PT) and $\Delta g^{\ddagger}$(NA)) which are usually less stable. Here again we use a simplified model without induced dipoles to reduce computer time. The results for the DI mechanism are quite sensitive to the protein constraint. In run 1, the protein constraint is 0.30 kcal/(mol·Å^2), and in run 2, it is 0.03 kcal/(mol·Å^2). The stronger constraint reduces the activation energy by an additional 6 kcal/mol, a 10^4 increase in reaction rate. Although this can be considered as a major sensitivity, we note that the 0.03 kcal/(mol·Å^2) is a much more reasonable constraint than the 0.3 kcal/(mol·Å^2) and consider the corresponding results as an interpolation toward the ideal value of zero. Thus, we report only simulations with a 0.03 kcal/(mol·Å^2) protein constraint.

Distance constraints, which involve active site residues, can have a large effect on the activation free energy. One example is the difference between runs 2 and 4 of Table 4. By decreasing the force constant for the constraint on the O2P−N(His 12), O1P−N(His 12) and the center from 3 to 1 kcal/(mol·Å^2), the $\Delta\Delta g^{\ddagger}_{w\rightarrow p}$ decreases by 10 kcal/mol, a change in reaction rate of 10^7. (Note, however, that the effect is much smaller when we include induced dipoles in the model.) Generally, more catalysis can be produced with stronger constraints. However, we like to limit the strength of the distance constraints, as we also do with the protein constraint. We therefore base our final assessment of the RNase energetics on the simulations in which the distance constraints are no greater than 3 kcal/(mol·Å^2), a value which we consider to be a reasonable replacement for the incompleteness of the model (see below). We also limit ourselves to no more than two of these constraints.

Finally, we examine the relationship between the constraints and the molecular completeness of the model, in particular, the completeness of the RNA substrate model. As shown experimentally, the uracil base of the substrate is essential to catalysis because it holds the substrate in place through formation of hydrogen bonds with active site residues.[48] We examined the significance of this base on the catalytic effect produced with our model by performing simulations in which we eliminated it from the substrate. The simplified model consisted of a 2-hydroxyethyl methyl phosphate as the substrate and did not include induced dipoles. No catalytic effect could be established for this mechanism when using "normal-sized" constraints (see Table 5). For example, in a simulation of the DI mechanism in which the constraints were set to 10 kcal/(mol·Å^2), an *anticatalytic* effect of 5 kcal/mol was established. While the detailed effect of having the uracil side chain is left to subsequent studies, we note that, as shown in Table 4, there is a catalytic effect of 10 kcal/mol with only 3 kcal/(mol·Å^2) constraints when the substrate is a uridine methyl phosphate. Without a detailed comparative study, we can only note that in the beginning of the nucleophilic attack step the RMS deviation of the EVB atoms are 2.4 Å in the substrate that does not include uridyl and only 1.3 Å in the one that does include uridyl. Thus, a part of the problem of the substrate without uracil is its initial location in the active site. More detailed studies of this effect will be conducted in the future.

In examining the effect of the uracil, one might try to consider the effect of corresponding position constraints in Tables 1 and 4. Here, however, the analysis is much less unique than in the above case. That is, the loss of catalysis upon removal of the uracil constraint in Table 1, row 2, is probably an overestimate since the conditions of row 3, which have smaller values for the additional constraints, give more catalysis. This might reflect insufficient averaging and will require further studies. In Table 4, on the other hand, we lose 3 kcal/mol of catalysis upon introducing the uracil constraint. This, again, might reflect incorrect averaging or some other effects. Again, a more careful examination of this point is left to future studies. However, regardless of these problems, the above-mentioned effects of removing the uracil is at present far more convincing than the effect of constraining the uracil position. Removing the uracil has indeed a very large effect on the calculated catalysis.

(47) Lee, F. S.; Warshel, A. A Local Reaction Field Method for Fast Evaluation of Long-Range Electrostatic Interactions in Molecular Simulations. *J. Chem. Phys.* **1992**, *97*, 3100.

