1. INTRODUCTION

DNA polymerase (pol)-catalyzed nucleotidyl transfer is a fundamental chemical reaction essential for cellular propagation.1,2 Because of the critical nature of this reaction, it has been the target of a wide variety of chemotherapeutic drugs to control aberrant cell growth.3 Initial characterization of the DNA synthesis reaction indicated that it was metal- and template-dependent.4,5 Structural characterization identified a bimetalic reaction scaffold at the pol active site with two closely placed magnesium or manganese ions of which one metal ion (Me1) facilitates incoming dNTP binding while the other metal ion (Me2) lowers the pKα of O3′ of the growing DNA strand (i.e., primer) terminus, leading to deprotonation and generating an oxyanion.6 A nucleophilic in-line attack of this oxyanion on Pα of the incoming nucleotide involves a penta-coordinated transition state that forms a phosphodiester bond (O3′−Pα) and breaks the incoming dNTP Pα−Oαβ bond, generating pyrophosphate (PPi) and DNA that has been extended by one nucleotide (dNMP).6,8,9 As with most enzymes, this reaction is reversible and is referred to as pyrophosphorolysis; activated PPi oxygen attacks phosphate of the DNA primer terminus, generating dNTP and DNA primer strand that has been shortened by one nucleotide.2

DNA polymerase β is a model pol that has been structurally and kinetically extensively studied, employing natural and chemically modified substrates.7,11−16 A recent study utilizing a modified dGTP substrate, the bridging oxygen between Pγ and Py of dGTP dramatically decreased the overall equilibrium enzymatic reaction strongly favors DNA synthesis over pyrophosphorolysis with natural substrates. The DNA polymerase β chemical reaction has been structurally and kinetically characterized, employing natural and chemically modified substrates. Substituting an imido-moiety (NH) for the bridging oxygen between Pβ and Py of dGTP dramatically decreased the overall enzymatic activity and resulted in a chemical equilibrium that strongly favors the reverse reaction (i.e., K ≪ 1). Using QM/MM calculations in conjunction with the utilization of parameters such as quantum mechanically derived atomic charges, we have examined the chemical foundation for the altered equilibrium with this central biological reaction. The calculations indicate that the rapid reverse reaction is likely due, in part, to the increased nucleophilicity of the reactive oxygen on the tautomeric form of imidodiphosphate.
the importance of the bridging atom between the β and γ-phosphates in the reaction path of pyrophosphorolysis. 13,16

When the bridging oxygen is replaced by a nitrogen atom that is protonated, the imidodiphosphate substrate can be found in various tautomeric forms. 17,18 Such tautomerizations, in general, are recognized as a key feature in many enzymatic reactions. In fact, Bax et al.19 summarized structural findings on bond distances and angles for PNP compounds using the small molecular crystal database (CSD) and found that the bridging nitrogen was not protonated when the phosphate oxygen of each phosphate is coordinated with a divalent metal ion. In this tautomeric form, the proton may have transferred to a phosphate oxygen or left PNP entirely. In the conversion of dihydroflavonols to anthocyanins, which are important plant secondary metabolites that act as photo-protectants and antioxidants, the importance of such tautomeric conversions is recognized. 20

The relative efficiency of nucleophilic attack of an oxygen atom at one of the phosphate groups in PPi or its analogues, depends on the identity of the bridging atom between the phosphates. The electronic environment of the attacking oxygen may be altered in an analogue, and the nucleophilicity of the attacking oxygen is a good indicator for the reactivity of the substrate. Liu et al.21,22 demonstrated that the Hirshfeld charges (and charges derived from similar methods, such as the CM5 method23) derived from quantum mechanical calculations, as well as Fukui functions based on these charges, provide a good theoretical framework to quantify the oxygen atom reactivity. In the present work, we used the energy barriers from the reaction path calculations and the chemical concept of nucleophilicity, quantified by charges and Fukui functions, to characterize the relative reactivity of various pyrophosphate derivatives.

2. METHODS

2.1. QM/MM Calculations. High-resolution crystallographic structures of the enzyme–DNA complexes were used as the initial configurations for theoretical calculations. After adding missing residues, heavy atoms, and hydrogen atoms to the initial structures (pdb ID: 4KLE), molecular dynamics simulations were carried out in a completely solvated aqueous medium and in a neutral environment by adding counter ions. Positions of all crystallographic water molecules were preserved initially. The total charge on PPi was taken to be −4e. All MD calculations were carried out with the Amber14SB force field within Gaussian 09.27 The calculations were performed with the electronic embedding option to accommodate the polarization of the QM region by the partial atomic charges in the classical region. Further details of the QM/MM calculations can be found in our previous work.14

2.2. QM Calculations on Clusters. To study the variations of the nucleophilicity of PPi and its derivatives, we used the QM systems and their smaller sub-systems from the QM/MM calculations described above. The entire QM sub system including the capping protons representing the boundary atoms from the QM/MM calculations was selected as the primary cluster system for charge calculations. To facilitate the discussion on how the strength of the nucleophile depends on contributions from various nearby components by peeling off different components from the QM system, instead of using the already polarized wavefunction from the QM/MM calculations, we decided to implement such a cluster model with implicit solvents for the calculations of charges. Various clusters selected from the primary system as described in the Results and Discussion section were used to analyze the effect of the presence of different components in the QM cluster. The charges at the atomic positions were calculated using the CMS charge model,28 which is an extension of the Hirshfeld population analysis, with DFT calculations at the B3LYP/6-311++G** level.29 Implicit solvent environment (corresponding to the solvent being water) was added by the use of the polarizable continuum model (PCM) in DFT calculations. CMS charges were calculated from the Gaussian output using the program CMSPAC version 2015.29 The Fukui functions were calculated using the method described in the work of Liu et al.21,22

