May 2008

The effect of oxidatively damaged DNA on the active site pre-organization during nucleotide incorporation in a high fidelity polymerase from *Bacillus stearothermophilus*

Ravindra Venkatramani
University of Pennsylvania

Ravi Radhakrishnan
University of Pennsylvania, rradhak@seas.upenn.edu

Postprint version. Published in *Proteins - Structure, Function, Bioinformatics*, Volume 71, Issue 3, May 2008, pages 1360-1372.
Publisher URL: http://dx.doi.org/10.1002/prot.21824

The effect of oxidatively damaged DNA on the active site pre-organization during nucleotide incorporation in a high fidelity polymerase from *Bacillus stearothermophilus*

Abstract

We study the effect of the oxidative lesion 8-oxoguanine (8oxo**G**) on the pre-organization of the active site for DNA replication in the closed (active) state of the Bacillus Fragment (BF), a Klenow analog from *Bacillus stearothermophilus*. Our molecular dynamics and free energy simulations of explicitly solvated model ternary complexes of BF bound to correct dCTP/incorrect dATP opposite guanine (**G**) and 8oxo**G** bases in DNA suggest that the lesion introduces structural and energetic changes at the catalytic site to favor dATP insertion. Despite the formation of a stable Watson-Crick pairing in the 8oxo**G**:dCTP system, the catalytic geometry is severely distorted to possibly slow down catalysis. Indeed, our calculated free energy landscapes associated with active site pre-organization suggest additional barriers to assemble an efficient catalytic site, which need to be overcome during dCTP incorporation opposite 8oxo**G** relative to that opposite undamaged **G**. In contrast, the catalytic geometry for the Hoogsteen pairing in the 8oxo**G**:dATP system is highly organized and poised for efficient nucleotide incorporation via the "twometal- ion" catalyzed phosphoryl transfer mechanism. However, the free energy calculations suggest that the catalytic geometry during dATP incorporation opposite 8oxo**G** is considerably less plastic than that during dCTP incorporation opposite **G** despite a very similar, well organized catalytic site for both systems. A correlation analysis of the dynamics trajectories suggests the presence of significant coupling between motions of the polymerase fingers and the primary distance for nucleophilic attack (i.e., between the terminal primer O3′ and the dNTP P_α atoms) during correct dCTP incorporation opposite undamaged **G**. This coupling is shown to be disrupted during nucleotide incorporation by the polymerase with oxidatively damaged DNA/dNTP substrates. We also suggest that the lesion affects DNA interactions with key polymerase residues, thereby affecting the enzymes ability to discriminate against noncomplementary DNA/dNTP substrates. Taken together, our results provide a unified structural, energetic, and dynamic platform to rationalize experimentally observed relative nucleotide incorporation rates for correct dCTP/incorrect dATP insertion opposite an undamaged/oxidatively damaged template **G** by BF.

Keywords

active site pre-organization, oxidative damage, high fidelity polymerase, 8-oxoguanine, DNA base pair synthesis, molecular dynamics, free energy

Comments

Postprint version. Published in *Proteins - Structure, Function, Bioinformatics*, Volume 71, Issue 3, May 2008, pages 1360-1372.
Publisher URL: http://dx.doi.org/10.1002/prot.21824

In Press at Proteins: Structure, Function, Bioinformatics, 2007

The effect of oxidatively damaged DNA on the active site pre-organization during nucleotide incorporation in a high fidelity polymerase from *Bacillus stearothermophilus*

Ravindra Venkatramani and Ravi Radhakrishnan[†]
Department of Bioengineering and Department of Biochemistry & Biophysics, University of Pennsylvania, 240 Skirkanich hall, 210 S. 33rd St, Philadelphia, PA 19104

Abstract

We study the effect of the oxidative lesion 8-oxoguanine (8oxo**G**) on the pre-organization of the active site for DNA replication in the closed (active) state of the Bacillus Fragment (BF), a Klenow analog from *Bacillus stearothermophilus***.** Our molecular dynamics and free energy simulations of explicitly solvated model ternary complexes of BF bound to correct dCTP/incorrect dATP opposite guanine (**G**) and 8oxo**G** bases in DNA suggest that the lesion introduces structural and energetic changes at the catalytic site to favor dATP insertion. Despite the formation of a stable Watson-Crick pairing in the 8oxo**G**:dCTP system, the catalytic geometry is severely distorted to possibly slow down catalysis. Indeed, our calculated free energy landscapes associated with active site pre-organization suggest additional barriers to assemble an efficient catalytic site, which need to be overcome during dCTP incorporation opposite 8oxo**G** relative to that opposite undamaged **G**. In contrast, the catalytic geometry for the Hoogsteen pairing in the 8oxo**G**:dATP system is highly organized and poised for efficient nucleotide incorporation via the "two-metal-ion" catalyzed phosphoryl transfer mechanism. However, the free energy calculations suggest that the catalytic geometry during dATP incorporation opposite 8oxo**G** is considerably less plastic than that during dCTP incorporation opposite **G** despite a very similar, well organized catalytic site for both systems. A correlation analysis of the dynamics trajectories suggests the presence of significant coupling between motions of the polymerase fingers and the primary distance for nucleophilic attack (i.e., between the terminal primer O3′ and the dNTP P_α atoms) during correct dCTP incorporation opposite undamaged **G**. This coupling is shown to be disrupted during nucleotide incorporation by the polymerase with oxidatively damaged DNA/dNTP substrates. We also suggest that the lesion affects DNA interactions with key polymerase residues, thereby affecting the enzymes ability to discriminate against non-complementary DNA/dNTP substrates. Taken together, our results provide a unified structural, energetic, and dynamic platform to rationalize experimentally observed relative nucleotide incorporation rates for correct dCTP/incorrect dATP insertion opposite an undamaged/oxidatively damaged template **G** by BF.

Key words: active site pre-organization, oxidative damage, high fidelity polymerase, 8-oxoguanine, DNA base pair synthesis, molecular dynamics, free energy

[†] Corresponding author email: rradhak@seas.upenn.edu

1. Introduction

DNA polymerases from family A, which include *Escherichia coli* polymerase (pol) I, T7 DNA pol I, and pol I from bacteria such as *Thermus aquaticus* and *Bacillus stearothermophilus,* specialize in the replication/repair of DNA. Shaped like a human right hand, with thumb, finger and palm domains, these enzymes bind to DNA to incorporate nucleotides selectively (dCTP opposite **G**, dATP opposite **T**, dTTP opposite **A,** and dGTP opposite **C**) via dexterous motions of their finger domains opposite unpaired template bases. Family A polymerases are typically "high fidelity" enzymes, exhibiting extremely small error rates (1 error in 10^6-10^8 dNTP incorporations[1]) during the DNA replication/repair cycle. The accuracy with which these enzymes function is crucial for preserving genomic integrity and failures often lead to severe biomedical implications [2-4].

The 8-oxoguanine (8oxo**G**) lesion is a prominent product generated by oxidative damage of DNA [2,5,6] . Propensities for the lesion to form 8oxo**G**:**A** mismatches[5-8] during DNA replication result in **G**:**C** → **T**:**A** transversion mutations [4,9,10], which are correlated with many types of cancers [10-12]. The rate of dATP (mis)incorporation opposite this lesion increases, in some cases, to an extent comparable to or greater than that of correct dCTP incorporation. Crystallographic captures of high fidelity polymerase/DNA post-insertion complexes [5,6,13] have suggested the structural origin for this altered behavior. That is, post-insertion complexes with the 8oxo**G:C** *anti:anti* base pair show large distortions in the template strand, presumably due to steric clashes between the O8 oxygen atom of the nucleotide base and the O4′ oxygen atom of the deoxy-ribose sugar of the lesion. In contrast, adopting a less conventional *syn* conformation and thus forming a stable Hoogsteen base pairing with an incoming dATP, enables the 8oxo**G:A** purine-purine mispair to fit snugly into the active site, without introducing severe template strand distortions. While the structural inferences have significantly enhanced our understanding of the polymerase behavior in the presence of 8oxo**G**, it is difficult to assess the effect of the lesion on the nucleotide incorporation cycle purely on the basis of post-/pre-insertion structures alone. We therefore employ molecular modeling and simulations as complimentary tools to address this central issue. In the past, such modeling studies have been highly valuable in understanding polymerase fidelity mechanisms and catalysis in T7 DNA pol I [14,15], in studying the effects of DNA lesions at the active site of various replicative and repair polymerases [16-21], and in identifying and demonstrating the relevance of substrate specific conformational changes and structural rearrangements to polymerase fidelity in human pol β [22-24].