(48) McPherson, A.; Brayer, G.; Cascio, D.; Williams, R. The Mechanism of Binding of a Polynucleotide chain to Pancreatic Ribonuclease. *Science* **1986**, *232*, 765.

Table 3. Effect of Simulation Time for a Simplified Model That Does Not Include Induced Dipoles[a]

simulation time (ps)	ΔG(PT)[b] (kcal mol^{-1})		ΔG(NA)[c]		$\Delta g^{\ddagger}$(NA) (kcal mol^{-1})		total $\Delta g^{\ddagger}$ = ΔG(PT) + $\Delta g^{\ddagger}$(NA)		$\Delta\Delta g^{\ddagger}_{w\rightarrow p}$
	water	protein	water	protein	water	protein	water	protein	
44	8	5	19	15	30	23	38	28	9
88	8	5	20	15	31	23	39	28	11

[a] Simulations of the dianion intermediate mechanism with uridine methyl phosphate as the substrate. The simulation times of 44 and 88 ps correspond to 4000 steps and 8000 steps, respectively, for each of the 11 mapping steps. The following conditions were used: 1-fs timestep; λ = 0.1; protein constraint = 0.03 kcal mol^{-1} Å^{-2} ; center constraint and uracil constraint = 3 kcal mol^{-1} Å^{-2}; distance constraints on O2−N and O1−N = 3 kcal mol^{-1} Å^{-2}. The results are an average of two or more simulations.

Table 4. Examining the Effect of Different Constraints in a Simplified Model That Does Not Include Induced Dipoles[a]

		constraints (kcal mol^{-1} Å^{-2})					ΔG(PT)[b] (kcal mol^{-1})		ΔG(NA)[c]		$\Delta g^{\ddagger}$(NA) (kcal mol^{-1})		total $\Delta g^{\ddagger}$ = ΔG(PT) + $\Delta g^{\ddagger}$(NA)		$\Delta\Delta g^{\ddagger}_{w\rightarrow p}$
	run	protein	uracil	center	O1P−N	O2P−N	water	protein	water	protein	water	protein	water	protein	
obsd							10		∼22		26		36	15	21
calcd	1[d]	0.30	0	3	3	3	12	0	16	10	25	18	37	18	19
	2[d]	0.03	0	3	3	3	11	5	17	15	31	24	42	29	13
	3[d]	0.03	3	3	3	3	8	5	19	15	30	23	38	28	10
	4[d]	0.03	0	1	1	1	10	9	12	17	24	22	34	31	3

[a] The calculations were done for the DI mechanism with uridine methyl phosphate as the substrate. The model did not include induced dipoles. The simulations involve the following: 1-fs timestep; 4000 steps/interval; λ = 0.1. [b] PT denotes the proton transfer step. [c] NA denotes the nucleophilic attack. [d] An average of two or more simulation sets which differed slightly by the position of the center atom.

Table 5. EVB Results for the Dianion Intermediate Mechanism with 2-Hydroxyethyl Methyl Phosphate as the Substrate[a]

	constraints (kcal mol^{-1} Å^{-2})				ΔG(PT)[b] (kcal mol^{-1})		ΔG(NA)[c]		$\Delta g^{\ddagger}$(NA) (kcal mol^{-1})		total $\Delta g^{\ddagger}$ = ΔG(PT) + $\Delta g^{\ddagger}$(NA)		$\Delta\Delta g^{\ddagger}_{w\rightarrow p}$
run	center	O2P−N	O1−N	O2′−P	water	protein	water	protein	water	protein	water	protein	
1	10	30	30	30	15	4	21	−15	30	11	45	15	30
2[d]	10	10	10	3	10	3	26	14	25	37	35	40	−5

[a] 1-fs timestep; 4000 steps/interval; λ = 0.1; 0.03 kcal mol^{-1} Å^{-2} protein constraint. [b] PT denotes proton transfer. [c] NA denotes nucleophilic attack. [d] An average of two simulation sets for the NA which differed slightly by the position of the center atom.

