BIOCHEMISTRY
including biophysical chemistry & molecular biology
Article

Biochemistry **2009,** *48,* 5963–5971 5963
DOI: 10.1021/bi900140c

A Computational Study of the Hydrolysis of dGTP Analogues with Halomethylene-Modified Leaving Groups in Solution: Implications for the Mechanism of DNA Polymerases[†]

Shina C. L. Kamerlin,* Charles E. McKenna, Myron F. Goodman, and A. Warshel*

Department of Chemistry, University of Southern California, 3620 McClintock Avenue, Los Angeles, California 90089

Received January 28, 2009; Revised Manuscript Received April 23, 2009

ABSTRACT: DNA polymerases make up a family of enzymes responsible for regulating DNA replication and repair, which in turn maintains the integrity of the genome. However, despite intensive kinetic, crystallographic, and computational studies, elucidation of the detailed enzymatic mechanism still presents a significant challenge. We recently developed an alternative strategy for exploring the fidelity and mechanism of DNA polymerases, by probing leaving group effects on nucleotidyl transfer using a series of dGTP bisphosphonate analogues in which the β,γ-bridging oxygen was replaced by a series of substituted methylene groups (X = CYZ, where Y and Z = H, halogen, or another substituent). Pre-steady state kinetic measurements of DNA polymerase-catalyzed incorporation of correctly base paired (R) and mispaired (W) analogues demonstrated a strong linear free energy relationship (LFER) between the polymerase rate constant (k_{pol}) and the highest pK_a of the free bisphosphonic acid corresponding to the leaving group. However, unexpectedly, the data segregated into two distinctly different linear correlations depending on the nature of the substituent. The discrepancy between the two lines was considerably greater when the dGTP analogue formed an incorrect (G·T) rather than a correct (G·C) base pair, although the reason for this phenomenon remains unexplained. Here, we have evaluated the complete free energy surfaces for bisphosphonate hydrolysis in aqueous solution and evaluated the corresponding LFER. Our study, which employs several alternative solvation models, finds a split of the calculated LFER for the mono- and dihalogen compounds into two parallel lines, reflecting their behavior in the polymerase-catalyzed condensation reaction. We suggest that the division into two linear subsets may be a generalized solvation phenomenon involving the overall electrostatic interaction between the substrates and their surroundings and would also be observed in polar solvents in the absence of the enzyme, if the reaction in solvent is in fact identical to that of the enzyme. However, the amplified differences between the LFER lines for the incorporation of matched and mismatched deoxynucleotides probably reflects the differences in the electrostatic interaction between the TS charges in the polymerase active site. An understanding of the mechanism of this reaction in solution could thereby provide a steppingstone for understanding the factors governing the fidelity of DNA polymerases.

Since the discovery of DNA polymerase in *Escherichia coli* more than 50 years ago (*1*, *2*), this enzyme, including both its high- and low-fidelity forms (*3*, *4*), has been under continuing and intense scrutiny in efforts to improve our understanding of its structures and functions. As the field has matured, explorations have moved beyond individual isolated polymerases to encompass interactions with accessory subunits, including protein complexes involved in DNA replication and repair (e.g., refs (*5*−*7*)). Nevertheless, a comprehensive characterization of the enzymatic mechanism has not yet been achieved and remains a challenge to biochemists.

A detailed understanding of the mechanisms governing DNA synthesis is essential for establishing not only how correctly paired deoxyribonucleotides are incorporated into DNA but also, equally importantly, how the polymerase discriminates against incorporating non-Watson−Crick (non-W−C) base pairs at the molecular level. X-ray crystallographic studies in which polymerase−primer-template−dNTP ternary complexes are captured at high resolution (*8*−*11*) show how correctly and incorrectly paired dNTPs are aligned within the polymerase active site. These geometric "snapshots" are unquestionably important in deducing how dNTP−p/t DNA orientations occurring in the polymerase ground state may either facilitate efficient phosphodiester bond formation for W−C base pairs or impede the incorporation of non−W−C base pairs (*11*), but the nature of the transition states occurring on the catalytic reaction pathway cannot be deduced from structural studies without additional theoretical "interpolations" (*12*−*15*). Such theoretical studies can

[†]This work was supported by NCI Grant 1 U19 CA105010-01 and NSF Grant MCB-0342276. All computational work was supported by the University of Southern California High Performance Computing and Communication Center.

*To whom correspondence should be addressed. E-mail: warshel@usc.edu or l.kamerlin@gmx.at. Phone: (213) 740-4114. Fax: (213) 740-2701.

 Published on Web 04/24/2009

greatly benefit from experimental information about the mechanism of the incorporation reaction. However, extracting such information in a unique way is challenging.

Recently, we developed a strategy for exploring the mechanism and fidelity from the perspective of the polymerase transition state. The approach is based on probing the nucleotidyl transfer mechanism for leaving group effects (downstream from the transition state) using a series of dNTP bisphosphonate analogues in which the β,γ-bridging oxygen of pyrophosphate was replaced by a series of substituted methylene groups (*16*, *17*) [X = CYZ, where Y and Z = H, halogen, or another substituent (Figure 1)]. For example, using the Y and Z = H, F, Cl, and Br set, the electrostatic properties of the bisphosphonate moiety can be tuned to provide a wide dynamic range in the basicity of the new oxyanion formed by $P_\alpha-O$ bond cleavage, which should reflect the pK_a values for the corresponding free bisphosphonic acids (*11*, *18*). We have proposed the construction of a linear free energy relationship (LFER) correlating the logarithm of the catalytic rate constant (k_{pol}) to the highest leaving group pK_a for such a series of synthetic dNTP analogues as a means of directly probing the energetics of the chemical steps of nucleotidyl transfer that relate the height of leaving group elimination barriers to other steps in the reaction pathway (*16*, *17*).

To implement this new approach, we have examined the LFER for human DNA polymerase β using a series of dGTP analogues (Figure 1) constructed from bisphosphonates having pK_a[4] values spanning the range of 7.8−10.5 (*16*, *17*). The G analogues were either correctly incorporated opposite template C or misincorporated opposite T. The resulting plots of log k_{pol} versus log pK_a are reproduced in Figure 2.

The data exhibited strong linear correlations between log k_{pol} and leaving group bisphosphonic acid pK_a[4] for correct (G·C, denoted R in the text) (Figure 2A) and incorrect (G·T, denoted W in the text) base pairs (Figure 2B). This result is expected in a LFER model in which a single activation barrier is rate-limiting for the chemical step for correct and incorrect substrates. However, the data unexpectedly segregated into two different linear correlations, one for a group comprised of the natural substrate with the unsubstituted methylene and monohalomethylene analogues and the other for the dihalomethylene analogues, with the difference being significantly greater for the mispairs; i.e., the LFER plots were fidelity-dependent (Figure 2). This remarkable difference between the C·G and T·G incorporation rates for the monohalomethylene versus the dihalomethylene analogue groups, as well as the diminished catalytic activity exhibited with the dihalogen-bridging analogues, is unexplained (*16*, *17*). It should be noted that monohalo nucleotides are mixtures of two diastereomeric forms, differing in configuration at the P−C−P carbon. However, were one form less effective as a substrate than the other, the observed turnover rates would be relatively decreased, not increased, even if saturating substrate concentrations were not used. We have considered several possibilities to account for the diminished activity of the dihalomethylene compounds in the enzymatic system, e.g., leaving group elimination might be hindered by intramolecular repulsion between the negative charge developing on the $\beta\alpha$ oxygen and the pro-*S* halogen (*17*), or alternatively, the halogens may engage in sterically debilitating steric/electrostatic interactions with nearby amino acid side chains in the pol β active site (*17*), but neither of these rationales is fully convincing. Thus, it is important to consider the possibility that the disparity in rates of release of the mono- and dihalobisphosphonate leaving groups reflects, at least to some extent, a trend that already exists in the uncatalyzed reference reaction in solution. Here, we would like to emphasize that our reference reaction corresponds to the same mechanism that occurs in the enzyme, which can be different from the predominant uncatalyzed reaction that occurs in solution.