3. RESULTS AND DISCUSSION

3.1. QM/MM Calculations. We previously reported results from a series of QM/MM calculations of pyrophosphorolysis using PPi as the substrate in the pol β-DNA complex active site. In that work, 14 with two magnesium ions occupying Mec and Mβ metal binding sites, we observed reaction energy continuing to increase well beyond a reasonable level for a biologically relevant reaction. The distance between the newly extended phosphate group and the former bridging oxygen on PPi, falls shorter than the corresponding distance at the transition state of the forward reaction.14 The energy profile did not change even when three metal ions were employed in the calculations, similar to the observed positions in the final product from time-lapse crystallography.30 Catalytically unrealistic reaction paths were the results of both calculations, with a lack of activation barriers and final energies over 50 kcal/mol.

In additional experiments under slightly different conditions, we re-evaluated the reaction path for the pyrophosphorolysis reaction of pol β, and the results remained the same as those mentioned above (Figure 1). The new calculation recalibrated the energy curve for the reverse reaction carried out for pol β with two metal ions where side chains of Arg183 and Arg254 were not considered to be in the quantum system because the results for the activation barrier calculated with the above-mentioned two residues did not seem to have any impact on the reverse reaction with the natural substrate. The reaction paths were virtually identical, even when the two arginine side chains were treated using molecular mechanics in place of
quantum mechanics. These energy curves serve as a reference for the problem addressed in the current study where the PPi substrate is replaced with PNP. Experimentally, as seen in Table 1, PNP preferentially promotes the reverse reaction ($K_{eq} \approx 0.002$). However, use of the same procedure to calculate the reaction path with PNP resulted in an energy barrier even more unrealistic than that observed with PPi.

Crystallographic structures of PNP and PPi yielded similar bond lengths and angles and no evidence for tautomeric forms.\textsuperscript{12} However, a shortened P−N bond distance was observed, suggesting partial double bond nature in the P−N bond. Subsequent solution NMR studies pointed to the possibility that the proton on the nitrogen was actually residing on phosphate oxygens.\textsuperscript{16} It was found that under mild conditions, the nitrogen was protonated. However, if the negative charge of the phosphate was neutralized by coordinating divalent metal ions, the imide group can be deprotonated.\textsuperscript{31} Another report pointed to the tautomerism activities of cyclo-imidodiphosphates.\textsuperscript{20} With this in mind, we decided to explore the reaction path of the PNP in its tautomeric form.

In the examination of the crystal structure of the immediate environment around PPi, one can identify a major difference in the interactions of the bridging atom between the β- and γ-phosphates. In the structure with PPi, the nearest water molecule points one of the protons toward the bridging oxygen atom. Crystallographically, this water molecule was not exactly located in the position found through molecular dynamics. In the structure with PNP, a water oxygen was observed in the crystal structure about 3 Å away from the bridging nitrogen of PNP. Because of the proton present on the nitrogen, the water oxygen is directed in this case, creating a somewhat different hydrogen-bonding network above this position. Potentially, such networks can contribute to alter the calculated reaction path.

In summary, we demonstrated that it is possible to observe a lower barrier for the reverse reaction when the nitrogen of PNP is unprotonated and is used as the substrate. In the following sections, we will examine a rationale for the reduction in the activation barrier by characterizing various

In the examination of the crystal structure of the immediate environment around PPi, one can identify a major difference in the interactions of the bridging atom between the β- and γ-phosphates. In the structure with PPi, the nearest water molecule points one of the protons toward the bridging oxygen atom. Crystallographically, this water molecule was not exactly located in the position found through molecular dynamics. In the structure with PNP, a water oxygen was observed in the crystal structure about 3 Å away from the bridging nitrogen of PNP. Because of the proton present on the nitrogen, the water oxygen is directed in this case, creating a somewhat different hydrogen-bonding network above this position. Potentially, such networks can contribute to alter the calculated reaction path.

Crystallographic structures of PNP and PPi yielded similar bond lengths and angles and no evidence for tautomeric forms.\textsuperscript{12} However, a shortened P−N bond distance was observed, suggesting partial double bond nature in the P−N bond. Subsequent solution NMR studies pointed to the possibility that the proton on the nitrogen was actually residing on phosphate oxygens.\textsuperscript{16} It was found that under mild conditions, the nitrogen was protonated. However, if the negative charge of the phosphate was neutralized by coordinating divalent metal ions, the imide group can be deprotonated.\textsuperscript{31} Another report pointed to the tautomerism activities of cyclo-imidodiphosphates.\textsuperscript{20} With this in mind, we decided to explore the reaction path of the PNP in its tautomeric form.

In the examination of the crystal structure of the immediate environment around PPi, one can identify a major difference in the interactions of the bridging atom between the β- and γ-phosphates. In the structure with PPi, the nearest water molecule points one of the protons toward the bridging oxygen atom. Crystallographically, this water molecule was not exactly located in the position found through molecular dynamics. In the structure with PNP, a water oxygen was observed in the crystal structure about 3 Å away from the bridging nitrogen of PNP. Because of the proton present on the nitrogen, the water oxygen is directed in this case, creating a somewhat different hydrogen-bonding network above this position. Potentially, such networks can contribute to alter the calculated reaction path.

