VOLUME 117, NUMBER 47
NOVEMBER 29, 1995

JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Structure-Energy Analysis of the Role of Metal Ions in Phosphodiester Bond Hydrolysis by DNA Polymerase I

Michael Fothergill,[⊥] Myron F. Goodman,[‡] John Petruska,[‡] and Arieh Warshel*,[†]

Contribution from the Department of Chemistry, and Department of Biological Sciences, Hedco Molecular Biology Laboratories, University of Southern California, Los Angeles, California 90089–1062

Received May 26, 1995[⊗]

Abstract: The detailed mechanism of DNA hydrolysis by enzymes is of significant current interest. One of the most important questions in this respect is the catalytic role of metal ions such as Mg^{2+}. While it is clear that divalent ions play a major role in DNA hydrolysis, it is uncertain what function such cations have in hydrolysis and why two are needed in some cases and only one in others. Experimental evaluation of the catalytic effects of the cations is problematic, since the cations are intimately involved in substrate binding. This problem is explored here by using a theoretical approach to analyze and interpret the key structural and biochemical experiments. Taking the X-ray structure of the exonuclease domain in the Klenow fragment of E. coli DNA polymerase I we use the empirical valence bond method to examine different feasible mechanisms for phosphodiester bond cleavage in the exonuclease site. This structure-function analysis is based on evaluating the activation free energies of different assumed mechanisms and comparing the calculated values to the corresponding experimentally observed activation energy for phosphodiester bond cleavage. Mechanisms whose calculated activation energies are drastically larger than the observed activation energy are eliminated and the consistency of the corresponding conclusion is examined in view of other available experimental facts including mutational and pH dependence studies. This approach indicates that phosphodiester bond hydrolysis involves catalysis by an OH^- ion from aqueous solution around the protein, rather than a general base catalysis by an active site residue. The catalytic effect of two divalent metal cations in the active site is found to be primarily electrostatic. The first cation provides a strong electrostatic stabilization to the OH^- nucleophile, while the second cation provides a very large catalytic effect by its interaction with the negative charge being transferred to the transition state during the nucleophilic attack step. The calculations also demonstrate that the second metal ion is not likely to be involved in a previously proposed strain mechanism. The two-metal ion catalytic mechanism is compared to the action of a single-metal cation active site and some general rules are discussed. Finally the relationship between the present computer modeling study and available experimental information on DNA hydrolysis is discussed, emphasizing that calculations of absolute rate constants should be, at least in principle, more effective in eliminating incorrect mechanisms than calculations of mutational effects.

1. Introduction

Recent structural and biochemical studies have shed new light on the molecular details of both the hydrolysis and formation of phosphodiester bonds in DNA and RNA.[1-4] Particular attention has been paid to the role of divalent metal ions in such important enzyme systems as polymerases[4] and

[⊥] Department of Chemistry.
[‡] Department of Biological Sciences.
[⊗] Abstract published in *Advance ACS Abstracts*, November 1, 1995.

(1) Uchimaru, T.; Uebayasi, M.; Tanabe, K.; Taira, K. *FASEB J.* **1993**, *7*, 137.
(2) Yarus, M. *FASEB J.* **1993**, *7*, 31.
(3) Steitz, T. *Current Biol.* **1993**, *3*, 31.
(4) Pyle, A. *Science* **1993**, *261*.

0002-7863/95/1517-11619$09.00/0

ribozymes.[4-6] Major clues have been obtained concerning the positions of metal ions in active sites involved in forming and breaking phosphodiester bonds, but the exact mechanisms have yet to be determined. It is not clear why different systems use different numbers of metal ions and different types of bases for catalysis.

Our studies of phosphodiester bond hydrolysis in catalytic sites containing a single metal cation[7,8] have indicated that the divalent cation plays a key role in stabilizing the nucleophile OH^-; without such stabilization this nucleophile would not reach a reasonable concentration in the active site. It was also demonstrated that a divalent cation contributes to catalysis by stabilizing the negative charge that is being transferred from OH^- to phosphate during the nucleophilic attack step. The simulations indicated that a single cation can perform both functions by moving toward the phosphate during the formation of the transition state. It is not clear, however, that such a catalytic configuration can be effective in all cases. In fact, the discovery of active sites with two divalent cations points to an alternative mechanism, having each cation catalyzing a different step of the hydrolysis.

The present work explores these issues, focusing on the catalytic reaction of the 3′−5′ exonuclease (exo) site in the Klenow fragment (Kf) of E. coli DNA polymerase I and examining the role of two metal ions in this site. The catalysis of phosphodiester bond cleavage in this site is simulated by using the corresponding X-ray structure and the empirical valence bond method. The simulations favor an "external OH^- catalysis" where an OH^- ion from aqueous solution around the protein serves as the nucleophile. The role of divalent cations in this mechanism is quantified and generalized.

2. Method of Analysis and General Considerations of the Energetics of Phosphodiester Bond Hydrolysis

While many ways are available to discuss phosphodiester hydrolysis in a qualitative way,[2,9] we are interested here in considering different mechanisms in a quantitative way, trying to identify the actual effect of the given protein environment on the assumed reaction mechanism. Currently the empirical valence bond (EVB) formulation[10] is the simplest and perhaps most effective method for quantitative analysis of phosphodiester bond hydrolysis. In this conceptual framework, one represents the reacting system in terms of "pure" valence bond (VB) resonance structures or states. For example, in the case of phosphodiester bond cleavage (Figure 1), one usually considers the protonated nucleophile (state I), the negatively charged nucleophile (state II), the pentacoordinated complex (state III), and the broken P−O bond (state IV). Such free energy functions have been found by simulation studies to behave as parabolic functions, (e.g., see discussion in ref 11).

The reaction potential surface is obtained by mixing the different VB states, and the corresponding free energy surface is usually analyzed along a reaction coordinate that represents the reorganization of the reacting fragments as well as the surrounding environment. The reaction coordinate transforms the system from configurations that stabilize the reactants to those that stabilize the product (see refs 7 and 10 for more details). While the actual free energy surface obtained by mixing the VB states is of great importance, it appears that one can gain useful insight by simply considering the free energies of the pure VB states.[10] In fact, the effect of the enzyme on the activation barriers is linearly correlated with its effect on these states.[11] The description of the energetics of enzyme catalysis in terms of the energy of the relevant VB states provides a powerful way to reach quantitative conclusions which are not available when one uses more customary approaches where the reaction is described by arrows that represent movement of electrons. For example, when one argues that a metal ion polarizes a bond and that this helps to catalyze a reaction, it is hard to assess how large this effect is. On the other hand, the EVB approach evaluates the interaction between the metal and the substrate charges in the relevant VB states and so *determines* the corresponding effect on the activation energy (e.g., see ref 8).