Like the uracil base, the other bases of the RNA substrate also help in binding. They mainly act through charge−charge interactions, not hydrogen bonding,[48] and thus, the catalytic effect of these bases is not as profound as that of the uracil base. Nonetheless, in comparison to the constraints of 30 kcal/(mol·Å^2) that are needed to correct for the omitted uracil base, constraints that are no greater than 3 kcal/(mol·Å^2) to correct for the binding interactions of the bases which are not present in our model seem reasonable and justifiable. (For further discussion, see section 4.)

In concluding this section we emphasize that when we use induced dipoles (and therefore a more complete model) we can reduce the magnitude of the constraints. The point is clearly seen from a comparison of the $\Delta\Delta g^{\ddagger}_{w\rightarrow p}$'s of Tables 1 and 4. This is consistent with our assessment that the constraints are acting as a correction potential for the "missing parts" of the model. When induced dipoles are included, we obtain a more correct description of our system and can use a smaller correction potential.

4. Discussion

The present EVB simulations reproduced the catalytic effect of RNase in the transphosphorylation reaction. This task has not yet been accomplished by other computational approaches. At this stage, we cannot determine conclusively what is the actual mechanism, since we estimate the error range for $\Delta\Delta g^{\ddagger}_{w\rightarrow p}$ to be ±6 kcal/mol and other alternative mechanisms were not examined. Furthermore, our experimental-based analysis of the reference reaction involves an ∼3 kcal/mol error range. At any rate, the overall picture that emerges from our simulations is an almost equal probability for both mechanisms, with a slightly stronger support (see below) for the DI mechanism than the MI mechanism. A recent instructive study of Herschlag[9] tried to exclude the triester (MI) mechanism based on the fact that replacing a nonbridging oxygen with a sulfur reduces the rate by a factor of ∼5. It was argued that a much larger effect is expected in view of the fact that the thiophosphate has pK_a's of 2.1, 7.3, and 12.3 while phosphate has pK_a's of 1.7, 5.4, and 10.1. Thus, the thiophosphate should have lower proton affinity. However, the effect is quite small (∼3 kcal/mol), and as shown here, the interaction between the protein and the substrate can easily lead to a different trend than that observed in the corresponding solution reactions. Regardless of the uncertainty with regards to the exact mechanism, we have demonstrated how RNase can produce a very large catalytic effect on reasonable and quite likely mechanisms. As evidenced by EVB studies of other enzymes, we expect that, after our analysis of the reference reactions becomes more quantitative and the EVB surface is further calibrated, we will be able to discriminate more between the alternative mechanisms. Furthermore, significant insight will come from future simulation studies of the effect of mutations on each assumed mechanism and from linear free energy relationships (LFER) studies.[26] Therefore, in future works, we hope to build upon these fundamental studies and make a more unique statement as to which mechanism RNase operates by.

The chance for more conclusive analyses should be increased by considering studies of the effect of metals in the hydrolytic cleavage of phosphodiesters (e.g., refs 49−52). Simulating these experiments may help in validating our potential surfaces and

(49) Bashkin, J. K.; Jenkins, L. A. The Role of Metals in the Hydrolytic Cleavage of DNA and RNA. *Comments Inorg. Chem.* **1994**, *16*, 77.

perhaps in finding common features between metal catalysis and the catalysis of RNase A that does not include any catalytic metal ion.

Considering the fact that the present study reproduced the very large catalytic effect of RNase for at least one reasonable mechanism, we may ask ourselves how the enzyme can provide such a large effect. In doing so, we will also address the issue of enzyme catalysis in a general way. We start by noting that many proposals have been put forward to account for the catalytic power of enzymes (e.g., refs 26, 53, and 54) and the use of computational approaches as well as other considerations should help to narrow down the available options. In the Introduction, we already touched upon two of these proposals: one being stabilization of the transition state and intermediates, the other being destabilization of the ground state. To discern how enzymes actually work we have to consider the fact that most enzymes have evolved through mutations that optimize k_{cat}/K_m. In such cases, as discussed thoroughly in other works (e.g., ref 26), it can be demonstrated that ground-state destabilization cannot provide an evolutionary advantage. Moreover, it is being recognized from analysis of mutation experiments[55] that catalytic residues work by either stabilization of both the ground state and transition state or by stabilizing only the transition state. Therefore, it is very likely that the increase of k_{cat} is achieved through the stabilization of the transition state (and intermediates) and not destabilization of the ground state.