Crystallographic structures of PNP and PPi yielded similar bond lengths and angles and no evidence for tautomeric forms.\textsuperscript{12} However, a shortened P−N bond distance was observed, suggesting partial double bond nature in the P−N bond. Subsequent solution NMR studies pointed to the possibility that the proton on the nitrogen was actually residing on phosphate oxygens.\textsuperscript{16} It was found that under mild conditions, the nitrogen was protonated. However, if the negative charge of the phosphate was neutralized by coordinating divalent metal ions, the imide group can be deprotonated.\textsuperscript{31} Another report pointed to the tautomerism activities of cyclo-imidodiphosphates.\textsuperscript{20} With this in mind, we decided to explore the reaction path of the PNP in its tautomeric form.

In summary, we demonstrated that it is possible to observe a lower barrier for the reverse reaction when the nitrogen of PNP is unprotonated and is used as the substrate. In the following sections, we will examine a rationale for the reduction in the activation barrier by characterizing various

Table 1. Summary of Observed Kinetic Parameters

| Substrate          | $k_{\text{forward}}$ ($\text{s}^{-1}$) | $k_{\text{reverse}}$ ($\text{s}^{-1}$) |
|--------------------|---------------------------------------|----------------------------------------|
| Forward/reverse    | (s$^{-1}$)                             | (s$^{-1}$)                              |
| dGTP/PPi           | 13.7b                                 | 0.03c                                  |
| dGDP(CH$_2$)P/P(CH$_2$)P | 1.7b                               | 2.0                                    |
| dGDP(N)P/PNP       | 0.07c                                 | 32.6c                                  |

aDNA substrate for the forward reaction was 1-nucleotide-gapped DNA; for the reverse reaction, the DNA substrate was nicked DNA. \textsuperscript{b}Taken from a previous study (Sucato et al., 2008). \textsuperscript{c}Taken from a previous study (Shock et al., 2017).

Figure 1. Energy profiles for the pol β-assisted pyrophosphorolysis with various P−X−P substrates. X = O (black), X = NH (red), and the tautomeric form with P−N≡P (green).

Figure 2. (a) Minimalist’s conformation of the reactive atoms as the largest system in the cluster calculation. All atoms with reported charges are marked. The hydrogens are in lighter gray, carbons in gray, oxygens in red, nitrogens in blue, and the two metal ions in light green. The pyrophosphate and the three residues of pol β involved in the catalytic scaffold are also marked. Water oxygens are marked with “W.” (b) Only two components in the quantum system undergo a chemical change: the phosphate group (from the terminal nucleotide) capped with two methyl groups and the pyrophosphate substrate that reacts with it during pyrophosphorolysis. (c) Tautomeric forms of the imidodiphosphate substrate used in the present study.
systems using simplified cluster models. Cluster calculations were performed on sub-components of the QM system used in the QM/MM study, and a reaction field scheme was adopted to account for solvent effects. Such stepwise isolation of individual system components allowed us to investigate features of the system where reactivity was observed. Thus,

Table 2. Charges and the Fukui Functions in $O_3P−O−PO_3$

| Atom | System 1 | System 2 | System 3 | System 4 | System 5 | System 6 | System 7 | System 8 | System 9 |
|------|----------|----------|----------|----------|----------|----------|----------|----------|----------|
| O$_{3a}$ | -0.61 | -0.62 | -0.65 | -0.64 | -0.66 | -0.54 | 0.12 | -0.62 | -0.49 | 0.13 | -0.66 |
| O$_{3b}$ | -0.49 | -0.49 | -0.55 | -0.58 | -0.67 | -0.55 | 0.12 | -0.52 | -0.48 | 0.04 | -0.68 |
| O$_{3c}$ | -0.66 | -0.66 | -0.70 | -0.69 | -0.68 | -0.58 | 0.10 | -0.64 | -0.51 | 0.13 | -0.67 |
| O$_{3d}$ | -0.52 | -0.52 | -0.60 | -0.63 | -0.66 | -0.48 | 0.18 | -0.52 | -0.45 | 0.07 | -0.67 |
| O$_{3e}$ | -0.64 | -0.65 | -0.69 | -0.68 | -0.68 | -0.55 | 0.13 | -0.64 | -0.47 | 0.17 | -0.67 |
| O$_{3f}$ | -0.58 | -0.59 | -0.62 | -0.70 | -0.69 | -0.55 | 0.14 | -0.65 | -0.45 | 0.20 | -0.69 |
| O$_{3g}$ | -0.34 | -0.34 | -0.36 | -0.36 | -0.36 | -0.33 | 0.03 | -0.33 | -0.30 | 0.03 | -0.35 |

Charges are evaluated using the CMS method, and the Fukui functions are calculated using $F_{index} = F(n) - F(n-1) \approx Q(n) - Q(n-1) = F$. When the magnesium ions were taken as classical point charges, the charges are displayed in red (with a (C) to denote they provide only a charge field for the QM calculation).