The practical implementation of the EVB method involves the use of the molecular simulation program ENZYMIX[12] where the different VB states are described by a set of potential functions that consider the bonding and charges of each state and its interaction with the surrounding environment (enzyme and/or solution). The free energy functions of the type diagramed in Figure 1 are evaluated by a free energy perturbation (FEP) and umbrella sampling procedure (see for example refs 7 and 10). The reaction free energy surface is then obtained by mixing the pure VB states and using an umbrella sampling procedure to obtain the corresponding free energies.[10] Without going into further detail, we only note that the simulations involve a mapping potential of the form

$$\epsilon_m = \sum \lambda_i^m \epsilon_i \qquad (1)$$

where ϵ_i is the potential surface that describes the ith VB state, λ_i being mapping parameters that allow one to transfer the system from state I to state IV and to evaluate the relevant free energies. The actual potential parameters used here to describe the ϵ_i for phosphodiester bond hydrolysis are very similar to those used for staphylococcal nuclease [Snase].[7] The parameters that describe the nonbonded interaction between the Mg^{2+} ion and the oxygen of either water or OH^- are $A = 70$ and $B = 32$ (see ref 7 for a definition of these parameters). These parameters were calibrated to reproduce the observed solvation energy of Mg^{2+} in water and the pK_a of water bound to Mg^{2+} in aqueous solution (see ref 13 for a related refinement). The main point in this and in any other EVB parametrization procedure is that the potential parameters are adjusted to reproduce experimental observations in solution (and sometimes in the gas phase), but they are *never* readjusted when the system is transferred from solution to the enzyme active site. This provides a rather unbiased way of performing a quantitative structure function correlation in proteins. It might be also useful to point out that at present it is very hard to obtain quantitative activation barriers by standard semiempirical approaches, since the energies for the reaction in the gas phase and solution might involve substantial errors (especially when one deals with phosphate hydrolysis). The EVB circumvents this problem by forcing the calculations to reproduce the relevant observed energies in solution. In this way the calculations reflect the *change* in solvation energy for moving the reacting fragments from water to the enzyme active site, without reflecting the uncertainties associated with the quantum mechanical part of the calcula-

(5) Piccirilli, J.; Vyle, J. S.; Caruthers, M. G.; Cech, T. R. *Nature* **1993**, *361*.
(6) Long, D.; Uhlenbeck, O. *FASEB J.* **1993**, *7*, 25.
(7) Åqvist, J.; Warshel, A. *Biochemistry* **1989**, *28*, 4680.
(8) Åqvist, J.; Warshel, A. *J. Am. Chem. Soc.* **1990**, *112*, 2860.
(9) Westheimer, F. H. *Science* **1987**, *235*, 1173.
(10) Warshel, A. *Computer Modeling of Chemical Reactions in Enzymes and Solutions*; John Wiley & Sons, Inc.: 1991.
(11) Warshel, A.; Schweins, T.; Fothergill, M. *J. Am. Chem. Soc.* **1994**, *116*, 8437.
(12) Lee, F. S.; Chu, Z. T.; Warshel, A. *J. Comput. Chem.* **1993**, *14*, 161.
(13) Aqvist, J. *J. Phys. Chem.* **1991**, *95*, 4587.

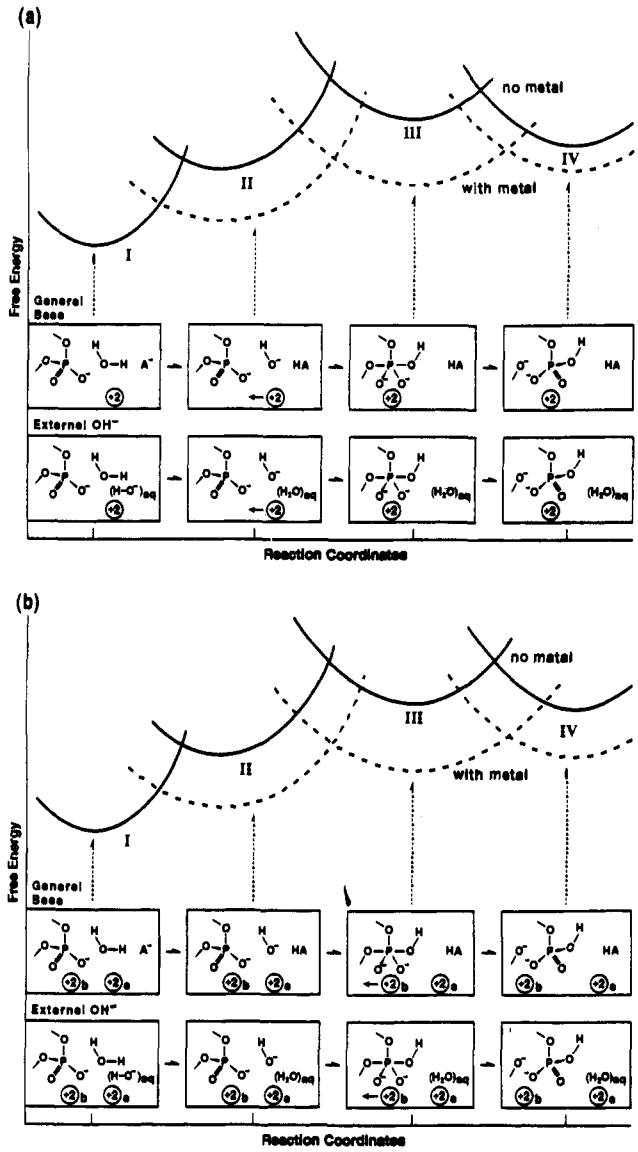

Figure 1. (a) The effect of a single metal ion on the energetics of the second and third resonance structures in general base and specific base mechanisms. As indicated by the figure, an effective metal should be able to move with the nucleophile toward the phosphate center. (b) A schematic representation of the effect of two metal ions on the energetics of phosphate hydrolysis. In this case, the first metal is involved in stabilizing the OH^- ion, while the second metal stabilizes the charge that comes to reside on the phosphate center.

tions (which would be crucial if we were to calculate the energetics of the reaction in water using a first principle approach).

Before we consider the results of the simulations, let us apply the EVB conceptual framework to set the stage for the actual calculations. We start by considering the first step of the reaction, the formation of the nucleophile OH^- (I → II in Figure 1a). At first glance, OH^- formation might appear easy to achieve, by simply ionizing a H_2O to give OH^- and H^+. However, because of the high pK_a of H_2O, this ionization costs 21.6 kcal/mol at room temperature in aqueous solution[14] (to obtain this number one has to consider both the pK_a and the concentration of water). Hence, the OH^- nucleophile of state II is unlikely to be found in the active site, unless its negative charge is stabilized much more effectively than in pure water. This is where a positively charged metal ion can play a major role in the two alternative mechanisms illustrated in Figure 1.

(14) Eigen, M.; de Maeyer, L. Z. *Elektrochem.* **1955**, *59*, 986.

The first mechanism involves a general base catalysis where the active site water donates a proton to a proton accepting residue (A^- in Figure 1a). In the presence of A^-, proton transfer (PT) from water to give OH^- and the conjugated acid (AH) has a free energy cost[8,10,15]

$$(\Delta G_{PT})^{p}_{\text{general base}} = 2.3RT(pK_a^p(H_2O) - pK_a^p(AH)) + \Delta V_{QQ}/\epsilon \quad (2)$$

where p designates the protein site where the basic group A^- is located along with a catalytic metal cation. $\Delta V_{QQ}/\epsilon$ is the change in effective electrostatic interaction between the protein donor and acceptor, which is usually small, since the effective dielectric constant for the interaction is usually large.[16]

A divalent cation in the proper environment and proper position may be able to reduce the pK_a of bound water by stabilizing the OH^- ion, or in the formulation of Figure 1a by stabilizing state II. Alternatively, an OH^- ion from aqueous solution may be brought into the active site to carry out "external OH^-" catalysis (Figure 1b). In this case, the free energy change for the proton transfer process is given by[15]

$$(\Delta G_{PT})^{p}_{\text{external OH}^-} = 2.3RT(pK_a^p(H_2O) - pH) \quad (3)$$

Here again, the metal ion helps to stabilize OH^- and thereby reduce the pK_a for water as proton donor.