To address this issue, it is necessary to explore the mechanism of the uncatalyzed solution reaction to understand the origin of the LFER in the enzyme. Such a study is also important for understanding the catalytic power of the enzyme, which reflects the difference between the reaction in the enzyme and in solution. Furthermore, characterization of the solution reaction is crucial for the calibration and validation of theoretical models that are being used for studying enzymatic reactions, which was demonstrated in previous studies (*19*). However, studying phosphate ester hydrolysis in solution is a difficult problem, as even the solution reaction can proceed through a variety of different reaction mechanisms (see, e.g., refs (*20*−*22*)). In broad terms,

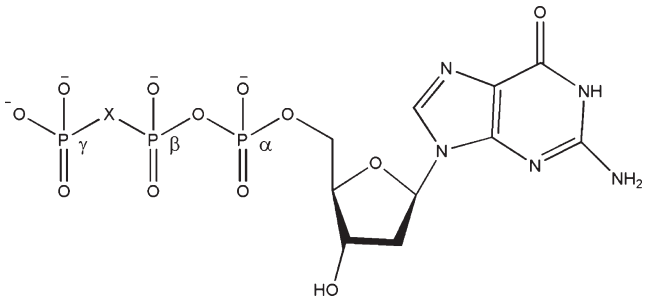

FIGURE 1: Structure of dGTP and β,γ-substituted analogues. X = CF_2, CFCl, CCl_2, O, CHF, CBr_2, CHCl, CHBr, and CH_2 (*16*, *17*).

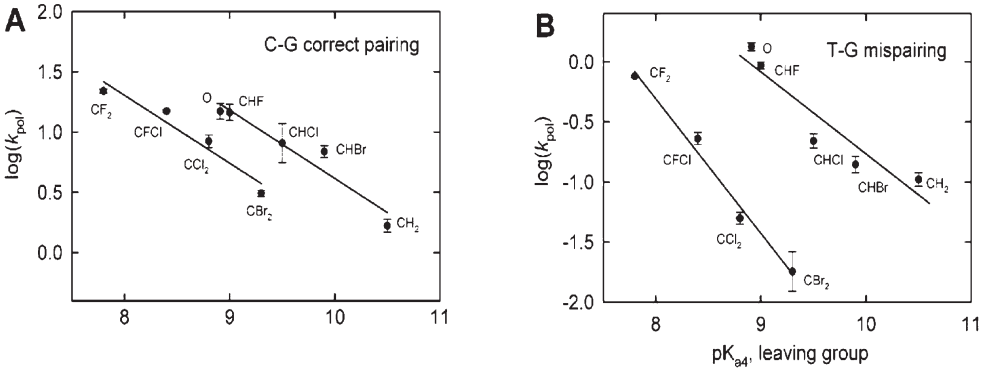

FIGURE 2: LFER for dGTP and analogues, incorporated opposite (A) the correct template base, C, and (B) the mispairing template base, T. This figure was reproduced from ref *17*. Copyright 2008 American Chemical Society.

phosphate hydrolysis can proceed via one of two mechanistic extremes, that is, pathways that can be associative or dissociative in nature, depending on the degrees of bond breaking with respect to the leaving group and bond formation with respect to the nucleophile, respectively, and despite decades of research, the precise nature of both the solution and enzyme-catalyzed reactions remains controversial (*23*−*33*).

Computational studies that map the full energy surface (*19*, *21*, *22*, *34*, *35*) for phosphate hydrolysis rather than examining isolated transition states have demonstrated that it is possible to accurately reproduce the LFER for phosphate hydrolysis in solution. Such studies examine the free energy surface by means of More O'Ferrall−Jencks (MFJ) plots (*36*, *37*), a sample of which is shown in Figure 3. MFJ plots are projections of the full free energy surface defined in terms of two reaction coordinates. At each point on the plot, these coordinates are frozen while all other degrees of freedom are allowed to be optimized without constraints. The full MFJ plot can thus be obtained by careful reaction coordinate pushing, and solvation can be simulated using a continuum model.

There are a number of advantages to using this approach. Firstly, it allows us to identify and directly compare different viable mechanisms in the presence of the same nucleophile and substrate. Secondly, the MFJ plot can be used to identify the approximate location of key stationary points, the precise geometry of which can then be obtained by unconstrained geometry optimizations. Finally, it should be noted that it can be quite difficult to study the relevant reference reaction in solution for an enzymatic reaction experimentally, as the enzyme can restrict the reaction in various ways so a specific mechanism can be followed, whereas in solution, there are multiple mechanistic possibilities. For instance, in the case of the system studied here, it is quite possible that experimentally, the nucleophile does not attack in the same position as in pol β, but rather that the terminal (γ) phosphate would react much faster, and as such, it would be possible to obtain only an upper limit for P−O bond cleavage between the α- and β-phosphates. However, it is essential to model precisely the same reaction as in the enzymatic system when studying the reference reaction, and this is an advantage of using a computational approach, where one can

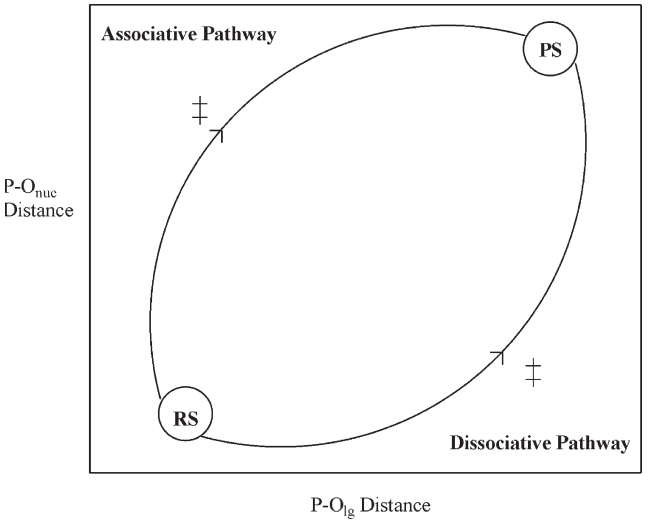

FIGURE 3: Sample More O'Ferrall−Jencks plot (*36*, *37*). RS and PS denote reactant and product states, respectively, and ≠ denotes a transition state, which can be either stepwise ($A_N + D_N$ or $D_N + A_N$) or concerted ($A_N + D_N$).

effectively place constraints on the reaction in such a way to accurately reproduce the correct reference state for the enzymatic reaction.

To address the issue of whether the disparity in the rates of release of the mono- and dihalobisphosphonate leaving groups reflects, at least to some extent, a trend that already exists in the uncatalyzed reference reaction in solution, and to gain more information about this reaction in general, we have carried out a computational analysis of the hydrolysis of these compounds in solution, using a variety of solvation models. Our study focused on the change in the rate of hydrolysis of the series described above as a function of the observed pK_a for the corresponding bisphosphonic acid, which he have examined by generating the full free energy surfaces for each system studied. The calculations indicated that the trend observed in the LFER of the reaction in pol β is already expected when the same reaction occurs in solution.