During nucleotide incorporation by polymerases, the basic reaction in the chemical step is the formation of a covalent phosphodiester bond between the incoming dNTP and the terminal primer base of the polymerase-bound DNA, prior to which the enzyme has to switch to an "active" state, usually through a large scale conformational rearrangement of the fingers (or the thumb) domain [25]. Subsequent to the chemical step, the enzyme usually switches back to an "inactive" state and the cycle repeats. Based on a large body of structural and biochemical works, the identity of the rate-limiting step in the nucleotide incorporation cycle depends not only on the type of polymerase, but also on the base pairs involved in the synthesis [25]. During correct incorporation, either the conformational change preceding the chemical step (*E. Coli* pol I [26]) or the chemical step itself (T4 Pol I [27]) is found to be rate-limiting, while for incorrect incorporation it is believed (largely based on indirect evidence across several polymerases) that the chemical step is rate-limiting.[25‡] In this study we compare and contrast the structural, energetic, and dynamic characteristics of four bacillus fragment (BF) polymerase/DNA/dNTP ternary complexes focusing on the active site pre-organization prior to the chemical step along the nucleotide incorporation pathway. The term active site pre-organization here refers to the assembly of a chemically potent ground state catalytic geometry for efficient phosphodiester bond formation. Such an organization can either proceed concurrently with, or follow the conformational change step.

We perform classical molecular dynamics simulations, free energy calculations, and correlation

‡ It is instructive to compare the energy barriers for the experimentally resolved conformational change and chemical steps in a specific polymerase system. Experiments on the Klenow fragment (Eger et. al. Biochemistry, 38, 9227) yield a free energy barrier of 15.2 kcal/mol for the conformational change, and 14.4 kcal/mol (upper-bound) for the chemical step during correct nucleotide incorporation. For incorrect nucleotide incorporation, the corresponding values reported are 15.2 and 19.8 kcal/mol, respectively.

analysis of fully atomistic, explicitly solvated models of the large fragment from *Bacillus stearothermophilus* (derived from X-ray crystal structure[28]**;** PDB id: 1LV5) for correct(dCTP)/incorrect(dATP) nucleotide insertion opposite undamaged **G** (control), and oxidatively damaged (8oxo**G**) template bases in the polymerase-bound DNA. Consistent with post-insertion structures from X-ray crystallography studies, we find that the unconventional *syn* conformation of the 8oxo**G** template base opposite an incoming dATP (which we refer to as the 8oxo**G**:d**A**TP system) leads to a distortion free catalytic site that is poised for a two-metal-ion catalyzed phosphoryl transfer, (see Section 3.1). Such a distortion free catalytic site also occurs for the undamaged Watson-Crick **G:**dCTP system. In contrast, the catalytic site geometry for the 8oxo**G:**dCTP mispair is highly distorted, despite the formation of a regular Watson-Crick base pair. Based on the energy barriers derived from our free energy simulations, we suggest that the rates for catalytic site pre-organization follow the trend **G:**dCTP > 8oxo**G:**dCTP >> **G:**dATP. Moreover, even though the catalytic site for the 8oxo**G:**dATP system is pre-organized to the same degree as the **G**:dCTP control system, its deformation characteristics are found to be considerably less plastic than the control, suggesting the trend for rates of active site assembly to be: **G:**dCTP > 8oxo**G:**dATP > 8oxo**G:**dCTP >> **G:**dATP, which correlates perfectly with the observed trend in the rates of nucleotide incorporation (k_{cat}/K_m) in kinetic studies [6]. We also find on the basis of a correlation analysis of active site motions of our control **G**:d**C**TP system that a dynamical coupling exists between the polymerase fingers domains and the reactive distances for phosphoryl transfer at the catalytic site. Such a coupling insinuates the role played by the polymerase fingers in the catalytic site pre-organization. Intriguingly, this dynamical coupling is disrupted when an oxidative damage is present, suggesting a higher free energy barrier for catalytic site pre-organization. Based on structural and energetic grounds, we quantify the involvement of key polymerase residues, namely, R615, Y714, Q797, and H829 identified in prior mutagenesis studies, in discriminating between correct and incorrect nucleotides at the active site, and show that the oxidative damage significantly diminishes this ability to discriminate. Collectively, these conclusions help provide a molecular basis for the altered specificity of the BF pol when it encounters an oxidative damage in the DNA.

2. Methods

We present a brief overview of our methods in this section. A more comprehensive description is provided as supplementary material.

2.1 System preparation and simulation protocols

Structures of the Bacillus fragment (BF) have been characterized extensively in the pioneering studies of Beese and co-workers at various stages of DNA replication [13,28-30], as well as in the presence of various lesions [6,31,32]. Starting from the crystal structure of a closed (active) ternary BF-DNA-dCTP complex (PDB id: 1LV5 [28]), we construct four model systems of explicitly solvated and neutral BF-DNA-dNTP ternary complexes, (see Figure S1 in supplementary material): (A) **G:**d**C**TP, (B) **G:**d**A**TP, (c) 8oxo**G:**d**C**TP, and (d) 8oxo**G:**d**A**TP. All structural modifications are carried out using the Insight II modeling software [33]. The NAMD simulation package [34,35] with the CHARMM27 [36] force field is used to minimize and equilibrate each model system. Parameters (i.e., partial charges and force constants) compatible with CHARMM27 for the 8oxo**G** residue are constructed according to the procedure described by Foloppe et. al. [37] using a genetic algorithm-based optimization scheme (Y. Liu, R. Radhakrishnan, unpublished). Following standard equilibration protocols, 10 ns constant number of atoms /volume/temperature (NVT) molecular dynamics (MD) production runs are carried out and data from the last 5 ns (during which the root-mean-squared deviations or RMSD of the protein backbone are found to be stable, see Figure S3 in supplementary material) are used for subsequent analysis. A listing of catalytic site parameters for the four models along with a comparison to experimental data is presented in Table S4 of the supplementary material.

2.2 Free energy simulations for pre-organization of catalytic sites

We perform umbrella sampling simulations in two-dimensions using the CHARMM program

(version c32a1) [38] by varying the distance between the O3′ of the DNA terminal primer and P_α of the incoming nucleotide (denoted here as the distance d_a) and the distance between O3′ and the catalytic Mg^{2+} (denoted here as distance d_b). The geometry of the time-averaged structure for the **G:**d**C**TP control system emerging from the simulation trajectory shows a catalytic site which only differs from the ideal two-metal-ion geometry postulated by experimental and theoretical studies in the values of these two distances (see section 3.1). Further, we postulate based on prior experience with pol β [39,40] and the hammerhead ribozyme, [41] that the pentacovalent transition state for phosphodiester bond formation can be realized by applying constraints on d_a and d_b, alone. We thus vary d_a and d_b in steps of 0.5 Å each, over a range of values $d_a^{max} - d_a^{min}$ and $d_b^{max} - d_b^{min}$, appropriately chosen to sample conformations of the active-site in the vicinity of the time-averaged reference structure. The value of (d_a^{max}, d_b^{max}) in units of Å is set to be (4.0,4.0), (5.0,5.0), (5.0,4.0), (4.0,4.0), for the **G:**d**C**TP (or **G:C**), **G:A**, 8oxo**G:C,** and 8oxo**G:A** systems, respectively, while the value of $d_a^{min} = d_b^{min}$ is set as 2.0 Å, for all four systems. At each grid point we perform two rounds of energy minimization using the steepest-descent method followed by Langevin dynamics (with friction coefficient γ=10 ps^{-1}) at 300 K with a 1 fs timestep. The first round of energy minimization, which brings the system to the desired grid point, is performed with a high value of forcing restraints (2000 kcal/mol/Å^2) and consists of 1000 steps of minimization and 1 ps of dynamics. The second round (umbrella sampling) is performed with a lower value of the forcing restraint (20 kcal/mol/Å^2), and consists of a 1000 steps of energy minimization and 10 ps of Langevin dynamics. We thus obtain 25, 49, 35 and 25 umbrella sampling windows for the **G:C**, **G:A**, 8oxo**G:C** and 8oxo**G:A** systems, respectively. Data from these windows are then used to construct unbiased probability distributions and free energy surfaces by employing the WHAM [42,43] algorithm. The error in free energies is estimated by calculating the standard deviations obtained from three independent umbrella sampling simulations performed for the **G:C** system and found to be ± 0.9 k_BT (0.54 kcal/mol). Since identical conditions are used to collect data, we expect the errors to be of the same order for the **G:A**, 8oxo**G:C,** and 8oxo**G:A** systems.