Table 3. Charges and the Fukui Functions in $O_3P−CH_2−PO_3$

| Atom | System 1 | System 2 | System 3 | System 4 | System 5 | System 6 | System 7 | System 8 | System 9 |
|------|----------|----------|----------|----------|----------|----------|----------|----------|----------|
| O$_{3a}$ | -0.64 | -0.65 | -0.69 | -0.67 | -0.70 | -0.55 | 0.15 | -0.65 | -0.48 | 0.17 | -0.69 |
| O$_{3b}$ | -0.50 | -0.50 | -0.54 | -0.58 | -0.68 | -0.57 | 0.11 | -0.53 | -0.47 | 0.06 | -0.70 |
| O$_{3c}$ | -0.67 | -0.67 | -0.71 | -0.69 | -0.58 | 0.11 | -0.65 | -0.47 | 0.18 | -0.68 |
| O$_{3d}$ | -0.54 | -0.54 | -0.60 | -0.66 | -0.68 | -0.50 | 0.18 | -0.53 | -0.47 | 0.06 | -0.70 |
| O$_{3e}$ | -0.65 | -0.66 | -0.70 | -0.69 | -0.68 | -0.57 | 0.11 | -0.64 | -0.57 | 0.07 | -0.67 |
| O$_{3f}$ | -0.65 | -0.63 | -0.67 | -0.71 | -0.70 | -0.58 | 0.12 | -0.66 | -0.55 | 0.11 | -0.70 |
| C$_{3g}$ | -0.20 | -0.20 | -0.21 | -0.21 | -0.21 | -0.20 | 0.01 | -0.20 | -0.18 | 0.02 | -0.21 |
| H$_1$ | 0.07 | 0.07 | 0.04 | 0.05 | 0.06 | 0.08 | 0.02 | 0.08 | 0.10 | 0.02 | 0.06 |
| H$_2$ | 0.07 | 0.07 | 0.05 | 0.05 | 0.05 | 0.08 | 0.02 | 0.07 | 0.10 | 0.03 | 0.06 |
| P$_1$ | 0.20 | 0.19 | 0.13 | 0.11 | 0.19 | 0.21 | 0.02 | 0.20 | 0.26 | 0.06 | 0.12 |
| P$_2$ | 0.20 | 0.20 | 0.13 | 0.15 | 0.20 | 0.29 | 0.09 | 0.20 | 0.30 | 0.10 | 0.12 |
| Mg$_{3m}$ | 0.77 | 0.74 | 2.0(C) | -- | -- | 1.34 | 1.40 | 0.06 | 2.0(C) |
| Mg$_{3n}$ | 0.82 | 2.0(C) | 2.0(C) | -- | -- | 1.32 | 1.37 | 0.05 | 2.0(C) |

Table 4. Charges and the Fukui Functions in $O_3P−NH−PO_3$

| Atom | System 1 | System 2 | System 3 | System 4 | System 5 | System 6 | System 7 | System 8 | System 9 |
|------|----------|----------|----------|----------|----------|----------|----------|----------|----------|
| O$_{3a}$ | -0.63 | -0.63 | -0.66 | -0.69 | -0.68 | -0.55 | 0.14 | -0.65 | -0.48 | 0.17 | -0.69 |
| O$_{3b}$ | -0.51 | -0.51 | -0.57 | -0.59 | -0.68 | -0.52 | 0.16 | -0.53 | -0.49 | 0.04 | -0.70 |
| O$_{3c}$ | -0.68 | -0.69 | -0.72 | -0.71 | -0.70 | -0.53 | 0.17 | -0.66 | -0.54 | 0.12 | -0.70 |
| O$_{3d}$ | -0.53 | -0.53 | -0.60 | -0.63 | -0.66 | -0.56 | 0.10 | -0.52 | -0.59 | 0.02 | -0.67 |
| O$_{3e}$ | -0.64 | -0.65 | -0.68 | -0.68 | -0.68 | -0.60 | 0.08 | -0.64 | -0.52 | 0.12 | -0.67 |
| O$_{3f}$ | -0.59 | -0.60 | -0.63 | -0.71 | -0.70 | -0.61 | 0.09 | -0.65 | -0.52 | 0.13 | -0.70 |
| N$_{3g}$ | -0.54 | -0.54 | -0.55 | -0.55 | -0.55 | -0.50 | 0.05 | -0.54 | -0.40 | 0.14 | -0.55 |
| H | 0.27 | 0.27 | 0.25 | 0.26 | 0.26 | 0.29 | 0.03 | 0.28 | 0.33 | 0.05 | 0.26 |
| P$_1$ | 0.30 | 0.29 | 0.23 | 0.19 | 0.20 | 0.27 | 0.07 | 0.29 | 0.37 | 0.08 | 0.20 |
| P$_2$ | 0.29 | 0.29 | 0.23 | 0.23 | 0.20 | 0.31 | 0.11 | 0.29 | 0.37 | 0.08 | 0.20 |
| Mg$_{3m}$ | 0.78 | 0.75 | 2.0(C) | -- | -- | 1.32 | 1.37 | 0.05 | 2.0(C) |
| Mg$_{3n}$ | 0.82 | 2.0(C) | 2.0(C) | -- | -- | 1.32 | 1.37 | 0.05 | 2.0(C) |
| P | 0.45 | 0.41 | 0.36 | 0.36 | 0.36 | 0.36 | 0.36 | 0.36 | 0.36 | 0.36 | 0.36 |
different bridging substituents between the two phosphate groups were tested for their influence in controlling the activation barrier of the reverse reaction.