THEORETICAL BACKGROUND

The main focus of this work is to study the effect of halogen substitution at the β,γ-bridging position of this series of dGTP analogues on the basic chemistry of these phosphate derivatives in solution rather than the reproduction of absolute energetics; therefore, we have used a simplified series of model compounds, in which the nucleoside is replaced by a methyl group, as shown in Figure 4. The general methodologies have been described in our earlier studies (*19*, *21*, *22*, *34*, *35*).

As we introduce our model system, we believe it to be important to remind the reader of the issue of the choice of the correct reference state for the enzymatic reaction, as even though this issue has been discussed in detail elsewhere (see, e.g., refs *38* and *39*), it still appears to cause confusion for many members of the community. Thus, we will briefly recapitulate the issue here. The definition of the word "catalysis" in a chemical context is to accelerate a chemical change by the addition of a catalyst (enzyme or otherwise). So, of course, a question that should immediately spring to mind is catalysis relative to what? Without an answer to this crucial question, statements such as "the enzyme binds the transition state stronger than the ground state" or "the catalytic groups are perfectly oriented" become meaningless, as they ignore the issue of how such differential binding is accomplished and what the actual catalytic groups are. As Wolfenden and coworkers have so elegantly demonstrated (*40*, *41*), many enzymes have evolved by optimizing k_{cat}/K_M, which can be very high relative to that of the uncatalyzed reaction of the substrate in solution (*41*) (k_{uncat}). However, such an observation is again meaningless without the definition of an adequate reference state, and as we have pointed out previously (*38*, *39*) and as described in, e.g., ref *41*, the most appropriate reference state is the uncatalyzed reaction in water. Now, in practice, the enzymatic reaction and the reaction in solution are not necessarily identical, since the solution reaction can proceed through a variety of different mechanisms, whereas the enzyme can restrict the reaction in several ways to follow a specific mechanism. Thus,

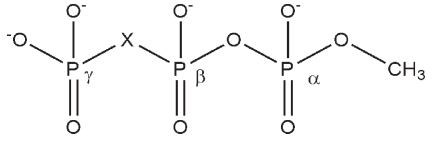

FIGURE 4: Model compounds. X = CF_2, CFCl, CCl_2, O, CHF, CBr_2, CHCl, CHBr, or CH_2.

it is helpful to consider many different mechanisms in solution, as well as the effect of having different environments. However, fortunately, a difference in mechanism here can be considered to be a "chemical effect" (such as having a general base instead of a water base), and such effects are well understood. The more important issue is to focus on the difference in the effect of the environment, that is, to compare the rate constant of a reaction that involves exactly the same mechanism, binding the same chemical groups, but being conducted in water instead of in the enzyme, which allows us to elucidate the catalytic enhancement of the reaction (i.e., the difference between the $\Delta G^{\ddagger}$ of the enzyme-catalyzed reaction and that of the same reaction in solution). This is the most proper reference reaction (see, e.g., ref *39*), which has also been understood by workers other than us (e.g., refs *42* and *43*). This is where computer simulations can be helpful, as it is possible to model the same reaction as in the enzyme even when this is not the preferred pathway in the absence of the enzyme, and by directly comparing the enzymatic reaction to the reference reaction in such a way, we are able to resolve controversies that would otherwise be hard to resolve experimentally (as was done in, e.g., ref *44* among many of our works). If we return to the compounds presented in Figure 4, of course, in principle the reaction could proceed via nucleophilic attack at any of the α-, *β*-, and *γ*-positions, and elucidating the actual position of attack in solution should be resolved experimentally. However, only nucleophilic attack at the α-position is biologically relevant in this case (as it is the correct reference state for the pol *β*-catalyzed reaction), and thus, this is the reaction that we have examined in this work. It must be clarified that we are focusing on the LFER of the enzymatic reaction in solution, and *not* on the LFER that will occur in the actual solution reaction, which needs to be verified by experimental studies. In other words, in trying to understand the origin of the LFER of the enzymatic reaction, we explored here whether this is due to the difference between the enzymatic environment and the solution environment. Of course, we recognize and anticipate the possibility that the enzyme changes the mechanism of the solution reaction, but this is not an issue in this work or a factor in what we define as "chemically filtered catalysis" (see ref *45*).

Separate energy surfaces were generated for each of the compounds in the series, both in the gas phase and in solution. In each case, the reaction surface was defined in terms of two reaction coordinates: the phosphorus oxygen distance to the leaving group [P−O_{lg} (Figure 3, *x*-axis)] and nucleophile [P−O_{nuc} (Figure 3, *y*-axis)]. At each point on the plot, only these two degrees of freedom were constrained, and all other degrees of freedom were allowed to freely optimize. P−O_{lg} distances were scanned in the range of 1.6−3.2 Å (in 0.2 Å increments), and P−O_{nuc} distances were scanned in the range of 1.65−3.0 Å (in 0.15 Å increments). It should be noted that, even though in principle the reaction could proceed via either an inline (nucleophile attacking from the opposite face as the departing leaving group) or a noninline (nucleophile attacking from the same face as the departing leaving group) mechanism (though both pathways have been suggested to proceed with fairly similar energetics), here the only constraints we have placed on the system are distance constraints on the phosphorus oxygen distances to the leaving group and nucleophilic oxygens, and thus, the nucleophile is left free to orient itself in such a way as to follow the lowest-energy path. Thus, the full free energy surface was obtained by precise reaction coordinate mapping, and the geometries and energies at each point on the plot were examined to ensure that the two-dimensional plot obtained is the true lowest-energy free energy surface.

All ab initio calculations were performed using the Gaussian software package (*45*), and Becke's three-level hybrid functionals, which combine Hartree−Fock exchange with density functional theory (DFT) exchange correlations (*46*). Initial gas-phase geometries were obtained using the 6-31 + G* basis set, followed by a single-point correction using the 6-311 + G** basis set to obtain the relevant energetics. Finally, solvation was simulated by applying a solvation correction to the gas-phase geometries, using the 6-311 + G** basis set and not only COSMO (*47*, *48*) but also the PCM(*49*−*53*) and Langevin Dipole (LD) (*54*) models for comparison. In the case of the COSMO and PCM models, the UFF model (which places spheres on all hydrogen atoms, thus treating them explicitly, using radii from the UFF force field) (*55*) was used rather than the standard UA0 model, to account for any possible proton transfer between the attacking water molecule and the phosphate. This approach has been demonstrated to accurately reproduce experimental activation barriers for phosphate hydrolysis (*19*, *22*, *34*, *35*).