Trying to account for transition state (rather than ground state) stabilization is not as simple as it might look. Many proposals (e.g., van der Waals effects, steric effects, entropic effects) cannot help in stabilizing the transition state. One of the most effective ways of stabilizing transition states in enzymes is provided by electrostatic effects. That is, in most enzymatic reactions (or at least those that involve large catalysis), the charge distribution of the reacting fragments changes in a major way along the reaction pathway. The active site can stabilize the transition state by providing an environment which generates a potential that is complimentary to the transition-state charge distribution. In particular, this can be accomplished by an active site with dipoles which are partially preoriented toward the transition-state charge distribution *before* the formation of the transition state.[56,57] This is important since, in the ground state of the reference reaction in water, the solvent dipoles are more or less in a random orientation. Thus, a significant part of the activation energy in water is invested in orienting the dipoles toward the transition state. The preorganized environment in the active site of RNase allows the enzyme to provide more electrostatic stabilization to the transition state and the pentacoordinated intermediate. This catalytic effect can be seen by considering the EVB free energy surfaces (the $\Delta g(x)$'s of eq 5)

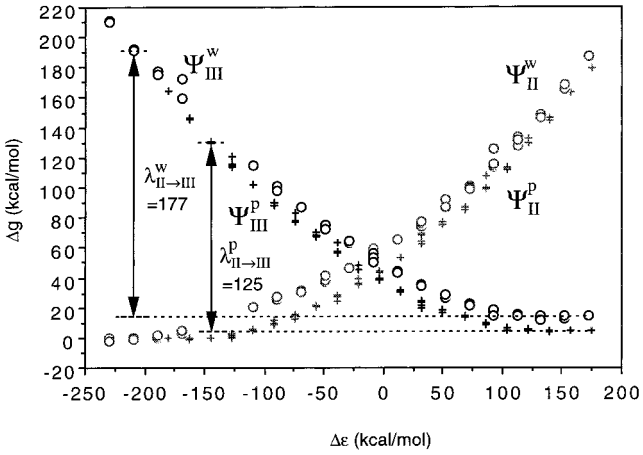


Figure 9. Diabatic free energy curves and the corresponding reorganization energy, λ, for water (○) and protein (+) for the nucleophilic attack step (II → III). λ is calculated as the difference between the gap energy at the minimum of Ψ_{II}, $\Delta\epsilon$, and the exothermicty, ΔG_0.

of Figure 9. As seen from the figure, the Δg curve of state III in the enzyme is pushed down and thus shifted to the left relative to its position in water. This leads to a reduction in the height of the intersection of state II and III and to a reduction of the corresponding $\Delta g^{\ddagger}$ (see ref 26 for discussion). This expression can be analyzed within the framework of the Marcus' formula for electron-transfer reactions or by the more quantitative modification introduced by Warshel and Hwang[58] and referred to here as the WH relationship:

$$\Delta g^{\ddagger}_{ij} = \frac{(\Delta G^0_{ij} + \lambda)^2_{ij}}{4\lambda_{ij}} - H_{ij} + \frac{H^2_{ij}}{(\Delta G^0_{ij} + \lambda)_{ij}} \qquad (6)$$

where λ is the so-called reorganization energy for the given chemical step and H_{ij} is the off-diagonal EVB term. As seen from Figure 9, both λ and ΔG^0 are reduced by ribonuclease. *Both* of these effects are the result of having preoriented dipoles in the enzyme active site.