3.2. Cluster Calculations. Various selected cluster systems were evaluated to isolate where the reactivity is concentrated and in the current case, for their charge redistribution in the complex assembly. In QM/MM systems, the high-level QM system consists of \( \text{P}^-\text{X}^-\text{P} \) (where \( \text{X} \) is \( \text{O}, \text{NH}, \) or \( \text{CH}_2 \)), a phosphate group representing the initial alpha-phosphate (with the two oxygen atoms capped with \( \text{CH}_3 \) groups and with two free oxygens), two divalent magnesium ions, three carboxyl groups representing pol active site aspartate groups, and two water molecules solvating the magnesium ions among other components. This description is taken as the minimum set of atoms required to represent the reactive system in the QM/MM analysis.

The following systems were selected in the charge calculations:

1. All atoms in the complete high-level QM system (Figure 2A).
2. As in system 1, but the catalytic magnesium was represented by a point charge with \( q = +2e \).
3. As in system 1, but both magnesium ions were represented by point charges with \( q = +2e \).
4. With only the \( \text{P}^-\text{X}^-\text{P} \) group and the phosphate group (see Figure 2B).
5. With only the \( \text{P}^-\text{X}^-\text{P} \) group.
6. With an electron removed from system 5.
7. The same as system 5, but the nucleotide binding magnesium was included.
8. With an electron removed from system 7.
9. The same as system 7, but the magnesium ion is replaced by a point charge with \( q = +2e \).

As shown in Tables 2–6, we tabulated the charges calculated from each system and included only selected atoms to facilitate the discussion. The metal-bound oxygens show reduced charges, and the positive charges (+2) of metal ions in the present ionization state were also reduced to less than a unit charge in all cases, with the values between 0.76 and 0.84. The charge neutralization on the metal ions was studied in detail in previous publications,\(^{14,32,33}\) and we use the effect here to highlight the effect of charge redistribution if metal ions are not included in the QM systems, as was performed in some theoretical work.\(^{34}\) In the present case, including the

| Atom | System 1 | System 2 | System 3 | System 4 | System 5 | System 6 | System 7 | System 8 | System 9 |
|------|----------|----------|----------|----------|----------|----------|----------|----------|----------|
| \( \text{O}_{\text{Me}} \) | -0.67    | -0.68    | -0.71    | -0.70    | -0.75    | -0.59    | 0.16     | -0.70    | -0.48    | 0.22     | -0.75    |
| \( \text{O}_{\text{Me}} \) | -0.53    | -0.53    | -0.58    | -0.63    | -0.72    | -0.56    | 0.16     | -0.55    | -0.49    | 0.06     | -0.73    |
| \( \text{O}_{\text{Me}} \) | -0.70    | -0.71    | -0.76    | -0.74    | -0.72    | -0.66    | 0.06     | -0.68    | -0.59    | 0.09     | -0.72    |
| \( \text{O}_{\text{Me}} \) | -0.53    | -0.53    | -0.58    | -0.66    | -0.68    | -0.62    | 0.06     | -0.53    | -0.59    | 0.03     | -0.69    |
| \( \text{O}_{\text{Me}} \) | -0.65    | -0.65    | -0.70    | -0.68    | -0.67    | -0.58    | 0.09     | -0.63    | -0.54    | 0.09     | -0.67    |
| \( \text{O}_{\text{Me}} \) | -0.46    | -0.46    | -0.48    | -0.52    | -0.51    | -0.47    | 0.04     | -0.48    | -0.44    | 0.04     | -0.51    |
| \( \text{H} \) | 0.31     | 0.31     | 0.28     | 0.29     | 0.29     | 0.32     | 0.03     | 0.31     | 0.34     | 0.03     | 0.29     |
| \( \text{N}_{\text{edge}} \) | -0.61    | -0.62    | -0.65    | -0.64    | -0.63    | -0.42    | 0.21     | -0.60    | -0.39    | 0.21     | -0.63    |
| \( \text{P}_1 \) | 0.32     | 0.31     | 0.25     | 0.24     | 0.25     | 0.34     | 0.09     | 0.34     | 0.42     | 0.08     | 0.25     |
| \( \text{P}_2 \) | 0.23     | 0.23     | 0.16     | 0.17     | 0.14     | 0.25     | 0.11     | 0.24     | 0.34     | 0.10     | 0.14     |
| \( \text{Me}_{\text{ext}} \) | 0.76     | 0.73     | 2.0(C)   | --       | 1.29     | 1.35     | 0.06     | 2.0(C)   |
| \( \text{Me}_{\text{int}} \) | 0.84     | 2.0(C)   | 2.0(C)   | --       | 1.29     | 1.35     | 0.06     | 2.0(C)   |
| \( \text{P} \) | 0.44     | 0.41     | 0.36     | 0.36     | --       | --       | --       | --       | --       | --       | --       |

3. As in system 1, but both magnesium ions were represented by point charges with \( q = +2e \).
4. With only the \( \text{P}^-\text{X}^-\text{P} \) group and the phosphate group (see Figure 2B).
5. With only the \( \text{P}^-\text{X}^-\text{P} \) group.
6. With an electron removed from system 5.
7. The same as system 5, but the nucleotide binding magnesium was included.
8. With an electron removed from system 7.
9. The same as system 7, but the magnesium ion is replaced by a point charge with \( q = +2e \).