RESULTS AND DISCUSSION

Figure 5 shows the free energy surface for the hydrolysis of the parent dGTP analogue (X = O) in solution, and the geometry of the transition state for this reaction is shown in Figure 6. As one can see in Figure 5, the reaction proceeds through a single concerted associative (A_ND_N) pathway, with P−O distances of 1.8 and 2.0 Å for the leaving group and nucleophile, respectively, in the transition state. The calculated activation barrier to this hydrolysis is 38.6 kcal/mol, as compared to an activation barrier

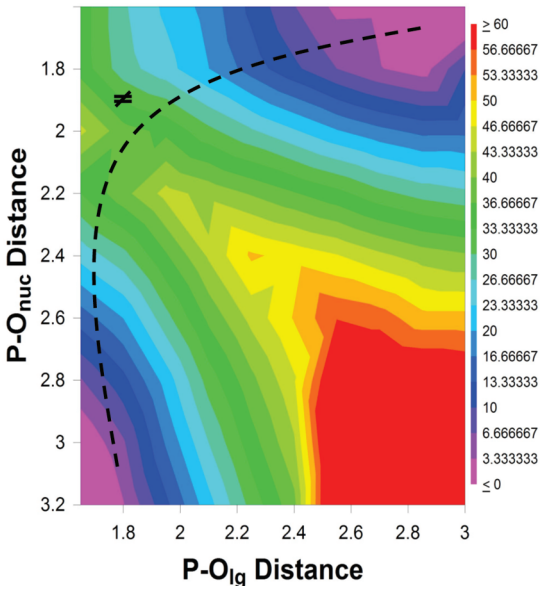

FIGURE 5: Free energy surface for the dGTP analogue where X = O. All distances are in angstroms, and ≠ denotes a transition state.

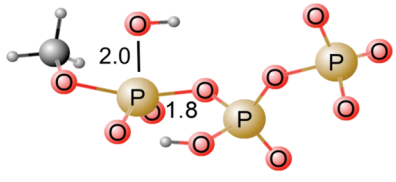

FIGURE 6: Concerted (A_ND_N) transition state for the parent dGTP analogue (X = O).

of 28.6 kcal/mol for the pol β-catalyzed hydrolysis of dGTP opposite a correct (R) template base (*17*). It should be noted that we have estimated free energies and rate constants on the basis of a temperature of 250 °C. This is due to the fact that such a high temperature that would be required for an experimental study of such a reaction in solution (see, for instance, refs *35* and *56* for a similar reaction), and we believe that that a comparison to an experimental value at the original temperature necessary for obtaining the experimental data is more reliable than comparison to the rate constants extrapolated to 25 °C.

The energy decomposition of the calculated activation barrier for the hydrolysis of the dGTP analogues examined in this work is shown in Table 1; the rate constants (obtained from transition state theory, assuming a temperature of 250 °C) are listed in Table 2, and the calculated LFER is shown in Figure 7. As with the enzymatic reaction (*17*) (Figure 2), the mono- and dihalo compounds show significantly different k_{calc} behavior, with the analogue closest to the native dGTP (X = O), the methylene analogue, and the monohalogenated compounds conforming to one linear relationship and the dihalogenated compounds conforming to a separate linear relationship that deviates significantly from that of methylene and monohalogenated analogues. However, qualitatively, the free energy surfaces for the hydrolysis of all the dGTP analogues in solution are similar (see the Supporting Information). In all cases, hydrolysis proceeds through a single concerted pathway, with only subtle variations in transition state geometries. In each case, a proton is transferred from the attacking water molecule to the phosphate concomitantly to P−O bond formation or cleavage, which can be seen from the transition state depicted in Figure 6.

Table 1: Energy Decomposition of the Calculated Activation Barrier for the Hydrolysis for the dGTP Analogues in Solution[a]

X	pK_a	ΔE_{gas}	ΔE_{pol}	$\Delta\Delta G_{solv}$	$\Delta G^{\ddagger}_{calc,sol}$	$\Delta G^{\ddagger}_{enz,R}$	$\Delta G^{\ddagger}_{enz,W}$
O	8.9	9.9	−1.4	30.1	38.6	28.6	31.1
CHF	9.0	9.1	−2.4	30.0	36.7	29.6	31.5
CHCl	9.5	28.3	1.0	14.0	43.3	29.3	33.0
CHBr	9.9	34.9	3.1	7.0	45.0	29.4	33.5
CH_2	10.5	31.7	2.5	11.6	45.8	30.9	33.6
CF_2	7.8	23.6	0.4	16.0	40.0	28.2	31.7
CFCl	8.4	19.4	0.5	21.0	40.9	28.6	33.0
CCl_2	8.9	22.9	0.2	23.7	46.8	29.2	34.5
CBr_2	9.3	23.8	0.3	21.6	45.7	30.3	35.5

[a] $\Delta G^{\ddagger}_{calc,sol}$ denotes the calculated activation barrier for each compound, and $\Delta G^{\ddagger}_{enz,R}$ and $\Delta G^{\ddagger}_{enz,W}$ denote the activation barriers for the incorporation of each compound opposite right (R) and wrong (W) template bases, respectively (*17*), as calculated from the rate constants provided in ref *17*.

Table 2: Calculated Activation Barriers and Rate Constants for the Hydrolysis of the dGTP Analogues in Solution[a]

X	pK_a	$\Delta G^{\ddagger}_{calc,sol}$	k	log(k)
O	8.9	38.6	9.80×10^{-4}	−3.01
CHF	9.0	36.7	6.30×10^{-3}	−2.20
CHCl	9.5	42.8	1.10×10^{-5}	−4.95
CHBr	9.9	45.0	2.15×10^{-6}	−5.67
CH_2	10.5	45.8	9.98×10^{-7}	−6.00
CF_2	7.8	40.0	2.63×10^{-4}	−3.57
CFCl	8.4	40.9	1.11×10^{-4}	−3.95
CCl_2	8.9	46.8	3.80×10^{-7}	−6.41
CBr_2	9.3	45.7	1.10×10^{-6}	−5.96

[a] $\Delta G^{\ddagger}_{calc,sol}$ denotes the calculated activation barrier for each compound, and k denotes the corresponding rate constant as obtained by transition state theory (assuming a temperature of 250 °C).

Table 3 shows the P−O distances with respect to the nucleophile and leaving group in the transition state for the hydrolysis of each of the dGTP analogues. From this table, one can see that not only does each hydrolysis proceed through a concerted A_ND_N transition state but also within small variation, these transition states are similar in geometry. Thus, despite the fact that the compounds clearly form two subsets, each conforming to a different linear relationship, the hydrolysis of each compound proceeds through a qualitatively similar reaction pathway and transition state.

The calculated activation barriers for the hydrolysis of the dGTP analogues in the gas phase and the corresponding calculated LFER are shown in Table 4 and Figure 7, respectively, for comparison to the solution LFER. There is no longer a segregation of the dGTP analogues into two subsets, but rather, despite the presence of some scattering, all compounds approach a single linear relationship, suggesting that the effect observed in the solution LFER is a solvation effect. Table 5 shows select atomic charges in both the reactant and transition states, which were obtained by performing charge fitting to the electrostatic potential at points selected according to the Merz−Singh−Kollman (*57*, *58*) scheme using the COSMO solvation model. From this, it can be seen that there are really only very small differences across the series but that the charges show a systematic trend, and thus, the only way to explore their meaning is by performing solvation calculations.