Now that we obtain a strong indication that RNase A and other enzymes provide their catalytic power by electrostatic stabilization, we can try to explain why RNase might be able to provide more catalysis to the DI mechanism than to the MI mechanism. In general, enzymes operate on mechanisms in which there are large changes in the charge distribution along the reaction pathway. In the MI mechanism there is little change between the charge distribution of the reactant state and the transition state of the nucleophilic attack. In both cases the charge resides on a nonbridging oxygen of the phosphate (see Figure 4). In the DI mechanism, on the other hand, there is a major change in the charge distribution along the reaction pathway. Thus it is easier for the enzyme to attain effective catalysis by providing an environment complementary to the change in charge distribution of the DI mechanism than the MI mechanism.

One of the instructive points of the present study is the dependence of $\Delta\Delta g^{\ddagger}_{w\to p}$ on the presence of the uracil base or on the presence of a constraint that fixes this side chain in the proper position. As discussed in section 3.3, we obtain a much larger catalytic effect in models that include the uracil side chain and in models that constrain the uracil to its correct position. Reproducing this effect is impressive since the base is located

(50) Bashkin, J. K.; Jenkins, L. A. Catalytic Hydrolysis of 2′,3′-Cyclic Adenosine Monophosphate by Aqua(2,2′:6′2″-terpyridine)copper(II): Breakdown of the Analogy between Activated Phosphodiesters and RNA. *J. Chem. Soc., Dalton Trans.* **1993**, 3631.

(51) Chin, J.; Zou, X. Catalytic Hydrolysis of cAMP. *Can. J. Chem.* **1987**, *65*, 1882.

(52) Liu, S.; Hamilton, A. D. Catalysis of Phosphodiester Transesterification by Cu(II)-Terpyridine Complexes with Peripheral Pendent Base Groups; Implications for the Mechanism. *Tetrahedron Lett.* **1997**, *38*, 1107.

(53) Fersht, A. R. *Enzyme Structure and Mechanism*; W. H. Freeman and Co.: New York, 1985.

(54) Jencks, W. P. *Catalysis in Chemistry and Enzymology*; Dover: New York, 1987.

(55) Warshel, A. Electrostatic Origin of the Catalytic Power of Enzymes and the Role of Preorganized Active Sites. *J. Biol. Chem.* **1998**, in press.

(56) Warshel, A. Energetics of Enzyme Catalysis. *Proc. Natl. Acad. Sci. U.S.A.* **1978**, *75*, 5250.

(57) Yadav, A.; Jackson, R. M.; Holbrook, J. J.; Warshel, A. Role of Solvent Reorganization Energies in the Catalytic Activity of Enzymes. *J. Am. Chem. Soc.* **1991**, *113*, 4800.

(58) Hwang, J.-K.; Warshel, A. Microscopic Examination of Free Energy Relationships for Electron Transfer in Polar Solvents. *J. Am. Chem. Soc.* **1987**, *109*, 715.