As shown in Tables 2–6, we tabulated the charges calculated from each system and included only selected atoms to facilitate the discussion. The metal-bound oxygens show reduced charges, and the positive charges (+2) of metal ions in the present ionization state were also reduced to less than a unit charge in all cases, with the values between 0.76 and 0.84. The charge neutralization on the metal ions was studied in detail in previous publications,\(^{14,32,33}\) and we use the effect here to highlight the effect of charge redistribution if metal ions are not included in the QM systems, as was performed in some theoretical work.\(^{34}\) In the present case, including the
nucleotide binding metal with the QM system helps maintain the charge distribution of negatively charged centers similar to both being in the QM system. Note that the phosphorus at the nucleophilic site (P) had slightly reduced positive charge in all cases. However, when both cations were treated classically, larger magnitudes of charges were observed at all negatively charged atom sites, and the smaller values resulted in positively charged phosphorus centers. The charge at the nucleophilic site (P) is substantially reduced (~20%) when both metal ions were treated classically. Interestingly, this system had a rather similar charge distribution to the system that comprised only the phosphate group and the P–X–P group (i.e., comparison of columns 3 and 4), indicating the polarization effects on the wavefunction; however, a classical charge distribution alone may not be significant. The charge on the nucleophilic oxygen center (Oβα) is lowest for PPi and the highest when PNP is in the unprotonated tautomeric conformation. Thus, this may be the best among the candidates for an effective nucleophilic attack.

Within the frontier orbital approach, one can make use of Fukui functions to predict the most nucleophilic or electrophilic sites within a molecule. Change in density of atomic sites is attributed to the result of a chemical reaction, and using the Fukui functions, electron density of a molecule at a given site can be quantified when the number of electrons is changed. In the present situation, the P–X–P system shares an electron with the electrophile (the phosphorous atom of the terminal nucleic acid, that is, primer terminus). By recalculating the charges at each center, after removing an electron from the molecule, one can find the site with the greatest change in charge. We first performed these calculations on P–X–P itself, and the results are tabulated under system 5 and system 6 with the corresponding Fukui index values. We focus on the first line (corresponding to the nucleophilic oxygen that attacks the phosphate) in Tables 2–6. The charges in system 5 are rather similar to those calculated in systems 3 and 4; also, the negative charges are slightly increased compared to system 1. However, this can serve as a first approximation. In the P–O–P system, the nucleophilic oxygen (Oγβ) does not differentiate much from the other two systems within this calculation. Similar results were found for P–NH–P. A modest favorability was seen for P–CH3–P, and the highest values were observed for both the nucleophile and metal-binding oxygens for the P–N–P system. When the nucleotide binding metal was included in the calculation, system 7 showed values similar to the original system (system 1), and clear differences appeared in the Fukui index. Again, the best nucleophile was found in the case of the P–N–P system (Table 5) with a Fukui index of 0.22.

One striking feature is that P–X–P was in a slightly constrained conformation within all complexes because of metal binding. The bond distances and angles of P–X–P in the complexes and in an optimized configuration under the same level of theory are summarized in Table 7. Small compression in the bond distances in some cases was observed along with large variations in the bond angles. Inter-oxygen distances of the oxygens bound to the nucleotide metal ion were found to be reduced by more than 0.6 Å, and this introduced angular constraints in the order of 10° or more in certain cases. Therefore, if one wishes to make predictions on reactivity using indexes such as the Fukui function, the structures derived from the complex should be used in such calculations.

4. CONCLUSIONS

The study of DNA synthesis normally focuses on the chain-lengthening DNA synthesis (i.e., forward) reaction. In the reverse process, that is, pyrophosphorolysis, the addition of pyrophosphate (or analogues) is also of mechanistic interest and biological interest because one avenue for HIV-1 drug resistance is attributed to a robust pyrophosphorolysis of the reverse transcriptase.55 Recently, we have obtained detailed results leading to the remarkable finding that the reverse reaction can occur much faster with PNP than PPi (∼1100-fold) with a dramatic decrease in the overall equilibrium constant.16 Here, we have shown with predictive computational tools (QM/MM and Fukui functions) and rational structure-based models that the rapid reverse reaction is likely due, in part, to the increased nucleophilicity of the reactive oxygen on the tautomeric form of PNP. These results should be of interest to those interested in the inhibition of DNA synthesis and more generally to those interested in the contributions of tautomers to chemical reactivity.

AUTHOR INFORMATION

Corresponding Author

Samuel H. Wilson — Genome Integrity and Structural Biology Laboratory, National Institute of Environmental Health Sciences, National Institutes of Health, Research Triangle Park, North Carolina 27709-2233, United States; orcid.org/0000-0002-1702-5293, Email: wilson5@niehs.nih.gov

Authors

Lalith Perera — Genome Integrity and Structural Biology Laboratory, National Institute of Environmental Health Sciences, National Institutes of Health, Research Triangle Park, North Carolina 27709-2233, United States

William A. Beard — Genome Integrity and Structural Biology Laboratory, National Institute of Environmental Health Sciences, National Institutes of Health, Research Triangle Park, North Carolina 27709-2233, United States

Lee G. Pedersen — Genome Integrity and Structural Biology Laboratory, National Institute of Environmental Health Sciences, National Institutes of Health, Research Triangle Park, North Carolina 27709-2233, United States