From the energy decomposition shown in Table 1, one can see that there is a significant difference in solvation between the monohalogenated compounds and their dihalogenated counterparts; that is, when the halogen substituent is chlorine or bromine, there is a somewhat larger solvation contribution ($\Delta\Delta G_{solv}$) to the overall activation barrier ($\Delta G^{\ddagger}_{calc}$) in the case of the dihalogenated compound than in the case of the monohalogenated compound. However, this trend is reversed in the case when the halogen substituent is fluorine. Here, $\Delta\Delta G_{solv}$ is much larger in the case of the monohalogenated compound than in the case of the dihalogenated compound. To verify that these observed trends are not coincidental, we simulated solvation

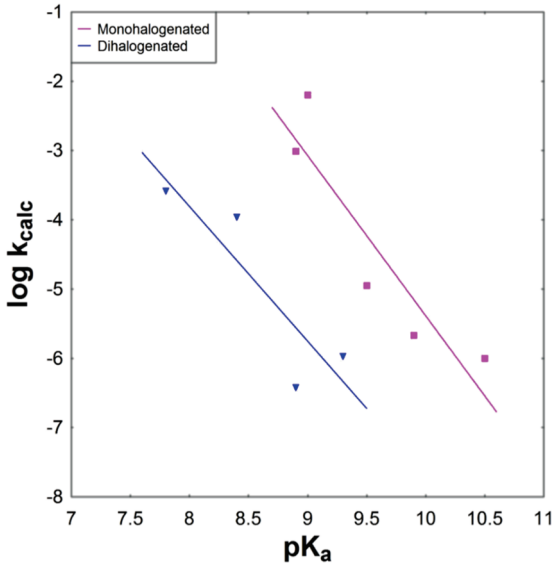

FIGURE 7: Calculated LFER for the hydrolysis of the dGTP analogues in solution.

Table 3: P−O Distances with Respect to the Nucleophile and Leaving Group in the Transition State for the Hydrolysis of the dGTP Analogues in Solution

X	pK_a	P−O_{nuc} (Å)	P−O_{lg} (Å)
O	8.9	2.0	1.8
CHF	9.0	2.0	1.8
CHCl	9.5	2.0	1.8
CHBr	9.9	2.0	1.95
CH_2	10.5	2.0	1.65
CF_2	7.8	2.0	1.8
CFCl	8.4	2.0	1.95
CCl_2	8.9	2.0	1.8
CBr_2	9.3	2.0	1.8

Table 4: Calculated Activation Barriers and Rate Constants for the Hydrolysis for the dGTP Analogues in the Gas Phase[a]

X	pK_a	$\Delta E^{\ddagger}_{calc,gas}$	k	log(k)
O	8.9	23.6	1862.1	3.27
CHF	9.0	30.5	2.45	0.39
CHCl	9.5	30.0	1.66	0.22
CHBr	9.9	34.9	0.04	−1.45
CH_2	10.5	32.0	0.56	−0.24
CF_2	7.8	25.5	263.0	2.42
CFCl	8.4	27.1	64.6	1.81
CCl_2	8.9	33.9	0.09	−1.03
CBr_2	9.3	33.7	0.11	−0.94

[a] $\Delta E^{\ddagger}_{calc,gas}$ denotes the calculated activation barrier for each compound, and k denotes the corresponding rate constant, as obtained from transition state theory (assuming a temperature of 250 °C).

Table 5: Merz−Singh−Kollman (*57*, *58*) Charges on the Nucleophilic Oxygen Atom, the Phosphorus Atom, and the Oxygen of the Departing Leaving Group, as Evaluated by the COSMO Solvation Model, in both the Reactant (RS) and Transition (TS) States

		RS			TS		
X	pK_a	O_{nuc}	P	O_{lg}	O_{nuc}	P	O_{lg}
O	8.9	−0.88	1.35	−0.67	−0.84	1.43	−0.84
CHF	9.0	−0.88	1.39	−0.64	−0.87	1.44	−0.79
CHCl	9.5	−0.88	1.31	−0.62	−0.79	1.56	−0.86
CHBr	9.9	−0.88	1.38	−0.65	−0.75	1.46	−0.81
CH_2	10.5	−0.88	1.37	−0.69	−0.73	1.62	−0.78
CF_2	7.8	−0.88	1.31	−0.53	−1.03	1.48	−0.79
CFCl	8.4	−0.88	1.32	−0.55	−0.75	1.48	−0.84
CCl_2	8.9	−0.88	1.32	−0.59	−0.72	1.48	−0.84
CBr_2	9.3	−0.88	1.31	−0.56	−0.73	1.54	−0.91

using the PCM. Table 6 shows a comparison of $\Delta\Delta G_{solv}$ values for each compound using both solvation models. From this table, one can see that even though the absolute value of the solvation contribution is dependent on the precise solvation model, the trend remains the same regardless of whether we simulate solvation using COSMO or the PCM. Finally, we also examined the trends in solvation by means of the LD model. Once again, the overall trend was the same as that seen in COSMO and PCM; i.e., for both chlorine and bromine substituents, $\Delta\Delta G_{solv}$ was much smaller for the mono- than for the dihalogenated compounds. It should be noted that our LD calculations were not calibrated to the case where the halogen substituent is fluorine, in view of the difficulty of the calculations, and their possible instability. Thus, we prefer to focus on the most stable cases (in this case chlorine and bromine).

Table 6: Comparison of the Solvation Contribution ($\Delta\Delta G_{solv}$) to the Total Activation Barrier ($\Delta G^{\ddagger}_{calc}$) As Obtained from the COSMO, PCM, and LD Solvation Models

X	pK_a	$\Delta\Delta G_{solv,COSMO}$	$\Delta\Delta G_{solv,PCM}$	$\Delta\Delta G_{solv,LD}$
O	8.9	30.1	29.8	30.9
CHF	9.0	30.0	29.7	39.2
CHCl	9.5	14.0	13.9	17.6
CHBr	9.9	7.0	19.7	11.1
CH_2	10.5	11.6	11.4	15.5
CF_2	7.8	16.0	15.5	13.9
CFCl	8.4	21.0	20.8	28.8
CCl_2	8.9	23.7	23.4	24.2
CBr_2	9.3	21.6	21.4	17.1

In summarizing the results described above, we believe it is important to clarify what we have been trying to examine in this work and what we have actually been able to determine. We start by noting that the observed trend in the reactivity of these compounds can be classified in terms of two different effects, i.e., (i) having a LFER that correlates $\Delta G^{\ddagger}$ to the leaving group pK_a and (ii) having two sets of LFER. The basic origin of the LFER is conceptually simple, but difficult to quantify computationally. That is, the true basis of the LFER is the relationship between $\Delta G^{\ddagger}$, i.e., the activation barrier of the reaction, and $\Delta G°$, i.e., the free energy difference between the reactants and products (see the discussion in ref *59* and references cited therein). The problem becomes more complex when we are dealing with several intersecting parabolas(*27*). Now, the dependence of the reaction rate on the leaving group pK_a comes indirectly from the correlation between the pK_a and $\Delta G°$. Thus, we would expect correlation between the intersections of the Marcus parabolas (and thus $\Delta G^{\ddagger}$) and the pK_a of the leaving group, if the intersection is correlated to the energy of the product state, and if the energy of the product state is in turn correlated to the pK_a. However, although we have succeeded in reproducing a LFER that correlates the calculated $\Delta G^{\ddagger}$ with the observed pK_a, we have encountered major difficulties in trying to quantify the origin of the LFER, since correlating $\Delta G^{\ddagger}$ to $\Delta G°$ gave unstable results. This is partly due to the fact that the nature of the proton transfer (PT) step changes in the different systems, where in some cases PT occurs before the rate-determining step and in others after this step. It should be noted that the pK_a values used in this work are those that were originally published in ref *17*, allowing us to maintain the same scale on the *x*-axis as that used for the corresponding LFER for the pol β-catalyzed reaction. We have to emphasize that our results and conclusions would not change at all if the pK_a scale were to be modified in both the protein and solution cases.