quite far from the reacting region. Apparently, having the base is crucial for keeping the transition state in the proper site, so it will benefit from the preoriented environment at this site. Interestingly, a recent work of Koshland and co-workers reported a related experiment where the change of a base in nicotinamide adenine dinucleotide phosphate reduced the rate by 5 orders of magnitude.[59] These authors have attributed the effect to the orbital steering mechanism. It should be noted that ref 59 has not presented any structure−energy considerations of the catalytic effect and thus could not rationalize the origin of the very large structural dependence of $\Delta\Delta g^{\ddagger}_{w\rightarrow p}$. The orbital steering concept postulated a very narrow dependence of the transition-state energy on the angle of approach of the reacting molecules. The corresponding rate acceleration has not been well-defined (see below) and is in conflict with the following. First, as was pointed out eloquently by Bruice and co-workers[60] and confirmed by subsequent calculations,[26] this proposal requires an unreasonably large force constant. Second, and in some respects more importantly, the proposal has not been defined correctly since it did not consider the corresponding effect of the reference reaction in water.[26] With a proper reference state one finds that the proposed catalytic effect disappears unless the enzyme can constrain the molecules in the ground state to the same narrow angular range.[26] Doing so, which is extremely unlikely in flexible systems, has little to do with orbitals since we have the same reacting orbitals in the protein and in solution. In this way we simply have another proposal of ground-state destabilization by orientational entropy. However, to obtain a catalytic effect of 5 kcal/mol by entropic factors, we need an enzyme that will restrict the angle of approach of the reacting fragments, in their reactant state, to less than 0.01°, assuming that in solution we have an angle of 30° (this would be the range of the potential surface of less than *RT* above the minimum value).[26] No such angular restriction was ever found in any realistic simulation. The problem is that the orbital steering and the related entropic proposal postulate large ground-state destabilization due to an assumed exact orientation between the reacting fragments and such effects with the flexibility of enzymes are inconsistent. Our electrostatic concept, on the other hand, is associated with the orientation of the reacting transition state relative to its environment (not relative to other reacting fragments) and a relatively small reorganization of this environment during the given reaction. Although the electrostatic model requires much less precise orientation than the orbital steering model, one may wonder how what looks as a rather small structural rearrangement can result in large free energy changes (the energy of flexible systems is not strongly dependent on small structural changes).[26] That is, in general, one has to think on two limiting equilibrium configurations, $\mathbf{r}_1$ and $\mathbf{r}_2$, of the reactive part of the substrate and its surrounding environment in the respective cases of the full substrate (S_1) and the substrate without the base (S_2). If the free energy of moving the reacting region of the substrate without the base from $\mathbf{r}_2$ to $\mathbf{r}_1$ is very small, then we will obtain similar $\Delta\Delta g^{\ddagger}_{w\rightarrow p}$ values for both substrates (since we create for both of them similar catalytic configurations). Fortunately, as stated above, the simulations did reproduce a large reduction of $\Delta\Delta g^{\ddagger}_{w\rightarrow p}$ upon deletion of the base or reducing its position constraint. This must mean that the overall structural changes upon transfer from $\mathbf{r}_2$ to $\mathbf{r}_1$ are quite large when one considers the full dimensionality of the many coordinates involved in this change (even if individual coordinates undergo small changes). Apparently, the free energy involved in changing $\mathbf{r}_2$ to $\mathbf{r}_1$ must be rather large and reflect significant correlated structural changes. Since the large $\Delta\Delta g^{\ddagger}_{w\rightarrow p}$ obtained in $\mathbf{r}_1$ appears to be associated with electrostatic factors[61] we conclude that the loss of catalysis upon motion to $\mathbf{r}_2$ is also an electrostatic effect. More detailed studies should involve detailed analysis of the difference between $\mathbf{r}_1$ and $\mathbf{r}_2$ and comparison of the relevant calculated and observed structure (a very limited comparison was made in section 3.3). This will help to validate our initial finding and to shed light on the nature of the above-mentioned electrostatic effects (i.e., the substrate and/or protein rearrangement as the result of omitting the base. In view of the many points covered in this work, we believe that a systematic examination of such an important issue should be left to subsequent studies.

In summary, the present study has not resolved the mechanistic controversy about RNaseA nor presented a definitive analysis of the catalytic power of this enzyme (this requires a consideration of all feasible mechanisms). However, it has been demonstrated that the enzyme provides an enormous environmental effect and that this effect is reproduced by our model in the mechanisms considered.

Acknowledgment. This work was supported by NIH Grant GM24492. We thank Dr. Jan Florián for insightful discussion and Dr. Jörg Bentzien for helpful comments regarding the manuscript.

JA981594X

(59) Mesecar, A. D.; Stoddard, B. L.; Koshland, D. E., Jr. Orbital Steering in the Catalytic Power of Enzymes: Small Structural Changes with Large Catalytic Consequences. *Science* **1997**, *277*, 202.

(60) Bruice, T. C.; Brown, A.; Harris, D. O. On the Concept of Orbital Steering in Catalytic Reactions. *Proc. Natl. Acad. Sci. U.S.A.* **1971**, *68*, 658.

(61) The EVB calculations reflect, of course, the electronic structure of the solute and therefore any orbital overlap effect. Yet, the calculated change in $\Delta\Delta g^{\ddagger}_{w\rightarrow p}$ is almost entirely due to electrostatic contributions.