2.3 Correlation analysis of the active site

In order to assess the effect of slower coupled motions of the DNA/polymerase/dNTP complex on the dynamics of the primary reactive distances in the presence/absence of the 8oxo**G** lesion, we performed a correlation analysis of the active site region (see Figure 1a). Our active-site is defined to include all heavy atoms of the incoming dNTP, seven residues of the DNA template strand (T_{n+2} , T_{n+1} , T_n , T_{n-1} , T_{n-2} , T_{n-3} ,and T_{n-4} , where, the index n refers to the **G**/8oxo**G** of the nascent base pair, indices n+2 and n+1, refer to residues in the single stranded template overhang, and indices n-1, n-2, etc., refer to bases in the DNA duplex), four residues of the DNA primer strand, (P_{n-1} , P_{n-2} , P_{n-3}, and P_{n-4}, where, the index n-1 refers to the terminal primer **A** which is part of the DNA duplex and indices n-2, n-3 and n-4 refer to residues further down in the duplex), the two Mg^{2+} ions, the highly conserved catalytic aspartates residues, (D830 and D653), of the polymerase, polymerase residues from the O and O1 helices, (residue numbers 696 through 726), in all 30 residues as depicted in Figure 1a forming the polymerase fingers, and three polymerase residues contacting the first three DNA base pairs at the active site, (A615, Q797, H829), along with Y714 at the base of the O helix. Using the geometries of the ternary complexes representing the structural average over the last 5 ns of our MD trajectories as references, variance-covariance matrices of atomic fluctuations in the Cartesian system of coordinates are constructed for the active site region of the 4 systems. The software program CARMA[44] is used to perform Principal Component Analysis (PCA) [45,46], to project/visualize the independent atomic motions in an MD trajectory, and sort them in the descending order of their variance. The top 10 principal component modes capture the bulk of the atomic fluctuations in our MD trajectories for all four systems studied, (namely, 70% for **G:C**, 84% for **G:A**, 72 % for 8oxo**G:C,** and 80% for 8oxo**G:A**).

3. Results and Discussion

3.1 Catalytic ground state geometry in BF ternary complex is optimized to favor dCTP insertion opposite an undamaged template G

In the "two-metal-ion" mechanism for phosphoryl transfer [47-49] believed to be applicable for a wide range of enzymes [50,51], the polymerase catalyzes the phospho-diester bond formation by coordinating two divalent metal ions at the catalytic site with the aid of highly conserved (usually aspartate) residues. The catalytic ion is believed to activate the 3'–hydroxyl group of the terminal primer to facilitate the formation of an O3′-P_α covalent bond. The nucleotidyl ion is thought to screen the negative charge in the triphosphate moiety of the incoming nucleotide and hence to facilitate the formation of the pentacovalent transition state and the subsequent cleavage of the $O3_\alpha$-P_α bond[15,51-55] [§]. Conformations close to the pentacovalent transition state as proposed by the two-metal-ion mechanism have recently been captured [56,57] in structural studies of other biomolecules; we consider the geometry obtained in the studies of Stahley and Strobel [57] as a reference for the "ideal-two-metal ion geometry", (see Figure 1c for a depiction, Figure 2 gives values for these distances from the studies of Stahley and Strobel [57]). In addition to these structures, computational delineations of transition state structures and reaction intermediates for T7 DNA pol [15], phosphoglucomutase [58], polymerase β [55,59], and the hammerhead ribozyme [41] have also provided us with detailed structural insights on the phosphoryl transfer pathway, including a characterization of the transition state geometry.

Since the input structure to our dynamics simulations of the **G:C** system is derived from crystal structures, this system serves as our control. Our simulations show (see Figure 3a) a tightly organized catalytic site with two Mg^{2+} ions (nucleotidyl MG1 and catalytic MG2), two aspartate residues (D831 and D653) highly conserved amongst high fidelity polymerases, and two bound water molecules. The coordination sphere of the catalytic Mg^{2+} (MG2) is completed by six oxygen ligands [55]: one each from D831 and D653, two from water molecules, one from the terminal primer (O3′), and the one from the non-bridging oxygen of P_α in the dCTP triphosphate moiety. The time-averaged structure of the ground state resulting from our classical simulations fulfills all but two requirements (see Figure 2, black line) for conforming to the "ideal-two-metal-ion" geometry: (1) the average O3′-P_α distance of 3.5 Å is greater than the 2.0 Å expected at the transition state for an associative or partially associative mechanism [56,60]; (2) the terminal primer hydroxyl remains protonated due to a large average O3′-MG2 distance of 2.7 Å . However, in our simulations, large (and simultaneous) fluctuations of both O3′-MG2 and O3′-P_α distances bring them close to the values of 2.1 Å and 3.2 Å, respectively; these fluctuations represent the correct trend for forming a reaction-competent two-metal-ion geometry [57]. Moreover, the observed proximity of the terminal primer hydroxyl group to the catalytic D830 and to the dNTP provide the possibility of multiple avenues for this initial deprotonation event [15,55]. In order to explore the energetics of transforming the ground state geometry of the **G:C** system to the ideal two-metal-ion geometry [49,53,56,57,61] for an associative transition state, we compute the free energy landscape by spanning the key reactive distances for catalysis O3′-P_α (or d_a) and the O3′-MG2 (or d_b). The energy profile (see Figure 3c) is described by a single well-characterized minimum in the vicinity of the simulation average (X). Moreover, for regions surrounding this minimum, the equal energy contours are elliptical with their major axes (which is also the principal direction of motion in the sub-space of the free energy landscape) aligned with slope~1, i.e., a direction corresponding to simultaneous increase/decrease of the reactive distances d_a and d_b. Moreover, structural analysis of conformations sampled in the vicinity of the minimum shows that other key interactions required for the two-metal-ion catalyzed reaction mechanism are preserved, (Figure 4, black symbols; see also Figure S5 in supplementary material). Thus, d_a and d_b can be transformed (in a correlated fashion) to be made closer to their respective values in the ideal two-metal-ion geometry without disrupting other key interactions required to stabilize the active site. The active site degrees of freedom, therefore, undergo plastic deformations, and the system appears primed for an inline attack, pending the deprotonation of the terminal primer hydroxyl group.

In contrast, in our results for the **G:dA**TP mismatch (i.e., dATP incorporation opposite **G** in the BF pol ternary complex) we observe a distorted catalytic site (Figure 3b) with significant increases in the

[§] Studies on nonenzymatic phosphoryl transfer reactions show that the reaction can proceed either through a dissociative (via a metaphosphate-like intermediate) or by an associative (via a pentacovalent intermediate) mechanism. In the case of some ribozymes and polymerases, theoretical evidence as well as structural arguments have suggested a partially or fully associative mechanism.