While the quantification of the precise origin of the observed LFER is left for further studies, we have demonstrated that the reason for the different behavior of the mono- and dihalogens can be explored by examining the energies of the corresponding transition states at the observed pK_a in solution. If the calculations can reproduce the difference in $\Delta G^{\ddagger}$ between the mono- and dihalogens at any given pK_a, then we may explore the reason for this computational difference. Indeed, after we found that the calculations reproduced the observed trend, our study indicated that the change in the TS charge distribution led to a difference in the corresponding solvation energy that we suggest is principally responsible for the segregation of the data into two separate LFER.

CONCLUSIONS

DNA polymerase catalysis and fidelity studies typically compare the incorporation of W−C correct pairs versus non−W-C mispairs, where the leaving group is the natural pyrophosphate. Recently (*16*, *17*), we altered the electronic and steric properties of the leaving groups by replacing PP_i with a series of halomethylene derivatives in the β,γ-bridging position. This strategy made possible the investigation of the catalytic mechanisms of right and wrong incorporations proceeding "downstream" from the transition state. Here, we have performed an extensive computational analysis, involving the evaluation of the complete free energy surface in solution for each of the systems explored, while using different simulation models. The calculations were used to explore whether the separate linear relationships of the mono- and dihalogenated dNTP analogues might at least to some extent from their intrinsic properties when they undergo hydrolysis in aqueous solution. A key element of our analysis is the use of MFJ plots to characterize transition states for the hydrolysis of the dGTP analogues presented in refs *16* and *17*, in both the gas phase and solution. In all cases, the reaction proceeds through a single concerted A_ND_N associative transition state, with no clear trend in $P-O_{nuc}$ and $P-O_{lg}$ distances across the series (cf. Table 3). The calculations reproduced an LFER with a trend similar to the one observed in the enzyme when using the calculated activation barriers and the observed pK_a values in solution. The only clear difference between the mono- and dihalogenated dGTP analogues occurred when the energy decomposition of the total calculated activation free energies in each case was examined (Table 1). Here, there is a clear difference in the solvation contribution to the total free energy between the monohalogenated and dihalogenated compounds. When the halogen is chlorine or bromine, there is a significantly larger $\Delta\Delta G_{solv}$ in the case of the dihalogenated than in the case of the monohalogenated compounds, and this trend is reversed when the halogen is fluorine.

We observe this phenomenon only in solution. In the gas phase, there is no clear pattern distinguishing hydrolysis rates for the mono- and dihalogens, with both sets of compounds falling roughly on one line (Figure 8). Thus, in this study the main cause for the discrepancy in the rate constants between the two series of compounds appears to arise from the difference in how these compounds are interacting with the solvent. To verify this, we simulated solvation using three different solvent models (COSMO, PCM, and LD), and in all three cases, we observed the same trend. The origin of this solvation effect was traced to a small but well-defined change in the TS charge distribution between the two systems. The difference in the interaction of these TS charges and the solvent leads to the split between the two LFERs. If this finding stands the test of time, then the origin of the discrepancy in the magnitude of this segregation in the polymerase reaction for the correct and incorrect base pairs is due, at least in part, to the change in the electrostatic interaction between the protein active site and the TS charges.

At this point, we must emphasize that the calculations reported here are extremely challenging, since we are dealing with small effects that compensate for the gas-phase trend in a way which is not fully intuitive. Thus, our results still present a somewhat speculative prediction (the experimental data for the solution hydrolyses of the halomethylene derivatives are not currently available to make a direct comparison with theory). However, we note that in cases where experimental data are available, we have

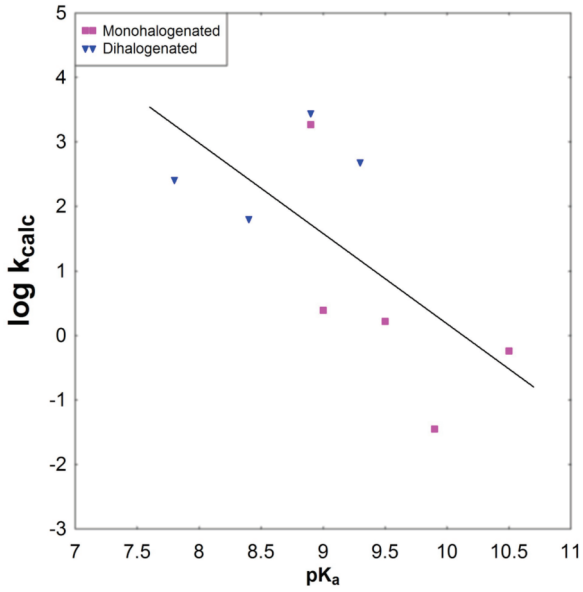

FIGURE 8: Calculated LFER for dGTP and analogues in the gas phase.

accurately reproduced experimental activation free energies for phosphate hydrolysis using our current protocol (*19*, *22*, *34*, *35*). Here, we have used exactly the same approach as in our previous studies, which have been shown to successfully reproduce solution LFER for phosphate hydrolysis (*19*, *22*). Thus, we believe that despite the absence of experimental rate constants for the reaction in solution, our predictions with regard to the trend we observe for the hydrolysis of these compounds are likely to represent the actual chemistry in solution. It is possible that the final judgment will be obtained by solution experiments, assuming that the mechanism in solution is the same as that in the enzyme. Otherwise, there will be a need for theoretical studies in which the solvent dielectric constant is varied, and the subsequent analysis of these experiments (if the mechanism is the same). Perhaps the most noteworthy result of the computational analysis is that the LFER in solution follows the same trend as that observed for the pol β-catalyzed reaction (*1*, *2*), with the native dGTP analogue including compounds with monohalogenated substituents conforming to one linear relationship and the compounds with dihalogenated substituents fitting to a second independent linear relationship (compare Figures 2 and 7).

As stated at the end of the previous section, our results suggest that the split of the LFER plots for the pol β-catalyzed incorporation of the dGTP analogues (*16*, *17*) into two subsets reflects a generalized solvation phenomenon (where we define such effects as the overall electrostatic interaction between the substrate and the enzyme/solvent system). Thus, the fact that the discrepancy between the two subsets of compounds is larger in the case of W than R nucleotide incorporation is likely to reflect the differences in the electrostatic interaction between the TS charges and the active site microenvironments induced by introduction of the mismatched versus matched nucleotides.

We believe that our results highlight the importance of the ability to differentiate between effects that are caused by the enzyme and effects that are a result of the chemistry of the substrate, and would in principle be observed even in the absence of the enzyme. We would like to re-emphasize here that in this case, we are examining a reference solution reaction that follows the same mechanism as that in the enzyme (namely nucleophilic attack at the α phosphate). Careful examination of the hydrolysis

of these dGTP analogues in solution should allow us to move on to the enzymatic reaction in order to understand why there is a larger discrepancy between the two subsets in the case of W rather than R nucleotide insertion. We believe that understanding the hydrolysis of these compounds in solution is an essential stepping-stone toward understanding the factors governing the fidelity of DNA polymerases. Of course, if the actual solution reaction will involve a different mechanism that that in the enzyme, we will have to continue to rely on the theoretical study of the chemically filtered reference reaction in the analysis of the enzymatic environment in the LFER for the *enzymatic* mechanism.