Although Figure 1a makes intuitive sense, it appears that the major role of a divalent cation in providing the nucleophile with electrostatic stabilization has not been generally appreciated. Some studies have assumed that RNA hydrolysis involves abstraction of a proton from a ribose hydroxyl (ROH) by an OH^- ion bound to Mg^{2+} in the active site.[2,17] In this mechanism, $Mg^{2+}OH^-$ is viewed as a general base or proton acceptor.[17] It has also been suggested that the metal provides a structural element that orients the attacking OH^-.[2] A recent study has suggested that the metal ion helps to "facilitate" OH^- formation,[18] but no specific details were given to show how this facilitation is accomplished, or what its magnitude is. The idea that metal ions contribute to catalysis by providing electrostatic stabilization to the OH^- nucleophile has received quantitative support by early computer simulation studies of staphylococcal nuclease (Snase)[7,8] and in more recent carbonic anhydrase studies.[19] In fact, the electrostatic role of the metal and the quantitative value of the EVB approach have been clearly established in the case of carbonic anhydrase where the large effect of the active site metal on the pK_a of the bound water is known experimentally (see discussion in ref 19).

In addition to the analysis of step I → II, we also apply the EVB formulation of Figure 1 to analyze the formation of the pentacoordinated state (state III). As is obvious from Figure 1, phosphodiester bond hydrolysis can be catalyzed by factors that stabilize state III. Thus a divalent cation that stabilizes the negative charge being transferred to phosphate in the II → III step can contribute in a major way to catalysis.[7,8] This intuitively clear point has not been generally accepted, although it has been recognized that metal ions can have an important role in stabilizing the transition state. It has been suggested that the metal ion can provide this effect by a strain mechanism that favors the transition state geometry[2,18,20] and by acting as a Lewis acid on the leaving group.[18] It was also suggested that

(15) Warshel, A. *Biochemistry* **1981**, *20*, 3167.
(16) Warshel, A.; Russell, S. T. *Quart. Rev. Biophysics* **1984**, *17*, 283.
(17) Haydock, K.; Allen, L. C. In *Progress in Clinical and Biological Research*; Rein, R., Liss, A. R., Eds.; 1985; Vol. 172A, pp 87–98.
(18) Steitz, T. A.; Steitz, J. A. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 6498.
(19) Åqvist, J.; Warshel, A. *J. Mol. Biol.* **1992**, *224*, 7.

the metal acts in neutralizing the charges of the phosphate, thus making it easier for the OH^- to attack, but it is not clear how to convert this concept to a quantitative estimate of the corresponding rate acceleration.

Here we perform an EVB analysis in an attempt to interpret the available structural and biochemical information and to draw general conclusions about metal-ion catalyzed mechanisms of phosphodiester bond hydrolysis. A recently reported computer simulation of hydrolysis in the active site of Snase, which involves a single divalent cation, is compared to a simulation in the exo site of DNA polymerase I Kf, which involves two divalent cations. These two limiting cases are considered schematically in Figure 1.

3. Results and Discussion

3.1. A Single-Metal Mechanism in Snase. The catalytic reaction of Snase provides an example of the effectiveness of a single divalent cation in phosphodiester bond hydrolysis. The main features of this mechanism simulated recently by the EVB method[7,8] are shown schematically in Figure 1a. The simulations indicated that the positively charged ion can reduce the pK_a of the catalytic water (thus helping in the I → II step) and subsequently also stabilizes the negative charge which comes to reside on the pentacoordinate transition state (helping to stabilize state III). This dual role of the divalent cation has a constraint in terms of evolutionary design. The cation must be able to both reduce the pK_a of bound water and also to move with the OH^- toward the $P(O_2^-)R_2$ center, which requires that the interaction between the metal and its ligands not be too strong.

Another important issue concerns the role of general base in Snase. Recent studies[21,22] question the role of negatively charged Glu 43 as a general base. It was argued that the pH profile of the native protein, being similar to that of mutants having Glu 43 replaced by other residues, does not support a general base mechanism. This argument is clearly relevant at high pH (see below) and is not inconsistent with calculation,[7] showing that the metal rather than the general base provides the most important catalytic effect of the first step. However, it is possible that the observed pH dependence can be described by two competing processes: general base mechanism at low pH and specific base catalysis by OH^- at higher pH. The energetics of the general base mechanism is determined by the pK_a difference between Ca^{2+} bound water and Glu 43 (see eq 2). On the other hand, the energy of the specific base mechanism is determined by the difference between the pK_a of Ca^{2+} bound water and the external pH (see eq 3). Obviously, the relative probabilities of these two processes depend on pH.

Another fundamental issue is related to the binding of metal ions by the active site. When the protein uses negatively charged residues to bind the divalent cation, it might overscreen the field that the cation exerts on the RO^- nucleophile, making the cation an ineffective catalyst. This issue does not appear so serious in the case of the single Ca^{2+} in Snase but becomes rather important in the case of two Mg^{2+} ions in the exonuclease site of DNA polymerase (see below).

3.2. Formation of OH^- Nucleophile in the Exo Site of the Klenow Fragment. A catalytic motif with two metal ions has been observed in the exonuclease site of DNA polymerase I Kf[18] and other systems of phosphodiester bond hydrolysis.[23-27] In order to explore the energetics of such hydrolytic motifs, we used the EVB method to simulate several possible mechanisms of the hydrolytic reaction in the exo site. The objective is to compare the calculated activation energies with the corresponding observed values in order to eliminate energetically unfavorable options.

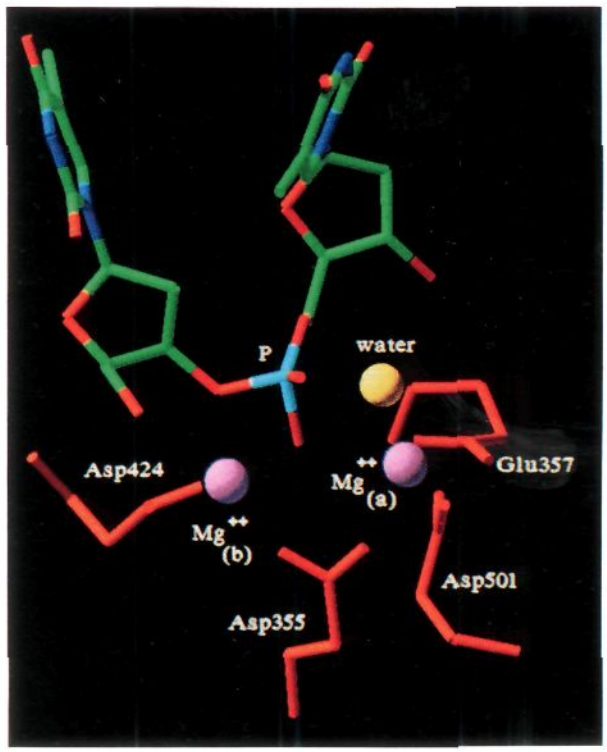

Figure 2. The substrate, two catalytic metals, and one water molecule, and some important residues in the exo site of the Klenow fragment. The rest of the protein and water molecules are not shown for clarity.