David D. Shock — Genome Integrity and Structural Biology Laboratory, National Institute of Environmental Health Sciences, National Institutes of Health, Research Triangle Park, North Carolina 27709-2233, United States

Table 7. Distances (Å) and Angles (deg) of (a) Oβ–P–X–PO3 (in the Optimized Structures from the Complex) and (b) Oβ–P–X–PO3 (in Optimized Structures of the Pyrophosphate Alone in the Reaction Field of Water)

|        | Pβ–X | Pβ–X | Pβ–X–Pγ | Oβα–Oγβ |
|--------|------|------|--------|---------|
| (a)    |      |      |        |         |
| X = O  | 1.68 | 1.66 | 123.9  | 3.00    |
| X = NH | 1.70 | 1.72 | 124.0  | 3.07    |
| X = N  | 1.64 | 1.60 | 121.7  | 2.90    |
| X = CH3| 1.81 | 1.83 | 116.1  | 2.99    |
| (b)    |      |      |        |         |
| X = O  | 1.69 | 1.69 | 136.5  | 3.69    |
| X = NH | 1.76 | 1.76 | 134.0  | 3.70    |
| X = N  | 1.71 | 1.62 | 127.8  | 3.82    |
| X = CH3| 1.89 | 1.89 | 129.3  | 3.71    |
References

(1) Bebenek, K.; Kunkel, T. A. Functions of DNA polymerases. In DNA Repair and Replication; Advances in Protein Chemistry; Yang, W., Ed.; Academic Press: San Diego, 2004; pp 137-165.

(2) Deutscher, M. P.; Kornberg, A. Enzymatic synthesis of deoxyribonucleic acid. 28. Phyrophosphorylation reactions of deoxyribonucleic acid polymerase. J. Biol. Chem. 1969, 244, 3019−3028.

(3) Torgovnick, A.; Schumacher, D. B. DNA repair mechanisms in cancer development and therapy. Front. Genet. 2015, 6, 157.

(4) Bessman, M. J.; Lehman, I. P.; Simms, E. S.; Kornberg, A. Enzymatic synthesis of deoxyribonucleic acid. II. General properties of the reaction. J. Biol. Chem. 1958, 233, 171−177.

(5) Beebe, L. S.; Steitz, T. A. Structural Basis for the 3′-5′ Exonuclease Activity of Escherichia-Coli DNA-Polymerase-I - a 2 Metal-Ion Mechanism. EMBO J. 1991, 10, 25−33.

(6) Steitz, T. A.; Steitz, J. A. A General 2-Metal-Ion Mechanism for Catalytic Rna. Proc. Natl. Acad. Sci. U.S.A. 1993, 90, 6498−6502.

(7) Beard, W. A.; Wilson, S. F. Structure and Mechanism of DNA Polymerase beta. Biochemistry 2014, 53, 2768−2780.

(8) Pelleiter, H.; Sawaya, M.; Kumar, A.; Wilson, S.; Kraut, J. Structures of Tertiary Complexes of Rat DNA-Polymerase-Beta. DNA Template-Primer, and Ddctp. Science 1994, 264, 1891−1903.

(9) Steitz, T. A. Visualizing polymerase molecules at work. EMBO J. 2006, 25, 3458−3468.

(10) Lin, P.; Pedersen, L. C.; Batra, V. K.; Beard, W. A.; Wilson, S. H.; Pedersen, L. G. Energy analysis of chemistry for correct insertion by DNA polymerase beta. Proc. Natl. Acad. Sci. U.S.A. 2006, 103, 13294−13299.

(11) Sucato, C. A.; Upton, T. G.; Kashemirov, B. A.; Osuna, J.; Oertell, K.; Beard, W. A.; Wilson, S. H.; Florian, J.; Warshel, A.; McKenna, C. E.; Goodman, M. F.; Wilson, S. H. Mapping Functional Substrate-Enzyme Interactions in the pol beta Active Site through Chemical Biology: Structural Responses to Acidity Modification of Incoming dNTPs. Biochemistry 2018, 57, 3934−3944.

(12) Larsen, M.; Willett, R.; Yount, R. G. Imidodiphosphate and Pyrophosphate - Possible Biological Significance of Similar Structures. Science 1969, 166, 1510−1511.

(13) Batra, V. K.; Oertell, K.; Beard, W. A.; Kashemirov, B. A.; McKenna, C. E.; Goodman, M. F.; Wilson, S. H. Mapping Functional Substrate-Enzyme Interactions in the pol beta Active Site through Chemical Biology: Structural Responses to Acidity Modification of Incoming dNTPs. Biochemistry 2018, 57, 3934−3944.

(14) Perera, L.; Freudenthal, B. D.; Beard, W. A.; Shock, D. D.; Pedersen, L. G.; Wilson, S. H. Requirement for transient metal ions revealed through computational analysis for DNA polymerase going in reverse. Proc. Natl. Acad. Sci. U.S.A. 2015, 112, E5228−E5236.

(15) Janssen, J. A.; Beard, W. A.; Pedersen, L. C.; Shock, D. D.; Moon, A. F.; Krahn, J. M.; Bebenek, K.; Kunkel, T. A.; Wilson, S. H. Time-lapse crystallography snapshots of a double-strand break repair polymerase in action. Nat. Commun. 2017, 8, 253.

(16) Shock, D. D.; Freudenthal, B. D.; Beard, W. A.; Wilson, S. H. Modulating the DNA polymerase beta reaction equilibrium to dissect the reverse reaction. Nat. Chem. Biol. 2017, 13, 1074−1080.