SUPPORTING INFORMATION AVAILABLE

Full free energy surfaces and transition states for all compounds studied in this work. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Kornberg, A., Lehman, I. R., Bessman, M. J., and Simms, E. S. (1956) Enzymic synthesis of deoxyribonucleic acid. *Biochim. Biophys. Acta 21*, 197–198.
2. Lehman, I. R., Zimmerman, S. B., Adler, J., Bessman, M. J., Simms, E. S., and Kornberg, A. (1958) Enzymatic synthesis of deoxyribonucleic acid v. chemical composition of enzymatically synthesized deoxyribonucleic acid. *Proc. Natl. Acad. Sci. U.S.A. 44*, 1191–1196.
3. Goodman, M. F. (2002) Error-prone repair DNA polymerases in prokaryotes and eukaryotes. *Annu. Rev. Biochem. 71*, 17–50.
4. Friedberg, E. C., Wagner, R., and Radman, M. (2002) Specialized DNA polymerases, cellular survival, and the genesis of mutations. *Science 296*, 1627–1630.
5. Zerbe, L. K., and Kuchta, R. D. (2002) The p58 subunit of human DNA primase is important for primer initiation, elongation and counting. *Biochemistry 41*, 4891–4900.
6. Mizuno, T., Yamagishi, K., Miyazawa, H., and Hanaoka, F. (1999) Molecular architecture of the mouse DNA polymerase α-primase complex. *Mol. Cell. Biol. 19*, 7886–7896.
7. Gavin, A.-C., and Superti-Furga, G. (2003) Protein complexes and proteome organization from yeast to man. *Curr. Opin. Chem. Biol. 7*, 21–27.
8. Batra, V. K., Beard, W. A., Shock, D. D., Pedersen, L. C., and Wilson, S. H. (2005) Nucleotide-induced DNA polymerase active site motions accomodating a mutagenic DNA intermediate. *Structure 13*, 1225–1233.
9. Batra, V. K., Shock, D. D., Prasad, R., Beard, W. A., Hou, E. W., Pedersen, L. C., Sayer, J. M., Yagi, H., Kumar, S., Jerina, D. M., and Wilson, S. H. (2006) Structure of DNA polymerase β with a benzo[c] phenanthrene diol epoxide-adducted template exhibits inutagenic features. *Proc. Natl. Acad. Sci. U.S.A. 103*, 17231–17236.
10. Batra, V. K., Beard, W. A., Shock, D. D., Krahn, J. M., Pedersen, L. C., and Wilson, S. H. (2006) Magnesium-induced assembly of a complete DNA polymerase catalytic complex. *Structure 14*, 757–766.
11. McKenna, C. E., Kashemirov, B. A., Upton, T. G., Batra, V. K., Goodman, M. F., Pedersen, L. C., Beard, W. A., and Wilson, S. H. (2007) (R)-β,γ-fluoromethylene-dGTP-DNA ternary complex with DNA polymerase β. *J. Am. Chem. Soc. 129*, 15412–15413.
12. Xiang, Y., Goodman, M. F., Beard, W. A., Wilson, S. H., and Warshel, A. (2008) Exploring the role of large conformational changes in the fidelity of DNA polymerase β. *Proteins: Struct., Funct., Bioinf. 70*, 231–247.
13. Florián, J., Goodman, M. F., and Warshel, A. (2003) Computer simulations of the chemical catalysis of DNA polymerases: Discriminating between alternative nucleotide insertion mechanisms for T7 DNA polymerase. *J. Am. Chem. Soc. 125*, 8163–8177.
14. Lin, P., Batra, V. K., Pedersen, L. C., Beard, W. A., Wilson, S. H., and Pedersen, L. G. (2008) Incorrect nucleotide insertion at the active site of a G:A mismatch catalyzed by DNA polymerase β. *Proc. Natl. Acad. Sci. U.S.A. 105*, 5670–5674.
15. Alberts, I. L., Wang, Y. A., and Schlick, T. (2007) DNA polymerase β catalysis: Are different mechanisms possible?. *J. Am. Chem. Soc. 129*, 11100–11110.
16. Sucato, C. A., Upton, T. G., Kashemirov, B. A., Batra, V. K., Martinek, V., Xiang, Y., Beard, W. A., Pedersen, L. C., Wilson, S. H., and McKenna, C. E. (2007) Modifying the β,γ leaving group bridging oxygen alters nucleotide incorporation efficiency, fideltiy, and the catalytic mechanism of DNA polymerase β. *Biochemistry 46*, 461–471.
17. Sucato, C. A., Upton, T. G., Kashemirov, B. A., Osuna, J., Oertell, K., Beard, W. A., Wilson, S. H., Florián, J., Warshel, A., McKenna, C. E., and Goodman, M. F. (2008) DNA polymerase β fidelity: Halomethylene-modified leaving groups in pre-steady-state kinetic analysis reveal differences at the chemical transition state. *Biochemistry 47*, 870–879.
18. McKenna, C. E., and Shen, P. D. (1981) Fluorination of methanediphosphonate esters by perchloryl fluoride. Synthesis of fluoromethanediphosphonic acid and difluoromethanediphosphonic acid. *J. Org. Chem. 46*, 4573.
19. Klähn, M., Rosta, E., and Warshel, A. (2006) On the mechanism of hydrolysis of phosphate monoester dianions in solution and proteins. *J. Am. Chem. Soc. 128*, 15310–15323.
20. Wilkie, J., and Gani, D. (1996) Comparison of inline and non-inline associative and dissociative reaction pathways for model reactions of phosphate monoester hydrolysis. *J. Chem. Soc., Perkin Trans. 2,* 783–787.
21. Kamerlin, S. C. L., and Wilkie, J. (2007) The role of metal ions in phosphate ester hydrolysis. *Org. Biomol. Chem. 5*, 2098–2108.
22. Rosta, E., Kamerlin, S. C. L., and Warshel, A. (2008) On the interpretation of the observed LFER in phosphate hydrolysis: A thorough computational study of phosphate diester hydrolysis in solution. *Biochemistry 47*, 3725–3735.
23. Friedman, J. M., Freeman, S., and Knowles, J. R. (1988) The quest for free metaphosphate in solution. Racemization at phosphorus in the transfer of the phospho group from aryl phosphate monoesters to tert-butyl alcohol in acetonitrile or in tert-butyl alcohol. *J. Am. Chem. Soc. 110*, 1268–1275.
24. Kirby, J. A., and Jencks, W. P. (1965) The reactivity of nucleophillic reagents towards the p-nitrophenyl phosphate dianion. *J. Am. Chem. Soc. 87*, 3209–3216.
25. Kirby, J. A., and Varvoglis, A. G. (1967) The reactivity of phosphate monoester hydrolysis. *J. Am. Chem. Soc. 89*, 415–423.
26. Florian, J., Aqvist, J., and Warshel, A. (1998) On the reactivity of phosphate monoester dianions in aqueous solution: Bronsted linear free-energy relationships do not have a unique mechanistic interpretation. *J. Am. Chem. Soc. 120*, 11524–11525.
27. Aqvist, J., Kolmodin, K., Florian, J., and Warshel, A. (1999) Mechanistic alternatives in phosphate monoester hydrolysis: What conclusions can be drawn from available experimental data?. *Chem. Biol. 6*, R71–R80.
28. Florián, J., and Warshel, A. (1998) Phosphate ester hydrolysis in aqueous solution: Associative versus dissociative mechanisms. *J. Phys. Chem. B 102*, 719–734.
29. Hengge, A. (2005) *Adv. Phys. Org. Chem. 40*, 49.
30. Mercero, J. M., Barrett, P., Lam, C. W., Fowler, J. E., Ugalde, J. M., and Pedersen, L. G. (2000) Quantum mechanical calculations on phosphate hydrolysis reactions. *J. Comput. Chem. 21*, 43–51.
31. Liu, Y., Gregersen, A., Hengge, A., and York, D. M. (2006) Transesterification thio effects of phosphate diesters: Free energy barriers and kinetic and equilibrium isotope effects from density functional theory. *Biochemistry 45*, 10043–10053.
32. Lopez, X., Dejaegere, A., Leclerc, F., York, D. M., and Karplus, M. (2006) Nucleophilic attack on phosphate diesters: A density functional study of in-line reactivity in dianionic, monoanionic, and neutral systems. *J. Phys. Chem. B 110*, 11525–11539.
33. Liu, Y., Gregersen, B. A., Lopez, X., and York, D. M. (2005) Density functional study of the in-line mechanism of methanolysis of cyclic phosphate and thiophosphate esters in solution: Insight into thio effects in RNA transesterification. *J. Phys. Chem. B 109*, 19987–20003.
34. Kamerlin, S. C. L., Florian, J., and Warshel, A. (2008) Associative versus dissociative mechanisms of phosphate monoester hydrolysis: On the interpretation of activation entropies. *Chem. Phys. Chem. 9*, 1767–1773.
35. Kamerlin, S. C. L., Williams, N. H., and Warshel, A. (2008) Dineopentyl phosphate hydrolysis: Evidence for stepwise water attack. *J. Org. Chem. 73*, 6960–6969.
36. Jencks, W. P. (1985) A primer for the bema hapothle. An empirical approach to the characterization of changing transition-state structures. *Chem. Rev. 85*, 511–527.
37. More O'Ferrall, R. A. (1970) Relationships between E2 and E1cB mechanisms of β-elimination.. *J. Chem. Soc. B,* 274–277.
38. Shurki, A., and Warshel, A. (2003) Structure/function correlations of proteins using MM, QM/MM and related approaches: Methods, concepts, pitfalls and current progress. *Adv. Protein Chem. 66*, 249–312.