The first simulation study examined the general base mechanism described in Figure 1b, with H_2O as a proton acceptor. This reaction has been described by the resonance structures depicted in the corresponding diagram (Figure 1b). The simulation was carried out using the ENZYMIX program,[12] starting with the X-ray structure of ref 20 (whose active site is presented in Figure 2), and involved equilibration of 20 ps and a subsequent FEP/Umbrella sampling procedure of 10 steps of 5 ps each. The corresponding results were averaged over three runs with different initial configurations generated during the process. The simulation used time steps of 1 fs at 300 K and with the SCAAS boundary conditions.[12] These boundary conditions together with the recently developed local reaction field (LRF) long-range treatment[12] allow us to obtain reliable free energies in highly charged active sites. The parameters used were described in the previous section and the coordinates of the different intermediates can be obtained from ref 28. The results of this simulation of the general base mechanism are summarized in Figure 3a. The proton transfer step in the protein appears less likely than the corresponding step in solution, and the Mg^{2+} ion does not help in catalyzing this step. Apparently, the Mg^{2+} ion does stabilize the OH^- ion since the pK_a of the bound water is reduced (see below). However, at the same time, the free energy of proton transfer from this water to the second water molecule is larger than the corresponding free energy in water (see Figure 3a). This energy increase reflects the fact that the energy of the proton transfer depends on the difference between the pK_a of the donor (the catalytic water) and the protonated acceptor (the H_3O^+ ion) and on the change in the effective electrostatic interaction between the donor and acceptor (the $\Delta V_{QQ}/\epsilon$ term of eq 2). It appears that the active site and the Mg^{2+} ion destabilize the H_3O^+ ion and/or reduces the ΔV_{QQ} term relative to the corresponding value in water. Another way to state this point is to say that the $OH^-H_3O^+$ ion pair is less

(20) Beese, L. S.; Steitz, T. A. *EMBO J.* **1991**, *10*, 25.

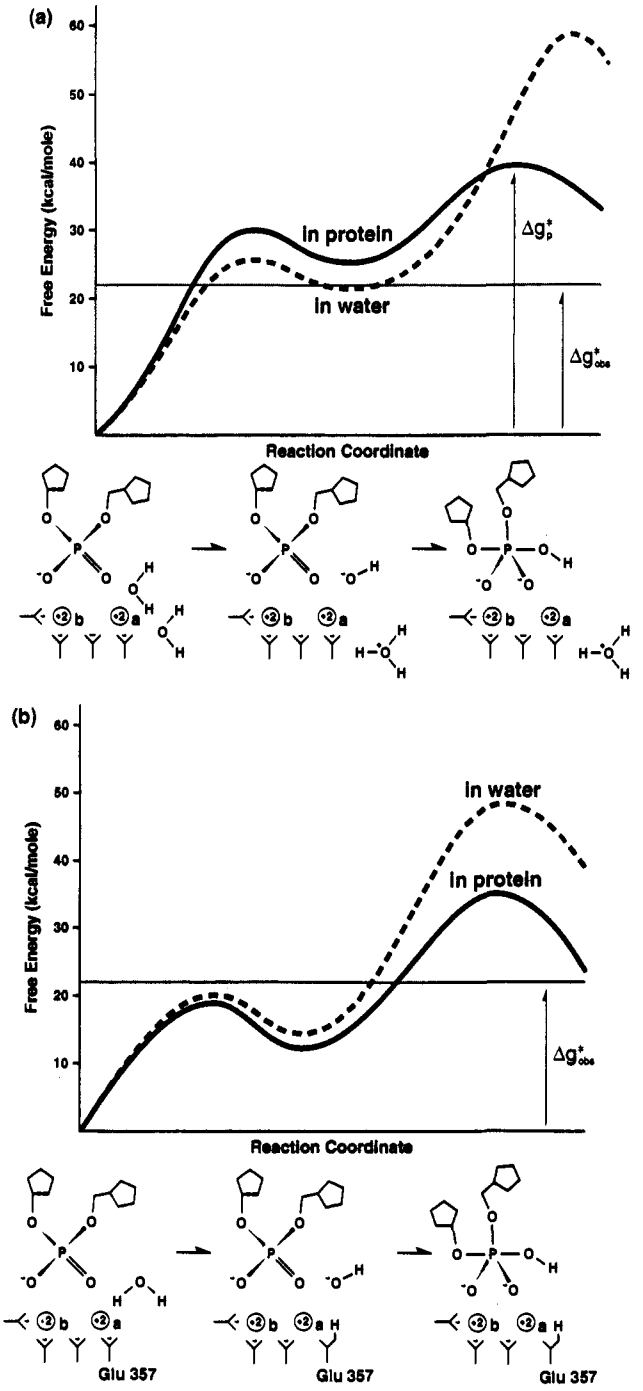

Figure 3. (a) The EVB free energy surfaces for a mechanism that involves a proton transfer between two water molecules in the exo site of the Klenow fragment. The calculated free energy barrier is drastically higher than the corresponding observed barrier thus making the mechanism considered rather unlikely. (b) The EVB free energy surface for a mechanism that involves a proton transfer between a water molecule and Glu 357. The calculated activation free energy is significantly higher than the corresponding observed barrier, thus making the mechanism considered unlikely.

stable in the active site than in solution, despite the fact that the isolated OH^- is more stable in the active site than in water. It is interesting to point out that the site of the Mg^{2+} ion could have provided more stabilization to the OH^- ion in the absence of the negatively charged ligands but then it would be harder to bind the cation. In any case, regardless of the exact reasons, our calculations indicate that the overall calculated activation barrier for the mechanism of Figure 3a is far too high to account for the observed rate constant (10^{-3} and 10^{-1} s^{-1} for duplex DNA and poly (dA), respectively, at pH = 7.5,[29] which corresponds to 22 and 19 kcal/mol activation barriers, respectively).