O3′- P_α and O3′-MG2 distances relative to those observed in the **G:C** control system. We attribute this crucial difference primarily to be due to an increase in the average base pair width by ~2 Å for the **G:A** purine-purine mismatch (this feature is also observed in the post-insertion crystal structures[13]), which misaligns the incoming dATP with respect to the primer strand and disrupts the key MG2-O3′ metal-ligand interaction. Further, we observe that in order to compensate for this loss of interaction, the **G:A** system allows an extra water molecule to associate with MG2. A summary of the relative distances that describe the relative alignments in the catalytic site is provided in Figure 2 (red line), which relative to our control, is far from optimum. The free energy landscape (Figure 3d) reveals a single minimum around the conformation resulting from the unconstrained MD simulation average (X). For regions surrounding this minimum, the major axes of the elliptical equal energy contours have a slope >>1, which signifies a restricted exploration of d_b as d_a changes. This causes a sharp increase in the free energy as we approach the ideal two-metal-ion geometry. Since the **G:A** system does not show a local minimum close to the optimal ground state geometry, using the location of the free energy minimum for control **G:C** as a reference (i.e., d_a^r= 3.51 Å and d_b^r= 2.66 Å), we estimate the free energy cost associated with reducing d_a and d_b for the **G:A** from their respective values at the free energy minimum (d_a= 4.74 Å and d_b= 4.72 Å) to the reference values (d_a^r,d_b^r) to be 30.7 k_BT (18.6 kcal/mol). This value is larger than the conformational change barrier (15.2 kcal/mol) reported in studies of mispair formation in *E. Coli* Klenow 26 and smaller than the barrier for the chemical step (19.8 kcal/mol). We note that this energetic estimate for the active site pre-organization is subject to the approximations of our classical force-field. Nevertheless, it provides an avenue to track structural changes in the active site at atomic detail and link them to energetic changes. Namely, the steric constraints imposed by the larger width of the nascent **G:A** basepair imply that a two-metal-ion geometry (i.e., a reduction in d_a and d_b, see below) can only be attained by inducing additional distortions at the template **G** and in the polymerase residues contacting the nascent basepair, namely, R702, K706 and Y714 of the O-helix and mismatch sensing R615 and Q797 , and thus at a high energy cost.

3.2 Catalytic ground state geometry is optimized to favor dATP incorporation opposite 8-oxoguanine

One striking feature of the structures resulting from our simulations is that, relative to the control (**G:C**), the catalytic site for correct dCTP incorporation opposite the lesion system (8oxo**G:C**) is distorted despite the formation of a stable Watson-Crick base pair (see Figure 5a). Therefore, the polymerase active site conforms to the two-metal-ion geometry only partially, with the structural distortions resembling those observed for the **G:A** system; in particular, average values of the O3′-P_α (4.50 Å) and O3′-MG2 (3.77 Å) are far from the two-metal-ion reference geometry. While the locations of the dominant minimum in the free energy landscape for the **G:A** and 8oxo**G:C** systems are similar, the free energy landscape for the 8oxo**G:C** system (see Figure 5c) shows a local minimum at d_a=3.4 Å and d_b=2.3 Å. Significantly, location of this local minimum is close to that of the dominant minimum for the **G:C** control system. The calculated value of the active site pre-organization energy, i.e., the free energy penalty associated with transforming the 8oxo**G:C** system from its ground state (d_a=4.50 Å and d_b=3.77 Å) to a state resembling the active site geometry for the **G:C** control (d_a=3.4 Å and d_b=2.3 Å) is significantly lower (7.86 kcal/mol or 11.5 k_BT), compared to that estimated for the **G:A** system (18.6 kcal/mol or 30.7 k_BT). Therefore, based on this comparison of the energetic barrier for active site pre-organization, we conclude that the rate of dCTP incorporation opposite 8oxo**G** will be more efficient than dATP incorporation opposite a template **G**. We assume that the rate is related to the activation free energy barrier ΔG using the transition state theory relationship, namely, rate=$(1/\tau)\times\exp(\Delta G/k_BT)$, where τ is a characteristic molecular relaxation time at the transition state. A relative comparison of these barriers with those associated with conformational change or chemical steps is not possible at present as experiments have not yet resolved these steps for oxidatively damaged substrates.

In contrast to the 8oxo**G**:**C** system, the Hoogsteen pairing observed in the 8oxo**G:A** system precludes any catalytic site distortions (see Figure 5b). In fact, the alignments in the catalytic site are optimal and similar to those observed in the **G:C** control. Figure 2 (green line) shows that the key interactions required for a two-metal-ion mechanism are preserved making 8oxo**G:A** mispair formation

more efficient than correct 8oxo**G:C** base pair formation. Based on the geometry of the catalytic site alone dATP incorporation opposite the lesion is expected to be as efficient as dCTP incorporation opposite **G**. However, an important distinguishing feature of the 8oxo**G:A** system from the other three systems lies in the lack of plasticity associated with its active site motions, which is revealed by our umbrella sampling simulations. In contrast to the G:C case (Figure 3c), the free energy landscape for 8oxo**G:A** (Figure 5d) exhibits multiple conformations in the vicinity of the simulation average (X); moreover, in Figure 4 (also see Figure S5 in supplementary material), obvious differences in the variations of metal-ligand interactions at the catalytic site are evident for the two systems: namely, while the **G:C** system is deformable in the limited subspace of d_a and d_b, with other key parameters close to ideal two-metal-ion values, the 8oxo**G:A** system shows a less orchestrated response as a result of a loss of active-site plasticity and several catalytic MG2 metal-ligand interactions (MG2-D830:$O1_\delta$, MG2-dNTP:$O2_\alpha$, MG2-bound waters) are disrupted. While both the **G:A** and 8oxo**G:C** systems also show similar disruptions in metal-ligand interactions for large values of d_a and d_b, they show a recovery of the plastic response as d_a and d_b are constrained to approach their ideal two-metal-ion values; the onset of such a recovery of the plastic behavior is not observed for the 8oxo**G:A** system (see Figure S5 in supplementary material). Thus, we conclude that the reaction landscape for active site pre-organization in the 8oxo**G:A** is more complex than those delineated for the **G:C**, **G:A**, and 8oxo**G:C** systems, and may not be best represented in terms of just d_a and d_b. In summary, while the absence of catalytic site distortions explains why 8oxo**G:A** mispair formation is preferred over 8oxo**G:C** in BF replication reactions, other factors such as the loss of plasticity at the catalytic site lower the incorporation efficiency in comparison to the **G:C** system.

3.3 Role of bound catalytic site waters

An analysis of structures and catalytic site distances from our free energy simulations shows three water molecules liganded to the catalytic Mg^{2+} for the **G:A** and 8oxo**G:C** systems where the catalytic geometries are far from the ideal two-metal-ion geometry. For these systems one of these water molecules serves as a surrogate to the disrupted MG2-O3′ metal-ligand interactions. Its presence also weakens MG2 metal-ligand interactions with the non-bridging oxygen ($O1_\alpha$/$O2_\alpha$) in the triphosphate moiety of the incoming nucleotide, (in Figure 2, these interactions are completely disrupted for the 8oxo**G:C** system). The loss of such critical interactions in these two systems not only hampers the assembly of the two metal-ion geometry, but also leads to a weakening of correlations between changes of the reactive distances d_a and d_b (Figures 3 and 5). We note that slopes $<<1$ or $>>1$ for the major axis of the energy contours are suggestive of a loss of correlation. While the **G:A** free energy landscape (see Figure 3d) shows a weaker correlation between changes in these coordinates, the correlation disappears in landscapes for 8oxo**G:C** system in regions where the extra water molecule (The optimally organized catalytic site of the control **G:C** shows only two water molecules) is bound (see Figures 5c). However, as the values of the coordinates d_a and d_b are reduced to approach their ideal two-metal-ion values, one of the waters is expelled and the coordination sphere of the catalytic Mg^{2+} becomes similar to that of the **G:C** control. Notably, the correlation in the reactive distances is also restored. We note that while a quantitative estimation of the free energy barrier for water expulsion would be very insightful, such a calculation is non-trivial because of the many-body nature of the appropriate reaction coordinate for water expulsion. Hence, such a calculation may be beyond the scope of the umbrella sampling method we have employed here. Nevertheless, our results suggest that such a barrier for water expulsion from the active-site may be significant because of its bearing on the plasticity of active-site deformations.