(17) Mann, D.; GudelChance, J.; Schartinger, J.; Gerwert, K.; Köttting, C. The protonation states of GTP and GppNHp in Ras proteins. J. Biol. Chem. 2018, 293, 3871−3879.

(18) Reynolds, M. A.; Gerlt, J. A.; Demou, P. C.; Oppenheimer, N. J.; Kenyon, G. L. N-15 and O-17 Nmr-Studies of the Proton Binding-Sites in Imidodiphosphate, Tetraethyl Imidodiphosphate, and Adenylyl Imidodiphosphate. J. Am. Chem. Soc. 1983, 105, 6475−6481.

(19) Bax, B.; Chung, C.-W.; Edge, C. Getting the chemistry right: protonation, tautomers and the importance of H atoms in biological chemistry. Acta Crystallogr., Sect. D: Struct. Biol. 2017, 73, 131−140.

(20) Sato, H.; Wang, C.; Yamazaki, M.; Saito, K.; Uchiyama, M. Computational study on a puzzle in the biosynthetic pathway of anthocyanin: Why is an enzymatic oxidation/ reduction process required for a simple tautomeration? PLoS One 2018, 13, e0199094.

(21) Liu, C.; Rong, C.; Lu, T. Information conservation principle determines electroneutrality, nucleoelectricity, and regioelectretivity. J. Phys. Chem. A 2014, 118, 3698−3704.

(22) Wang, B.; Rong, C. Y.; Chattaraj, P. K.; Liu, S. B. A comparative study to predict regioelectretivity, electrophility and nucleoelectricity with Fukui function and Hirshfeld charge. Theor. Chem. Acc. 2019, 138, 124.

(23) Marenich, A. V.; Jerome, S. V.; Cramer, C. J.; Truhlar, D. G. Charge Model S: An Extension of Hirshfeld Population Analysis for the Accurate Description of Molecular Interactions in Gaseous and Condensed Phases. J. Chem. Theory Comput. 2012, 8, 527−541.

(24) Case, D. A.; Betz, R. M.; Cerutti, D. S.; Cheatham, T. E.; Darden, T. A.; Duke, R. E.; Giese, T. J.; Goette, A. W.; Homeyer, I.; Isadi, S.; Janovskis, P.; Kaus, J.; Kovalenko, A.; Lee, T. S.; LeGrand, S.; Li, P.; Lin, C.; Luchko, T.; Luo, B.; Madej, B.; Merz, D. K.; Merz, G.; Monroe, J. B.; Nguyen, H.; Nguyen, H. T.; Omleny, I.; Onufriev, A.; Roe, D. R.; Roitberg, A.; Saguì, C.; Simmerling, C. L.; Bellotto-Smill, W. M.; J. Swails, Walker, R. C.; Wang, J.; Wolf, R. M.; Wu, X.; Xiao, L.; Kollman, P. A. AMBER 2016. University of California: San Francisco, 2016.

(25) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. J. Chem. Phys. 1983, 79, 926−935.

(26) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Ochterski, J.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Gaussian 09; Gaussian, Inc.: Wallingford CT, 2009.

(27) Janssen, J. A.; Beard, W. A.; Pedersen, L. C.; Shock, D. D.; Moon, A. F.; Krahn, J. M.; Bebenek, K.; Kunkel, T. A.; Wilson, S. H. Time-lapse crystallography snapshots of a double-strand break repair polymerase in action. Nat. Commun. 2017, 8, 253.

(28) Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. Generalized Born Solvation Model SM12. J. Chem. Theory Comput. 2013, 9, 609−620.
(29) Duanmu, K.; Wang, W.; Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. CMSPAC version 2015; University of Minnesota: Minneapolis, MN, 2015.

(30) Freudenthal, B. D.; Beard, W. A.; Shock, D. D.; Wilson, S. H. Observing a DNA Polymerase Choose Right from Wrong. Cell 2013, 154, 157−168.

(31) Dauter, M.; Dauter, Z. Deprotonated imidodiphosphate in AMPPNP-containing protein structures. Acta Crystallogr., Sect. D: Biol. Crystallogr. 2011, 67, 1073−1075.

(32) Batra, V. K.; Perera, L.; Lin, P.; Shock, D. D.; Beard, W. A.; Pedersen, L. C.; Pedersen, L. G.; Wilson, S. H. Amino Acid Substitution in the Active Site of DNA Polymerase beta Explains the Energy Barrier of the Nucleotidyl Transfer Reaction. J. Am. Chem. Soc. 2013, 135, 8078−8088.

(33) Perera, L.; Freudenthal, B. D.; Beard, W. A.; Pedersen, L. G.; Wilson, S. H. Revealing the role of the product metal in DNA polymerase beta catalysis. Nucleic Acids Res. 2017, 45, 2736−2745.

(34) Matute, R. A.; Yoon, H.; Warshel, A. Exploring the mechanism of DNA polymerases by analyzing the effect of mutations of active site acidic groups in Polymerase beta. Proteins 2016, 84, 1644−1657.

(35) Das, K.; Martinez, S. E.; Arnold, E. Structural Insights into HIV Reverse Transcriptase Mutations Q151M and Q151M Complex That Confer Multinucleoside Drug Resistance. Antimicrob. Agents Chemother. 2017, 61, e00224.