It is useful to point out that the calculated reaction rate that should corresponds to the observed k_{cat} is evaluated from the overall activation barrier rather than any of the individual activation barriers for the assumed reaction steps. This is entirely justified when the reaction profile is similar to that of Figure 3. In such a case we can assume a kinetic scheme of the form

$$B + S + H_2O \underset{k_{II\to I}}{\overset{k_{I\to II}}{\rightleftharpoons}} BH^+ + S + OH^- \xrightarrow{k_{II\to III}} BH^+ + (SOH)^- \tag{4}$$

where B and S are the base and substrates, respectively, and since in our case $k_{II\to III} < k_{II\to I}$ we can write

$$\begin{aligned} \text{rate}_{I\to III} &= k_{cat}[B][S][H_2O] \qquad (5)\\ &\simeq k_{II\to III}[BH^+][S][OH^-]\\ &\simeq k_{II\to III}\left(\frac{k_{I\to II}}{k_{II\to I}}\right)[B][S][H_2O]\\ &\simeq A\exp(-(\Delta g_{II\to III} + \Delta G^0_{I\to II})/k_BT)[B][S][H_2O]\\ &= A\exp(-(\Delta g^{\ddagger}_{I\to III}/k_BT)[B][S][H_2O] \end{aligned}$$

where the preexponential factor, A, is $\sim 6 \times 10^{12}$ s^{-1} for reactions in room temperature in condensed phase with significant activation barriers (see ref 10). Although we assumed a quasi-equilibrium for the first step, the same result would be obtained with more rigorous treatments as long as $\Delta G_{I\to II} > 0$ and $\Delta g^{\ddagger}_{I\to III} > \Delta g^{\ddagger}_{I\to II}$. Considering that the above mechanism yields calculated energies inconsistent with the observed energetics, it is important to consider alternative mechanisms. In particular we might consider a mechanism that makes the Mg^{2+} ion a more effective charge stabilizing element. The simplest option is to protonate one of the ionized acids that ligate this ion, and Glu 357 appears to be a likely candidate. In order to explore this possibility we simulated the corresponding mechanism by the EVB approach. The results of the simulation are presented in Figure 3b.

As seen in Figure 3b, the general base mechanism provides some catalysis for the first step of the reaction, but the overall activation energy is still too high (ca. 35 ± 5 kcal) to account for the observed rate. One might wonder whether this finding is consistent with recent site directed mutagenesis experiments[29] which demonstrated that the Glu357Ala mutation reduces the rate constant by three orders of magnitude. Unfortunately, as in the case of most mutation studies, it is difficult to reach unique mechanistic conclusions based only on the observed effect of the mutations (see for example ref

(21) Hale, S. P.; Poole, L. B.; Gerlt, J. A. *Biochemistry* **1993**, *32*, 1479.
(22) Judice, J. K.; Gamble, T. R.; Murphy, E. C.; de Vos, A. M.; Schultz, P. G. *Science* **1993**, *261*, 1578.
(23) Davies, H. F.; Hostomska, Z.; Hostomsky, Z.; Jordan, S. R.; Matthews, D. A. *Science* **1994**, *252*, 88.
(24) Yang, W.; Hendrikson, W.; Crouch, R. J.; Satow, Y. *Science* **1990**, *249*, 1398.
(25) Kim, E. E.; Wyckoff, H. W. *J. Mol. Biol.* **1991**, *218*, 449.
(26) Pelletier, H.; Sawaya, M. R.; Kumar, A.; Wilson, S. H.; Kraut, J. *Science* **1994**, *264*, 1891.
(27) Davies, J. F.; Almassy, R. J.; Hostomska, Z.; Hostomsky, Z. *Cell* **1994**, *76*, 1.
(28) The coordinates of the different intermediates and transition state are available from the authors upon request.
(29) Derbyshire, V.; Grindley, N. D. F.; Joyce, C. M. *EMBO J.* **1991**, *10*, 17.

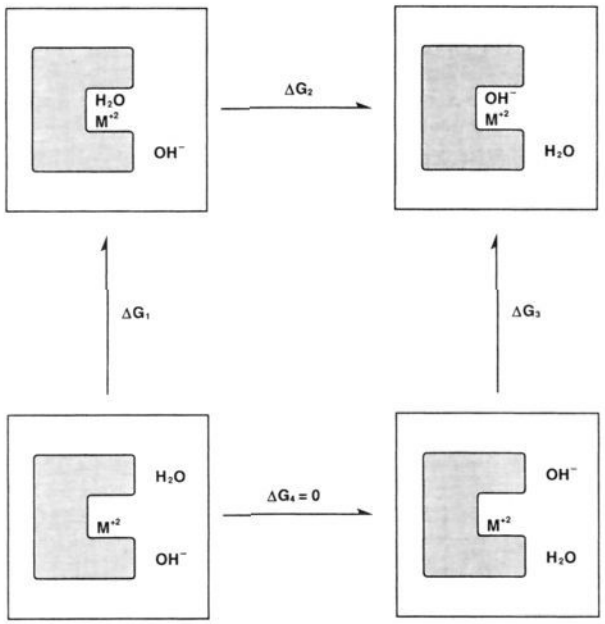

Figure 4. The thermodynamic cycle used to calculate the pK_a of the metal bound water.

30). As in previous cases[30] one might try to use the EVB method to elucidate the origin of the observed effect of the mutation, but such a study is not justified for a mechanism whose calculated activation barrier is entirely inconsistent with the observed rate constant.

Considering the failure of the general base mechanisms to account for the observed activation barrier of the exo site, we examined the external OH^- mechanism depicted in the lower diagram of Figure 1. The evaluation of the energetics of this mechanism is based on the use of the thermodynamic cycle of Figure 4 that describes the process of replacing an active site water by an OH^- ion from the bulk solution. The corresponding free energy change is given in terms of the ΔG_i's of Figure 4

$$\begin{aligned}\Delta G_2 &= \Delta G_1 + \Delta G_4 + \Delta G_3 = \Delta G_1 + \Delta G_3 \qquad (6)\\ &= \Delta G[(H_2O)^p \rightarrow (H_2O)^w] + \Delta G[(OH^-)^w \rightarrow (OH^-)^p]\\ &= \Delta G[(H_2O + OH^-)^p \rightarrow (OH^- + H_2O)^w]\\ &= \Delta G[(H_2O \rightarrow OH^-)^p] - \\ &\qquad \Delta G[(H_2O \rightarrow OH^-)^w] = \Delta\Delta G\end{aligned}$$

where w and p designate water and protein sites, respectively. As pointed out in our early studies[15,16,31] and in related studies (e.g., refs 32−35), one can determine the free energy associated with the process $XH \rightarrow X^-$ (e.g., $XH = H_2O$) inside a protein by "mutating" XH to X^- in the protein active site and in water. The free energies of these "mutation" processes are summarized in Table 1 and give (in kcal/mol):

Table 1. Energetics of Creating OH^- in the Protein Active Site[a]

environment	$\Delta G(H_2O \rightarrow OH^-)^{site}$	$\Delta\Delta G$	pK_a^{site}
protein	-115.1 ± 3	-6.1 ± 3	11.2 ± 2
water	-109.0 ± 2	0	15.7
Mg^{2+} in water	-114.9 ± 3	-5.9 ± 3	11.4 ± 2

[a] The calculations were performed with the program ENZYMIX [11] using the SCAAS boundary conditions and the LRF treatment of long range interactions. The calculations involve an average over four runs with different starting configurations generated during the equilibration process. Site designates the indicated environment. Energies are given in kcal/mol, and $\Delta\Delta G$ designates $\Delta G(H_2O \rightarrow OH^-)^{site} - \Delta G(H_2O \rightarrow OH^-)^w$.

$$\Delta\Delta G = \Delta G(H_2O \rightarrow OH^-)^p - \Delta G(H_2O \rightarrow OH^-)^w = -115.1 - (-109.0) \pm 3 = -6.1 \pm 3 \quad (7)$$

This free energy difference can be expressed in terms of the relevant pK_a changes (see refs 15 and 16) as

$$pK_a^p = pK_a^w + \Delta\Delta G/2.3RT = 15.7 - (6.1 \pm 3)/1.38 = 11 \pm 2 \quad (8)$$

The calculated value of 11 is quite close to the value of $pK_a \simeq 10$ that can be deduced from the observed titration curve of ref 29.