3.4 A dynamical coupling between the polymerase fingers and catalytic site reactive distances aids the pre-organization toward a two-metal-ion geometry

A correlation analysis of the dynamics trajectories (see section 2.3) reveals that for the **G:C** control system, motions in the polymerase fingers are highly cooperative with motions of the nascent **G:**dCTP base pair (see correlation plot in Figure S6 and Table S7 in supplementary material). Strong correlations also exist between each base pair in the DNA duplex as well as between successive bases within the template and primer strands. A visualization of the top 10 modes from a principal component analysis

(section 2.3) for the control **G:C** case shows motions of the polymerase fingers causing delocalized (large scale) template-primer strand distortions (similar to writhing and stretching modes of the DNA) and synchronized with fluctuations in the reactive distance d_a, (see Movie S8). The plastic response of the catalytic geometry (comprising of D830, D653, the two Mg^{2+} ions and bound waters) to these dominant modes ensures that an optimal active site is maintained. This leads us to conclude that DNA primer and template strands act as a relay to transmit these cooperative fluctuations and the polymerase drives the fluctuations of the reactive distances d_a and d_b using a mechanical coupling existing between the O helices and DNA template and primer strands. The picture that emerges is that specific modes (delocalized over a larger spatial domain) of the system couple to the fluctuations of catalytic reactive distances to possibly drive the system towards the transition state for phosphoryl transfer. This picture is bolstered by the accumulating evidence in other enzyme complexes of a pre-chemistry phase, in which delocalized (global) and local motions driven by thermal fluctuations of the enzyme-substrate complex synergistically orchestrate the assembly of an optimal catalytic site [39,62-65] . We note that the terminology "dynamic coupling" is introduced here to signify the coupling between fast (local) and slow (delocalized) modes. In our view, the coupling (when present) is significant in altering the free energy landscape. We do not suggest that the coupling necessarily alters the temporal relaxation dynamics. This interpretation is consistent with the discussion in a recent review article on this topic [66]. Our observations are also consistent with the response of these systems to external forces as measured in single molecule studies: in these studies, the replication rates of two high fidelity A family polymerases were shown to be significantly affected by the application of a stretching force on the template strand of the DNA substrate [67-69]. We have shown elsewhere that a force-response of the rate of phosphodiester bond formation arises possibly as a direct result of the dynamical coupling phenomena [70].

For the **G:A** system, (see Table S7 and Figure S9), we observe that a significant number of the correlations prevail, and are even enhanced relative to **G:C**. Notably, however, the correlations between the O-helix and dATP are weaker due to the increased purine-purine width, and consequently, the correlations between the O3' and P_α atoms involved in the nucleophilic attack are also weakened. As a result, the dynamical coupling does not drive fluctuations in the reactive distance in the **G:A** system.

3.5 Oxidative damage weakens the dynamical coupling

For the two systems containing oxidative lesions, many correlations between the polymerase fingers and the nascent base pair (8oxo**G:**dNTP) are significantly weakened relative to the control, (see Table S7, Figures S10 and S11). While correlations within the bound DNA fragment as well those within the catalytic site are maintained, we record reductions in correlations between O-helix and dNTP for 8oxo**G:A**, and between O1-helix and T_n for both 8oxo**G:C** and 8oxo**G:A** systems. These variations produce a reduction in correlation between the terminal primer O3′ and the dNTP P_α,. Even though the top 10 principal component modes for 8oxo**G:C** and 8oxo**G:A** show large DNA strand distortions, a visual comparison with those of the **G:C** system suggests that, in stark contrast to the control, the DNA strand distortions are localized to the single stranded template overhang and the lesion and do not propagate to the catalytic site (see Movie S8). The loss of dynamic coupling between the polymerase fingers and the active-site due to the presence of the oxidative lesion exerts a negative influence on the active-site plasticity and accentuates the important role played by the fingers in the active-site pre-organization. This notion is consistent with previous studies, which have also emphasized the various roles played by the fingers domain in high fidelity polymerases: (1) The induced-fit conformational change in high-fidelity polymerases involves the fingers (or thumb for pol β) domain and plays a discriminatory role [71-73] in nucleotide incorporation by producing a stable active conformation only when the correct nucleotide is bound in the active-site. (2) The role played by Y714 (Y766 in Klenow) located at the base of the O-helix of the polymerase fingers in the pre-chemistry conformational change by limiting the template base to access the active-site only in the closed or active conformation of the polymerase is well appreciated from prior crystallographic studies [74].(3) In their studies on BF, Johnson et. al.[28] characterize a "pre-insertion site" which is distinct from the insertion site (the active site where the incoming nucleotide is paired opposite a template base) and where the loop connecting the O and O1-helices sequesters an unpaired template base in the active (open) state of the enzyme. As the polymerase

switches to its closed (active) state through the conformational change involving the fingers and the thumb subdomains, the template base is moved to the insertion site while the loop blocks the pre-insertion site preventing access to other template bases in the single stranded DNA overhang . The latter two roles of the polymerase fingers are instrumental in preventing frameshift mismatch errors by the polymerase during DNA synthesis [28].

3.6 Effect of oxidative damage on DNA-polymerase interactions implicated in enzyme fidelity

Mutagenesis experiments on pol I homologues of BF (*E. Coli*, *Taq* and *T7* pol I) reveal that mutations of four highly conserved residues R615, Y714, Q797, and H829, (R668, Y766, Q849, and H881, respectively for *E. Coli* Klenow fragment), significantly affect fidelity[75,76]. While mutation of residues R615 and Y714 lead to a decrease, mutations of Q797 and H829 show an increase in replication fidelity. Three of these residues, R615, Q797, and H829, are part of a "H-bonded-track", and have been shown to participate in the catalytic activity of the polymerase,[77] with a mutation of any one of these residues greatly reducing polymerase-DNA binding. Experiments have proposed a role for a fourth residue, Y714, in stabilizing the template strand base paired opposite the incoming nucleotide in the closed (active) state of the polymerase[75]. In Table I, we list the interactions of these four residues with the first four DNA base pairs including the nascent pair (**G**:dCTP) for the **G:C** system as inferred from prior experiments, (and confirmed by an analysis of structures obtained from our classical simulations). Residues R615, Q797, and H881 contact the DNA minor groove acceptors (N3 for purines and O2 for pyrimidines) in the template (Q797) and primer (R615, H829) strands, respectively, to influence fidelity at the mismatch extension stage [75-78] . In addition, R615 and Y714 contact the nascent base pair directly to influence fidelity at the nucleotide insertion stage.

We have calculated the non-bonded (van der Waals and electrostatic) interaction energies of each of these four residues with the first four bases in the DNA fragment and the incoming dNTP for all four model systems **G:C**, **G:A**, 8oxo**G:C**, and 8oxo**G:A**, to assess the effect of the lesion on key interactions. In Table II, we present only the strongest pair-wise (residue-DNA) interactions averaged over the last 2ns of our MD trajectories along with their standard deviations. The strongest interactions are all negative signifying an attractive complementarity between the polymerase active site and the DNA+dNTP substrate for all four model systems. For the **G:C** control the interactions are strongest around the nascent base pair and decrease as we move further down in the DNA duplex. Residues R615 and H829 show strong interactions with pairs of bases along the primer strand, which emphasizes their dual role in aligning the primer and dNTP for efficient catalysis as well as in extending newly formed base pairs[77,78]. On the other hand residues Y714 and Q797 show one-on-one interactions with the template strand consistent with their proposed functions of either stabilizing (Y714) or extending base pairs (Q797)[75,76]. Table II also shows that the presence of a mispair (**G:A**) at the active site weakens almost all interactions with the strongest effect on the residues which interact with the primer strand. The interactions of R615 with the incoming dNTP, in particular, as well as interactions of R615 and H829 with the terminal primer **A** are significantly weakened to provide a greater degree of conformational flexibility for the mispair and primer at the catalytic site. However, the flexibility is limited as polymerase retains its contacts with the dATP and the terminal primer. Due to the increased purine-purine width R615 forms a hydrogen bond with the dATP at the N3 base position instead of the regular O4′ sugar position which might prove to be vital in stabilizing the mispair in a distorted catalytic geometry to slow down catalysis[78]. On the template side, the interactions of Y714 are also reduced and this further decreases the stability of the **G:A** wobble base pair. Q797 and H829 interactions with the DNA minor groove are not affected, with the latter actually showing a significant increase. These observations are consistent with the roles assigned to the four residues by experiments [75-78]. Notably, the residues implicated in sensing during the nucleotide insertion stage, namely R615 and Y714, are affected by the presence of a mispair at the active site highlighted by the significantly weakened interaction with the DNA fragment.