39. Warshel, A., Sharma, P. K., Kato, M., Xiang, Y., Liu, H., and Olsson, M. H. M. (2006) Electrostatic basis for enzyme catalysis. *Chem. Rev. 106*, 3210–3235.
40. Wolfenden, R., and Snider, M. J. (2001) The depth of chemical time and the power of enzymes as catalysts. *Acc. Chem. Res. 34*, 938–945.
41. Lad, C., Williams, N. H., and Wolfenden, R. (2003) The rate of hydrolysis of phosphomonoester dianions and the exceptional catalytic proficiencies of protein and inositol phosphatases. *Proc. Natl. Acad. Sci. U.S.A. 100*, 5607–5610.
42. Schowen, R. L. (1978) *Transition states in biological processes*, Plenum, New York.
43. Rogers, G. A., and Bruice, T. C. (1974) Synthesis and evluation of a model for the so-called charge-relay system of the serine esterases. *J. Am. Chem. Soc. 96*, 2473–2481.
44. Schweins, T., Langen, R., and Warshel, A. (1994) Why have mutagenesis studies not located the general base in ras p21. *Nat. Struct. Biol. 1*, 476–484.
45. Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Montgomery, J. A., Jr., Vreven, T., Kudin, K. N., Burant, J. C., Millam, J. M., Iyengar, S. S., Tomasi, J., Barone, V., Mennucci, B., Cossi, M., Scalmani, G., Rega, N., Petersson, G. A., Nakatsuji, M., Hada, M., Ehara, K., Toyota, R., Fukuda, J., Hasegawa, M., Ishida, T., Nakajima, Y., Honda, Y., Kitao, O., Nakai, H., Klene, M., Li, X., Knox, J. E., Hratchian, H. P., Cross, J. B., Adacmo, C., Jaramillo, J., Gomperts, R., Stratmann, R. E., Yazyev, O., Austin, A. J., Cammi, R., Pomelli, C., Ochterski, J., Ayala, P. Y., Morokuma, K., Voth, G. A., Salvador, P., Dannenberg, J. J., Zakrzewski, V. G., Dapprich, S., Daniels, A. D., Strain, M. C., Farkas, O., Malick, D. K., Rabuck, A. D., Clifford, K., Cioslowki, J., Stefanov, B. B., Liu, G., Liashenko, A., Piskorz, P., Komaromi, I., Martin, R. L., Fox, D. J., Keith, T., Al-Laham, M. A., Peng, C. Y., Nanayakkara, A., Challacombe, M., Gill, P. M. W., Johnson, B. G., Chen, W., Wong, M. W., Gonzalez, C., and Pople, J. A. (2004) *GAUSSIAN 03* , revision c.02, Gaussian, Inc., Pittsburgh, PA.
46. Becke, A. D. (1993) Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys. 98*, 5648–5652.
47. Barone, V., and Cossi, M. (1998) Quantum calculation of molecular energies and energy gradients in solution by a conductor solvent model. *J. Phys. Chem. A 102*, 1995–2001.
48. Klamt, A., and Schuurmann, G. J. (1993) COSMO: A new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. *J. Chem. Soc., Perkin Trans. 2,* 799–805.
49. Tomasi, J., Mennucci, B., and Cancès, M. T. (1991) *THEOCHEM 464*, 464.
50. Cancès, M. T., Mennucci, B., and Tomasi, J. (1997) A new integral equation formalism for the polarizable continuum model: Theoretical background and applications to isotropic and anisotropic dielectrics. *J. Chem. Phys. 107*, 3032–3041.
51. Mennucci, B., and Tomasi, J. (1997) Continuum solvation models: A new approach to the problem of solute's charge distribution and cavity boundaries. *J. Chem. Phys. 106*, 5151–5158.
52. Mennucci, B., Cancès, M. T., and Tomasi, J. (1997) Evaluation of solvent effects in isotropic and anisotropic dielectrics and in ionic solutions with a unified integral equation method: Theoretical bases, computational implementation, and numerical applications. *J. Phys. Chem. B 101*, 10506–10517.
53. Cossi, M., Scalmani, G., Rega, N., and Barone, V. (2002) New developments in the polarizable continuum model for quantum mechanical and classical calculations on molecules in solution. *J. Chem. Phys. 117*, 43–54.
54. Florián, J., and Warshel, A. (1999) Calculation of hydration entropies of hydrophobic, polar, and ionic solutes in the framework of the Langevin dipole solvation model. *J. Phys. Chem. B 103*, 10282–10288.
55. Rappé, A. K., Casewit, C. J., Colwell, K. S., Goddard, W. A.III, and Skiff, W. M. (1992) UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations. *J. Am. Chem. Soc. 114*, 10024–10035.
56. Schroeder, G. K., Lad, C., Wyman, P., Williams, N. H., and Wolfenden, R. (2006) The time required for water attack at the phosphorus atom of simple phosphodiesters and of DNA. *Proc. Natl. Acad. Sci. U.S.A. 103*, 4052–4055.
57. Besler, B. H., Merz, K. M.Jr., and Kollman, P. A. (1990) Atomic charges derived from semiempirical methods. *J. Comput. Chem. 11*, 431–439.
58. Singh, U. C., and Kollman, P. A. (1984) An approach to computing electrostatic charges for molecules. *J. Comput. Chem. 5*, 129–145.
59. Warshel, A. (1991) *Computer modeling of chemical reactions in enzymes and solutions*, John Wiley and Sons, New York.