Using eq 3 and the calculated pK_a of the active site water (pK_a^p), an evaluation of the energetics of the first step of the specific base catalysis can be made at any given pH. At pH = 7.5, the formation of the OH^- nucleophile in the active site costs about 5 kcal/mol (see Figure 5), and other pH ranges will be considered in the concluding remark section. This is drastically less than the free energy needed for the equivalent step in the general base mechanisms considered in Figure 3a,b. Note that the same free energy is required for either a proton transfer from an active site water to an OH^- ion in solution or replacement of the water by this ion (a proton transfer to OH^- involves penetration of this ion to the active site and basically making OH^- and water from the original water molecule and OH^-). We did not evaluate the activation energy for a penetration of an OH^- ion to the active site, but this energy is likely to be significantly smaller than the overall barrier of the reaction (~22 kcal/mol) and would not change the $(\Delta G_p)_{I \rightarrow II}$ of Figure 5.

One might wonder why the metal is not as effective in the general base mechanism of Figure 3b, as it is in the external OH^- mechanism. The answer is that although the metal helps in reducing the pK_a of the donor water, it also reduces the pK_a of the acceptor, A^-. Furthermore, at relatively high pH the −pH term of eq 3 is more negative than the $-pK_a(AH)$ term of eq 2 for typical biological acids. This point will also be considered in the last section.

After evaluating the energetics of the formation of the OH^- ion (about 5 kcal/mol) we performed FEP/umbrella sampling calculations for the second step of the external OH^- catalytic mechanism (Figure 4) and found the activation energy for the step to be around 20 kcal/mol. Thus, the calculated barrier for this step was found to be around 25 kcal/mol, in reasonably good agreement with the observed activation barrier (22 kcal/mol).

If we neglect the relatively small effect of the ionizable protein residues (as printed out in ref 16, this contribution corresponds typically to an effective dielectric constant of about 40) we can obtain the pH profile by using eq. (3) for $(\Delta G_p)_{I \rightarrow II}$. In this way the rate constant relative to the corresponding value at pH = 7.5 will change exponentially with a dependence of $\exp\{RT(pH-7.5)\}$ on the solution pH.

(30) Warshel, A.; Naray-Szabo, G.; Sussman, F.; Hwang, J.-K. *Biochemistry* **1989**, *28*, 3629.
(31) Warshel, A.; Sussman, F.; King, G. *Biochemistry* **1986**, *25*, 8368.
(32) Merz, K. M. *J. Am. Chem. Soc.* **1991**, *113*, 3572.
(33) Yang, A.-S.; Gunner, M. R.; Sampogna, R.; Sharp, K.; Honig, B. *Proteins* **1993**, *15*, 252.
(34) Bashfold, D.; Karplus, M. *Biochemistry* **1990**, *29*, 10219.
(35) Del Buono, G. S.; Figueirdo, F. E.; Levy, R. M. *Proteins* **1994**, *20*, 85.

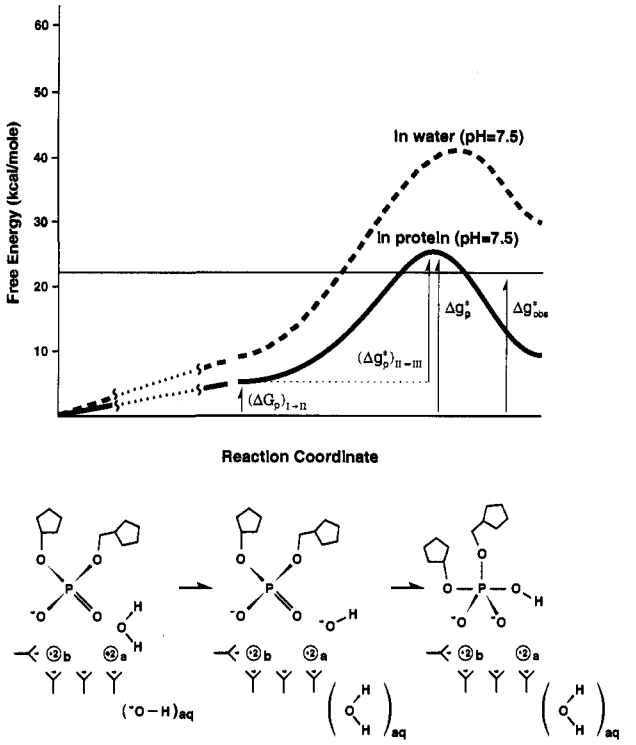

Figure 5. The calculated free energy profile for an external OH^- catalysis mechanism in the exo site of the Klenow fragment. The calculated activation barrier in this mechanism is of the same order of magnitude as the corresponding observed barrier. The first step of the reaction (I → II) is presented in a schematic way (i.e., straight lines) since we only evaluate the ΔG for this step (eq 3) and not the corresponding barrier ($\Delta g^{\ddagger}_{I\to II}$). Note, however, that as long as this barrier is small, the overall barrier for the reaction is given by $\Delta g^{\ddagger} = \Delta G_{I\to II} + \Delta g^{\ddagger}_{II\to III}$.

In analyzing the role of the active site in this mechanism, we note (see Figure 5) that the enzyme provides a large catalytic contribution to the nucleophilic attack step. Apparently, as indicated schematically in Figure 1(b), the second $Mg^{2+}_{(b)}$ ion exerts its electrostatic effect in this step by stabilizing the charge which comes to reside on the phosphate center. This effect will be examined in the next section.

3.3. The Catalytic role of the Second Metal in the Exo Site; the Importance of Electrostatic Effects and the Problems with a Strain Mechanism. As indicated in Figure 5, the enzyme exerts a large catalytic effect in the nucleophilic attack step. A major part of this effect appears to be associated with the electrostatic interaction between $Mg^{2+}_{(b)}$ and the negative charge that comes to reside on the phosphate center.