For the systems with oxidative damage there is minimal change in the interactions of R615 with the incoming nucleotide (for both dCTP and dATP) and the terminal primer base relative to that observed for the **G:C** control. Further, the reduced width of the 8oxo**G:A** Hoogsteen pair allows R615 to contact the dNTP at the regular O4′ position as seen for the regular Watson-Crick 8oxo**G:C** and **G:C** systems. However, relative to the control, the interactions of Y714 with 8oxo**G**, H829 with the terminal

primer base and Q797 with the minor groove are all significantly weaker. While the first two destabilize the active site during nucleotide insertion, the third would likely affect the extension of the lesion. The degree of weakening of interactions is smaller for the 8oxo**G:A** system, which in our view could rationalize why the 8oxo**G:A** system is stable at the well-aligned catalytic geometry, whereas the 8oxo**G:C** system is only metastable.

The presence of these DNA polymerase interactions running as a relay parallel to the helix axis along the template and primer strands, coupled with stacking interactions between the adjacent base pairs, provide a molecular basis for the dynamical coupling between the polymerase fingers and the catalytic site observed for the **G:C** control (see section 3.4 and Movie S8). In this picture, the weakening of Y714-8oxo**G**(T_n) interactions lead to a loss of correlations between the polymerase fingers and the DNA template strand for systems with oxidative damage. Moreover, for these systems, there are significant increases in the interactions of Y714 with the incoming nucleotide and a weakening of Q797 and H829 interactions with the DNA (Table II), which leads to a loss of correlation between template and primer strand motions. In our view, these factors collectively lead to the observed disruption in the dynamical coupling for systems with oxidative damage

4. Conclusions

We have identified and quantified structural and energetic changes during active-site pre-organization in the context of correct and incorrect nucleotides bound to the active BF polymerase/DNA complex in the presence and absence of the oxidative lesion 8oxo**G** in the DNA template base at the active-site. Our calculated free energy barriers for active site pre-organization suggest that the rate of assembling an efficient two-metal-ion geometry (catalytic site pre-organization) follows the trend: **G:C** > 8oxo**G:C** >> **G:A.** Moreover, the stability of well organized two-metal-ion geometry for the 8oxo**G:A** system suggests the trend: 8oxo**G:A** > 8oxo**G:C**. We have also identified a dynamical coupling between the polymerase fingers and catalytic site degrees of freedom aiding catalytic site pre-organization in the **G:C** control system which underscores their role in the assembly of an optimal active site. Intriguingly, such an orchestrated coupling is absent for the systems with oxidative damage, which suggests the trend: **G:C** > 8oxo**G:A** for the rate of active-site pre-organization. Taken together, the trend in rates of active site pre-organization emerging from our studies is **G:C** > 8oxo**G:A** > 8oxo**G:C** >> **G:A**, which shows that the oxidative lesion enables the polymerase to favor dATP (mis)incorporation over correct dCTP incorporation. Moreover, this inferred trend correlates perfectly with relative nucleotide incorporation rates observed in kinetic experiments [6], suggesting that this pre-chemistry step is a significant determinant of polymerase fidelity. Our comparative analysis of correlations for the four systems strongly supports the participation of the polymerase fingers in a discriminatory role towards catalytic site pre-organization, Namely that, while the dynamical coupling between the polymerase fingers and the catalytic site aids in active-site pre-organization for the **G:C** system, it exerts a negative influence on the pre-organization of the active site for the **G:A** system. Thus, our observations identify a new discriminatory role for the polymerase fingers induced by the dynamical coupling in BF complexes devoid of lesions. For the 8oxo**G** systems the presence of the lesion disrupts the dynamical coupling to similar extents during correct/incorrect nucleotide incorporations thereby affecting the ability of the polymerase fingers to discriminate against incorrect nucleotides. Therefore, we suggest that the mutagenic potential of 8oxoG can partly be ascribed to this loss of discrimination. Collectively, these observations further extend the significance of the fingers domain in the fidelity mechanisms in A-family polymerases.

The analysis we have presented provides valuable insights on the mutagenic potential of 8oxo**G** during its interaction with high fidelity polymerases. Kinetic studies resolving the barriers for the conformational change and chemical steps during nucleotide incorporation opposite 8oxo**G** are required validate our predictions and to completely understand the role of the lesion in promoting DNA mispair formation by high fidelity polymerases during the nucleotide incorporation cycle. Nevertheless, the paradigm for nucleotide incorporation emerging from our collective analysis is that the pre-organization of the catalytic site is not only achieved on the basis of context-specific structural (catalytic site geometries) and energetic (stability and plasticity of the catalytic site) attributes of the active site but also

strongly influenced by the slow dynamical modes of the system.

Acknowledgments

We thank Dr. Tamar Schlick for providing us with the preprint of a manuscript [39] prior to publication. We also thank Ms Yingting Liu for her help and for providing the genetic algorithm based optimization code used for the refinement of CHARMM parameters for the 8oxo**G** base. We acknowledge funding from the Department of Bioengineering, University of Pennsylvania. Computational resources were provided in part by the National Center for Supercomputing Alliance under Grant DAC1103423 and National Partnership for Advanced Computational Infrastructure under Grant MCB060006.

References

1. Kunkel TA. DNA Replication Fidelity. J Biol Chem 2004;279(17):16895-16898.
2. Lindahl T. Instability and decay of the primary structure of DNA. Nature 1993;362:709-715.
3. Friedberg EC. DNA damage and repair. Nature 2003;421:436-439.
4. Hollstein M, Shomer B, Greenblatt M, Soussi T, Hovig E, Montesano R, Harris CC. Somatic point mutations in the p53 gene of human tumors and cell lines: Updated compilation. Nucl Acids Res 1996;24(1):141-146.
5. Brieba LG, Eichman BF, Kokoska RJ, Doublie S, Kunkel TA, Ellenberger T. Structural basis for the dual coding potential of 8-oxoguanosine by a high-fidelity DNA polymerase. Embo J 2004;23(17):3452-3461.
6. Hsu GW, Ober M, Carell T, Beese LS. Error-prone replication of oxidatively damaged DNA by a high-fidelity DNA polymerase. Nature 2004;431:217-221.
7. Furge LL, Guengerich FP. Pre-steady-state Kinetics of Nucleotide Insertion following 8-oxo-7,8-dihydroguanine Base Pair Mismatches by Bacteriophage T7 DNA Polymerase exo$^-$. Biochemistry-Us 1998;37(10):3567-3574.
8. Shibutani S, Takeshita M, Grollman AP. Insertion of specific bases during DNA synthesis past the oxidation-damaged base 8-oxodG. Nature 1991;349:431-434.
9. Fortini P, Pascucci B, Parlanti E, D'Errico M, Simonelli V, Dogliotti E. 8-Oxoguanine DNA damage: at the crossroad of alternative repair pathways. Mut Res Fund Mol Mech Mutag 2003;531(1-2):127-139.
10. Cheng KC, Cahill DS, Kasai H, Nishimura S, Loeb LA. 8-Hydroxyguanine, an abundant form of oxidative DNA damage, causes G—>T and A—>C substitutions. J Biol Chem 1992;267(1):166-172.
11. Jackson AL, Loeb LA. The contribution of endogenous sources of DNA damage to the multiple mutations in cancer. Mut Res Fund Mol Mech Mutag 2001;477(1-2):7-21.
12. Mambo E, Chatterjee A, de Souza-Pinto NC, Mayard S, Hogue BA, Joque MO, Dizdaroglu M, Bohr VA, Sidransky D. Oxidized guanine lesions and hOgg1 activity in lung cancer. Oncogene 2005;24(28):4496-4508.
13. Johnson SJ, Beese LS. Structures of Mismatch Replication Errors Observed in a DNA Polymerase. Cell 2004;116(6):803-816.
14. Florian J, Goodman MF, Warshel A. Computer simulations of protein functions: Searching for the molecular origin of the replication fidelity of DNA polymerases. Proc Natl Acad Sci USA 2005;102(19):6819-6824.
15. Florian J, Goodman MF, Warshel A. Computer Simulation of the Chemical Catalysis of DNA Polymerases: Discriminating Between Alternative Nucleotide Insertion Mechanisms for T7 DNA Polymerase. J Am Chem Soc 2003;125(27):8163-8177.
16. Geacintov NE, Broyde S, Buterin T, Naegeli H, Wu M, Yan SX, Patel DJ. Thermodynamic and structural factors in the removal of bulky DNA adducts by the nucleotide excision repair machinery. Biopolymers 2002;65(3):202-210.
17. Perlow RA, Broyde S. Extending the Understanding of Mutagenicity: Structural Insights into Primer-Extension Past a Benzo[a]pyrene Diol Epoxide-DNA Adduct. J Mol Biol 2003;327(4):797-818.
18. Wang LH, Broyde S. A new anti conformation for N-(deoxyguanosin-8-yl)-2-acetylaminofluorene (AAF-dG) allows Watson-Crick pairing in the Sulfolobus solfataricus P2 DNA polymerase IV (Dpo4). Nucl Acids Res 2006;34(3):785-795.
19. Wu X, Shapiro R, Broyde S. Conformational Analysis of the Major DNA Adduct Derived from the Food Mutagen 2-Amino-3-methylimidazo[4,5-f]quinoline. Chem Res Toxicol 1999;12(10):895-905.