The catalytic importance of the metal ion $Mg^{2+}_{(b)}$ has been recognized before e.g., refs 2, 18, and 20. However, the details of the proposed catalytic effects differ from the electrostatic stabilization found in this work. It has been proposed[18,20] that the metal ion B acts both as a Lewis acid in facilitating the leaving of the 3′ oxyanion (step III → IV of Figure 1) and by inducing strain on the O−P−O bond angle. The first proposed effect is related to electrostatic stabilization of state IV and is quite reasonable if step III → IV is the rate-limiting step. However, for cases where the formation of the pentacoordinated intermediate is the rate-limiting step, it is necessary to focus on the role of metal B in stabilizing state III. The strain hypothesis can be analyzed by examining the effect of the metal on the bending potential of the O−P−O angle. It should be noted that by strain one means a nonelectrostatic steric effect, and if the interaction with the metal stabilizes the 90° configuration (which corresponds to the pentacoordinated state) more

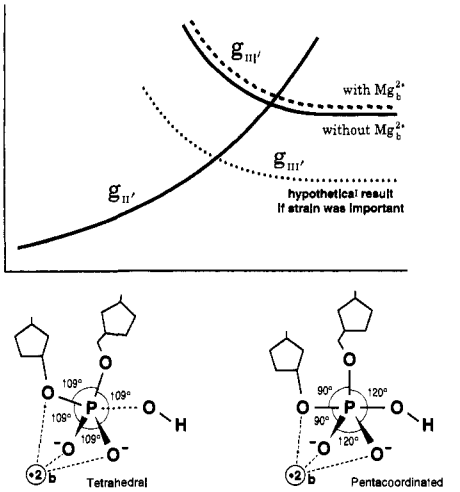

Figure 6. Examining the possible role of Mg^{2+}_{b} in a strain mechanism. The calculations were done by using special potential functions, $\epsilon_{II'}$ and $\epsilon_{III'}$, with the same bending parameters as that of state II and III, respectively, and with the charge distribution of state III for both states but with no residual charge on the attacking OH group. Using these functions and eq 9 allow us to force the O−P−O angle to change and to examine the effect of the metal on this bending. As seen from the figure, the steric interaction between the metal to the phosphate oxygen does not stabilize state III′. This finding indicated that Mg^{2+}_{b} acts by electrostatic rather than by steric stabilization.

than the tetrahedral geometry of the reactant state, then strain can make a significant contribution to the catalytic process. In order to evaluate this effect we have to force the O−P−O angle to change from 109° to 90° with and without the metal and to determine the corresponding effect of the metal. The simplest way to force the angle to change is to represent the system by two fictitious resonance structures (II′ and III′) with the charge distribution of state III for the phosphate atoms in both states and with *no* residual charges on the attacking OH^- group in both states and then to use a mapping potential of the form

$$\epsilon_m = (1 - \lambda_m)\epsilon_{II'} + \lambda_m\epsilon_{III'} \qquad (9)$$

where $\epsilon_{II'}$ and $\epsilon_{III'}$ are identical to the potential surfaces of state II and state III but with the charge distribution of state II′ and state III′. Now the mapping process (that involves a change of λ from zero to one) moves the systems from configurations that stabilize $\epsilon_{II'}$ to those that stabilize $\epsilon_{III'}$ and changes the O−P−O angle from 109° to 90° without transferring any negative charge to the phosphate. This mapping process allows us to determine the strain contribution without including the electrostatic effect that is associated with the transition from the real state II to the real state III. Thus, the results of FEP calculations that change λ_m from 0 to 1 with and without $Mg^{2+}_{(b)}$ (Figure 6) should give the direct contribution of $Mg^{2+}_{(b)}$ to strain. As seen from Figure 6, no strain contribution was found. In fact, adding the "strain" effect of the $Mg^{2+}_{(b)}$ ion slightly destabilizes the transition state.

As far as the strain mechanism is concerned, it is hard to see how the small change of the $Mg^{2+}\cdots O^-$ distance that accompanies the formation of the pentacoordinated state could lead to a significant interaction change and to a strain mechanism. It appears that the Mg^{2+} ion just retains the same interaction with the phosphate oxygens as they adjust their positions during the transition from state II′ to state III′. Another way to think about this model is to realize that two oxygen atoms can be considered as ligands to the Mg^{2+} ion and that the bending force constant that involves ligands around transition metals is rather small. Finally, the fact that the inclusion of electrostatic

effects accounts quantitatively for the observed rate acceleration makes the strain effect even more unlikely.

4. Concluding Remarks

The present work explores the special features of the two-metals mechanisms in the exo site of DNA pol I by using a computer simulation model for correlating the structure and energetics of this system. It is found that the first metal ion ($Mg^{2+}_{(a)}$) provides electrostatic stabilization to the OH^- nucleophile and that the second metal ion ($Mg^{2+}_{(b)}$) stabilizes the negative charge that comes to reside on the phosphate during the nucleophilic attack step. While this overall picture seems to be a rather simple manifestation of the qualitative considerations of Figure 1b, the details of and the actual way by which the catalysis is accomplished are not so obvious and require a quantitative structure-function analysis of the type presented in this work. Apparently the fact that the Mg^{2+} ions are bound to the protein site by negatively charged acidic residues (e.g., Asp and Glu) makes it hard for the cations to stabilize the OH^- charge. The negative field of these groups shields the positive field of the $Mg^{2+}_{(a)}$ ion and reduces its ability to lower the pK_a of the catalytic water. Using Glu 357 as a base for the proton of the catalytic water could have provided an effective way to overcome this problem (the Glu charge would have been neutralized during the proton transfer process). However, the present calculations have indicated that Glu 357 does not provide a reasonable base. Apparently, it is advantageous to use an OH^- ion from solution as a base rather than to use a mechanism that involves a proton transfer to another water or to a protein residue. Note, that substitution of the catalytic water by OH^- from the bulk solvent and proton transfer from this water to an OH^- (that was brought to the active site from bulk solvent) involve similar energy considerations and are determined by the pK_a of the water bound to the ($Mg^{2+}_{(a)}$) ion. Also note that the external OH^- mechanism is supported by the experimentally observed pH profile of the exo reaction[29] and that a similar pH profile is observed in the action of ribozymes.[6]

In considering the energetics of the OH^- formation it might be useful to point out that our study does not support any special entropic factors associated with the orientation of the nucleophile (see also chapter 9 of ref 10). In particular we do not find that Glu357 orients the OH^- ion as proposed before in refs 18 and 20. Basically the negative charge of Glu357 repels the negatively charge of the OH^- ion and trys to keep it as far as possible. This repulsion persists even when the hydrogen of the OH^- ion points toward the acid and provides the opportunity of forming an (O^- H—O^-) interaction. Perhaps the misunderstanding is associated with the confusion of the stable hydrogen bond of the (O^- H—O) system with the present (O^- H—O^-) system. Similarly, we do not find support for the idea[2] that the role of the Mg^{2+} ion is to orient the attacking OH^- and supposedly help entropically by restricting the angular orientational entropy of this nucleophile. Our calculations do not produce a drastic angular constraint on the orientation of the OH^- ion in state II. More importantly, if the OH fragment is fixed at the transition state, there is no way to use its orientational entropy in increasing k_{cat}/K_M (as explained in chapter 9 of ref 10, this quantity reflects the difference between the free energy of the transition state in the protein and the reactants in water and it is not affected by intermediate states). This again emphasizes the overwhelming importance of electrostatic considerations where the energetics of state II is mainly determined by the electrostatic interaction of the OH^- ion with its surrounding. It is also instructive to comment here on the proposal[2,17] that the $Mg^{2+}_{(a)}OH^-$ system serves as a base for forming OH^- ion from water. The energy requirement of this process is equivalent to the energy of moving OH^- from the first solvation shell to the second one which is simply a part of our step II → III. But this mechanism neglects the crucial energy of forming the OH^-, thus implying that the OH^- is readily available in the active site at any pH is in disagreement with the observed pH profile. Furthermore, the calculations of ref 17 are gas phase calculations and have no quantitative relevance to the corresponding process in the enzyme or solution.