20. Zhang L, Rechkoblit O, Wang LH, Patel DJ, Shapiro R, Broyde S. Mutagenic nucleotide incorporation and hindered translocation by a food carcinogen C8-dG adduct in Sulfolobus solfataricus P2 DNA polymerase IV (Dpo4): modeling and dynamics studies. Nucl Acids Res 2006;34(11):3326-3337.
21. Wang YL, Arora K, Schlick T. Subtle but variable conformational rearrangements in the replication cycle of Sulfolobus solfataricus P2 DNA polymerase IV (Dpo4) may accommodate lesion bypass. Protein Sci 2006;15(1):135-151.
22. Radhakrishnan R, Schlick T. Orchestration of cooperative events in DNA synthesis and repair mechanism unraveled by transition path sampling of DNA polymerase β's closing. Proc Natl Acad Sci USA 2004;101(16):5970-5975.
23. Yang L, Arora K, Beard WA, Wilson SH, Schlick T. Critical role of Magnesium Ions in DNA Polymerase β's Closing and Active Site Assembly. J Am Chem Soc 2004;126(27):8441-8453.
24. Arora K, Schlick T. In Silico Evidence for DNA Polymerase-β's Substrate-Induced Conformational Change. Biophysical J 2004;87:3088-3099.
25. Joyce CM, Benkovic SJ. DNA Polymerase Fidelity: Kinetics, Structure, and Checkpoints. Biochemistry-Us 2004;43(45):14317-14324.
26. Eger BT, Benkovic SJ. Minimal Kinetic Mechanism for Misincorporation by DNA-Polymerase-I (Klenow Fragment). Biochemistry-Us 1992;31(38):9227-9236.
27. Capson TL, Peliska JA, Kaboord BF, Frey MW, Lively C, Dahlberg M, Benkovic SJ. Kinetic Characterization of the Polymerase and Exonuclease Activities of the Gene 43 Protein of Bacteriophage T4. Biochemistry-Us 1992;31(45):10984-10994.
28. Johnson SJ, Taylor JS, Beese LS. Processive DNA synthesis observed in a polymerase crystal suggests a mechanism for the prevention of frameshift mutations. Proc Natl Acad Sci USA 2003;100(7):3895-3900.
29. Kiefer JR, Mao C, Braman JC, Beese LS. Visualizing DNA replication in a catalytically active Bacillus DNA polymerase crystal. Nature 1998;391(6664):304-307.
30. Kiefer JR, Mao C, Hansen CJ, Basehore SL, Hogrefe HH, Braman JC, Beese LS. Crystal structure of a thermostable Bacillus DNA polymerase I large fragment at 2.1 A resolution. Structure 1997;5:95-108.
31. Hsu GW, Kiefer JR, Burnouf D, Becherel OJ, Fuchs RPP, Beese LS. Observing Translesion Synthesis of an Aromatic Amine DNA Adduct by a High-fidelity DNA Polymerase. J Biol Chem 2004;279(48):50280-50285.
32. Hsu GW, Huang XW, Luneva NP, Geacintov NE, Beese LS. Structure of a High Fidelity DNA Polymerase Bound to a Benzo[a]pyrene Adduct That Blocks Replication. J Biol Chem 2005;280(5):3764-3770.
33. Insight II molecular modelling software. San Diego: Molecular Simulations Inc.; 2000.
34. Kale L, Skeel R, Bhandarkar M, Brunner R, Gursoy A, Krawetz N, Phillips J, Shinozaki A, Varadarajan K, Schulten K. NAMD2: Greater Scalability for Parallel Molecular Dynamics. J Comput Phys 1999;151(1):283-312.
35. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kale L, Schulten K. Scalable molecular dynamics with NAMD. J Comput Chem 2005;26(16):1781-1802.
36. MacKerell AD, Bashford D, Bellott M, Dunbrack RL, Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S, Joseph-McCarthy D, Kuchnir L, Kuczera K, Lau FTK, Mattos C, Michnick S, Ngo T, Nguyen DT, Prodhom B, Reiher WE, Roux B, Schlenkrich M, Smith JC, Stote R, Straub J, Watanabe M, Wiorkiewicz-Kuczera J, Yin D, Karplus M. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. J Phys Chem B 1998;102(18):3586-3616.
37. Mackerell AD, Banavali NK. All-atom empirical force field for nucleic acids: II. Application to molecular dynamics simulations of DNA and RNA in solution. J Comput Chem 2000;21(2):105-120.
38. Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M. Charmm - a Program for Macromolecular Energy, Minimization, and Dynamics Calculations. J Comput Chem 1983;4(2):187-217.
39. Radhakrishnan R, Arora K, Wang Y, Beard WA, Wilson SH, Schlick T. Regulation of DNA Repair Fidelity by Molecular Checkpoints: "Gates" in DNA Polymerase β's Substrate Selection. Biochemistry-Us 2006;45(51):15142-15156.
40. Radhakrishnan R, Schlick T. Correct and incorrect nucleotide incorporation pathways in DNA polymerase β. Biochem Biophys Res Comm 2006;350(3):521-529.
41. Radhakrishnan R. Coupling of Fast and Slow Modes in the Reaction Pathway of the Minimal Hammerhead Ribozyme Cleavage. Biophys J 2007;In press(published online June 1 [Epub ahead of print]).
42. Roux B. The calculation of the potential of mean force using computer simulations. Comput Phys Comm 1995;91(1-3):275-282.
43. Kumar S, Rosenberg JM, Bouzida D, Swendsen RH, Kollman PA. Multidimensional free-energy calculations using the weighted histogram analysis method. J Comput Chem 1995;16(11):1339-1350.