While the role of the metal ion in stabilizing the nucleophile is intuitively clear there seems to be confusion about what is meant by this effect. For example, some workers associate this effect with the particular type of the metal ion (e.g., Mg^{2+} or Ca^{2+}) rather than with the combined effect of the metal and its specific site. Evaluating the interaction of the OH^- ion with the cation does not mean anything without considering the "dielectric" effect associated with the given site (i.e., the ligands and the rest of the protein). Each metal ion can have different effects depending on its particular site so that the electrostatic effect of the cation and the protein rather than the metal ion alone is what counts.

Some readers might wonder about the logic of our proposed mechanism since any factor that stabilizes the OH^- ion and makes it less basic would also make it less nucleophilic. This issue which might seem rather confusing when addressed by most available concepts about chemical reactivity is uniquely analyzed by the EVB approach. That is, stabilizing state II in Figure 1 (which corresponds to making a less basic OH^- ion) does not change the energy at the minimum of state III and the corresponding $\Delta G^0_{I\to III}$. However, our transition state is not at the minimum of state III but rather at the intersection of the parabolae of state II and state III. The downward shift of the parabola of state II pushes the intersection of these two parabolae down and to the right, toward the minimum of state III. In other words, stabilizing the OH^- ion makes the transition state for II → III more product-like and reduces its energy. The present argument is also related to the above discussion of the problems associated with entropic effects as catalytic factors. That is, fixing the OH^- ion in state II will only destabilize rather than stabilize state II.

The present study has not examined in detail the effect of pH on phosphodiester hydrolysis (we only gave a very rough estimate, neglecting the effect of the change in the ionization state of the protein residues). A systematic evaluation of the pH dependence of the rates, which is entirely feasible within the EVB formulation, should consider the activation energy for different ionization forms of the reacting system (e.g., the mechanism when Glu 357 is in its protonated form) and when other protein residues are in different ionization forms (e.g., ionized and neutral Tyr497). The rate constants for all the relevant ionization states should then be weighted according to the relevant concentration of each state in the given pH. It is expected, for example, that a general base catalysis that involves Glu357 will be important at low pH when the concentration of OH^- is low.

One may wonder why the external OH^- mechanism is more effective than the general base mechanism of Figure 3b. The reasons are quite instructive. Both processes have the same proton donor, but in the first case we have an OH^- as a base and in the second case a glutamate ion. The OH^- ion is a much stronger base (the pK_a difference for a case when the two bases are in solution is ~10 pK_a units), but at pH ≃ 7 the concentration of the OH^- is so small that it corresponds to a loss of 7 pK_a units in $\Delta g_{I\to II}$, relative to the general base mechanism where

the base has a unit concentration. However, the effective pK_a of Glu357 in the protein is significantly lower than in solution due to the effect of $Mg^{2+}_{(a)}$. Another way to see this point is to realize that a proton transfer from the bound H_2O to an OH^- ion in water leads to a significant electrostatic stabilization where H_2O Mg^{2+} becomes OH^-Mg^{2+}, while in the general base mechanism (H_2O $Mg^{2+}COO^-$) becomes ($OH^-Mg^{2+}COOH$) with much less increase in electrostatic stabilization.

Comparing the catalytic effect of the exo site of DNA pol I to that of Snase, we argue that the double and single metal mechanisms represent two alternative ways to catalyze the same reaction. The first mechanism (Figure 1a) uses a single metal to both catalyze the formation of the nucleophile (by stabilizing its negative charge) and to help the migration of the nucleophile charge to the phosphate. In this case, the metal should be able to move with the nucleophile toward the phosphate. On the other hand, in the two metal mechanism (Figure 1b), each metal has a different role. The first one stabilizes the nucleophile charge, and the second stabilizes the negative charge that is transferred to the phosphate in the second step of the reaction.

At present it is hard to determine uniquely what the requirements for one-metal and two-metal active sites are, and we can only evaluate the catalytic effect using the X-ray structure of the given site. However, some interesting points can be deduced from the present study; sites that hold the metal in a very rigid environment may reduce its ability to participate in a single-metal catalytic mechanism since the metal cannot move effectively toward the phosphate center during the nucleophilic attack step. However, such sites would be quite useful in a two-metal mechanism. Another possible advantage of a two-metal site can be provided in cases where the rate limiting step involves the dissociation of the leaving group (step III → IV), or when this step occurs in concert with step II → III. This effect may be examined by substrates with better leaving groups where the energy of state IV is lowered.

It might be useful to comment about the relationship between the present paper and experimental "reality". Apparently there is a tendency in some circles to confuse interpretation of experiments with experimental facts and at the same time to view computer simulations as exercises that simply explore the behavior of different assumed models. Such a view overlooks the emerging role of computer modeling in providing structure-function correlations. For example, the present approach estimates the activation energy for different assumed mechanisms by calibrating the EVB parameter using information about elementary steps of the reaction in water and then using them without any readjustment for modeling the assumed mechanism in the enzyme active site. The calculated activation barriers for each of the assumed mechanisms are then compared to the single observed activation barrier (which corresponds to what is in principle an unknown mechanism). If the calculated barrier is very different than the observed barrier, then this assumed mechanism is eliminated. One may still argue that the calculations are not sufficiently quantitative to allow such an elimination process. However, the same type of calculations have reproduced experimentally observed electrostatic energies with a reasonable accuracy,[12] and as argued and demonstrated before (e.g., ref 30), the error range in our calculations is determined by their ability to evaluate electrostatic energies in proteins. Thus we like to view our approach as a valid attempt and at present perhaps the only way of using structural-based energy considerations in mechanistic analysis (we are not aware of a direct experimental approach that can do so).

Of course, one may try to use mutation experiments to validate our proposed mechanism. But unfortunately mutation experiments only provide circumstantial evidence since they can be interpreted in different ways and essentially the only way to make the interpretation more unique is to use computer simulation approaches (e.g., see ref 30). In general calculating absolute activation free energies is much more effective in eliminating different mechanisms than in reproducing the effect of mutations which could be consistent with several mechanisms. For example, the three orders of magnitude reduction in rate constants as a result of the Glu357Ala mutation could be interpreted as a loss of the general base mechanism of Figure 3b or as reduction of the shielding of the Mg^{2+}_a that makes it more effective in stabilizing the OH^- ion in the external OH^- mechanism of Figure 5. Nevertheless, the effect of the Tyr497Phe mutation[29] is consistent with the external OH^- mechanism. That is, the ionization of Tyr497 at a high pH creates a negatively charged ionized Tyr residue that should destabilize the OH^- nucleophile and increase its pK_a and therefore increase $\Delta g^{\ddagger}_{I\rightarrow II}$ relative to its value in the absence of Tyr497. This is consistent with the fact that the native protein mutation slows down the reaction at high pH, while this effect is not observed in the Tyr497Phe mutant. However, a more serious analysis of the pH effect on the native and mutant enzyme must involve actual calculations, and such a study will be reported elsewhere.

Acknowledgment. We are grateful to Dr. Lorena Beese for providing the protein coordinates. This work was supported by NIH Grant GM 24492 and DOE Grant DE-FG03-94ER61945 to A.W. and NIH Grant GM 21422 to M.F.G.

JA951726B