44. Glykos NM. Software news and updates - Carma: A molecular dynamics analysis program. J Comput Chem 2006;27(14):1765-1768.
45. Amadei A, Linssen ABM, Berendsen HJC. Essential dynamics of proteins. Proteins Struct Funct Gen 1993;17(4):412-425.
46. Amadei A, Linssen ABM, deGroot BL, Aalten DMFv, Berendsen HJC. An Efficient Method for Sampling the Essential Subspace of Proteins. J Biomol Struct Dyn 1996;13(4):615-625.
47. Freemont PS, Friedman JM, Beese LS, Sanderson MR, Steitz TA. Cocrystal Structure of an Editing Complex of Klenow Fragment with DNA. Proc Natl Acad Sci USA 1988;85(23):8924-8928.
48. Beese LS, Steitz TA. Structural basis for the 3'-5' exonuclease activity of Escherichia coli DNA polymerase I: A two metal ion mechanism. EMBO J 1991;10(1):25-33.
49. Steitz TA. A mechanism for all polymerases. Nature 1998;391:231-232.
50. Yang W, Lee JY, Nowotny M. Making and Breaking Nucleic Acids: Two-Mg2+-Ion Catalysis and Substrate Specificity. Mol Cell 2006;22(1):5-13.
51. Strater N, Lipscomb WN, Klabunde T, Krebs B. Two-Metal Ion Catalysis in Enzymatic Acyl- and Phosphoryl-Transfer Reactions. Angewandte Chem 1996;35(18):2024-2055.
52. Gregersen BA, Lopez X, York DM. Hybrid QM/MM Study of Thio Effects in Transphosphorylation Reactions: The Role of Solvation. J Am Chem Soc 2004;126(24):7504-7513.
53. Mildvan AS. Mechanisms of signaling and related enzymes. Proteins Struct Funct Gen 1997;29:401-416.
54. Gregersen BA, Lopez X, York DM. Hybrid QM/MM Study of Thio Effects in Transphosphorylation Reactions. Journal of the American Chemical Society 2003;125(24):7178-7179.
55. Lin P, Pedersen LC, Batra VK, Beard WA, Wilson SH, Pedersen LG. Energy analysis of chemistry for correct insertion by DNA polymerase β. Proc Natl Acad Sci USA 2006;103(36):13294-13299.
56. Lahiri SD, Zhang GF, Dunaway-Mariano D, Allen KN. The Pentacovalent Phosphorus Intermediate of a Phosphoryl Transfer Reaction. Science 2003;299:2067-2071.
57. Stahley MR, Strobel SA. Structural Evidence for a Two-Metal-Ion Mechanism of Group I Intron Splicing Science 2005;309:1587-1590.
58. Webster CE. High-Energy Intermediate or Stable Transition State Analogue: Theoretical Perspective of the Active Site and Mechanism of beta-Phosphoglucomutase J Am Chem Soc 2004;126(22):6840.
59. Radhakrishnan R. Mechanism of nucleotide incorporation in DNA polymerase β. Biochem Biophys Res Comm 2006;347(3):626-633.
60. Ferrin LJ, Mildvan AS. NMR studies of the conformations and interactions of substrates and ribonucleotide templates bound to the large fragment of DNA polymerase I. In: McMacken R, Kelly TJ, editors. DNA Replication and Recombination. Volume 47. New York: Alan R. Liss, Inc.; 1987. p 75-87.
61. Steitz TA, Steitz JA. A general two-metal-ion mechanism for catalytic RNA. Proc Natl Acad Sci USA 1993;90(14):6498-6502.
62. Thorpe IF, Brooks CL. Conformational Substates Modulate Hydride Transfer in Dihydrofolate Reductase. J Am Chem Soc 2005;127(37):12997-13006.
63. Thorpe IF, Brooks CL. The coupling of structural fluctuations to hydride transfer in dihydrofolate reductase. Proteins 2004;57(3):444-457.
64. Benkovic SJ, Hammes-Schiffer S. Enzyme Motions Inside and Out. Science 2006;312:208-209.
65. Hammes-Schiffer S, Benkovic SJ. Relating Protein Motion to Catalysis. Annu Rev Biochem 2006;75:519-541.
66. Olsson MHM, Parson WW, Warshel A. Dynamical Contributions to Enzyme Catalysis: Critical Tests of a Popular Hypothesis. Chem Rev 2006;106(5):1737-1756.
67. Bustamante C, Bryant Z, Smith SB. Ten years of tension: single-molecule DNA mechanics. Nature 2003;421:423-427.
68. Wuite GJL, Smith SB, Young M, Keller D, Bustamante C. Single-molecule studies of the effect of template tension on T7 DNA polymerase activity. Nature 2000;404:103-106.
69. Maier B, Bensimon D, Croquette V. Replication by a single DNA polymerase of a stretched single-stranded DNA. Proc Natl Acad Sci USA 2000;97(22):12002-12007.
70. Venkatramani R, Radhakrishnan R. A computational study of the force dependence of phosphoryl transfer during DNA synthesis by a high fidelity polymerase. Phys Rev Lett 2007;submitted (under revision).
71. Steitz TA. DNA polymerases: Structural diversity and common mechanisms. J Biol Chem 1999;274(25):17395-17398.
72. Rothwell PJ, Waksman G. Structure and mechanism of DNA polymerases. Adv Prot Chem 2005;71:401-440.
73. Arora K, Beard WA, Wilson SH, Schlick T. Mismatch-induced conformational distortions in polymerase beta/DNA complex support an induced-fit mechanism for fidelity. Biochemistry-Us 2005;44:13328-13341.

74. Patel PH, Suzuki M, Adman E, Shinkai A, Loeb LA. Prokaryotic DNA polymerase I: Evolution, structure, and "base flipping" mechanism for nucleotide selection. J Mol Biol 2001;308(5):823-837.
75. Loh E, Loeb LA. Mutability of DNA polymerase I: implications for the creation of mutant DNA polymerases. DNA Rep 2005;4(12):1390-1398.
76. Minnick DT, Bebenek K, Osheroff WP, Turner RM, Jr., Astatke M, Liu L, Kunkel TA, Joyce CM. Side Chains That Influence Fidelity at the Polymerase Active Site of Escherichia coli DNA Polymerase I (Klenow Fragment). J Biol Chem 1999;274(5):3067-3075.
77. Singh K, Modak MJ. Presence of 18-A Long Hydrogen Bond Track in the Active Site of Escherichia coli DNA Polymerase I (Klenow Fragment). Its Requirement in the Stabilization of Enzyme-Template-Primer Complex. J Biol Chem 2003;278(13):11289-11302.
78. McCain MD, Meyer AS, Schultz SS, Glekas A, Spratt TE. Fidelity of Mispair Formation and Mispair Extension Is Dependent on the Interaction Between the Minor Groove of the Primer Terminus and Arg668 of DNA Polymerase I of Escherichia coli. Biochemistry-Us 2005;44(15):5647-5659.

Figure captions

1) **Figure 1:** (a) The active site region for the BF-DNA-dNTP ternary complex used in our correlation analysis. The circled region depicts the catalytic site. (b) The fully solvated and neutralized ternary complex used in our MD simulations. The circled region depicts the location of the active site fragment. (c) The catalytic site, depicting the key components required for a two-metal-ion assisted phosphoryl transfer reaction.

2) **Figure 2:** Average values (solid-lines) and standard deviations of metal-ligand and nucleophilic attack distances for the four model systems (**G:C**, **G:A**, 8oxo**G:C** and 8oxo**G:A**) obtained from 5ns classical simulations. Also shown are values of metal ligand distances expected for an ideal two-metal-ion geometry (solid bars) taken from reference [58].

3) **Figure 3:** The average catalytic site geometry obtained from 5ns classical simulations for the **G:C** (a) and **G:A** (b) systems. The primary reactive distances for catalysis, O3′-P_α (d_a) and O3′-MG2 (d_b) are shown. Corresponding free energy landscapes are shown in figures (c) and (d) respectively with distances d_a and d_b given in Å. The location of the average simulation geometry is marked by X in these plots. The scale for energies (shown on the right for each free energy plot) is in units of k_BT ~0.6 kcal/mol (for a T value of 300 K). The errors in free energies are estimated to be ~ ± 0.9 k_BT (0.54 kcal/mol).

4) **Figure 4:** Colored symbols represent values of various metal-ligand distances for the four model systems (**G:C**, **G:A**, 8oxo**G:C** and 8oxo**G:A**) obtained during our umbrella sampling simulations. The nucleophilic attack distance dNTP:P_α—**A**:O3′ and the MG2—**A**:O3′ distance were constrained to a range of values between 5.0—2.0 Å (see Methods). For all systems except the control **G:C** (which shows two water molecules at the catalytic site) the catalytic Mg^{2+} is liganded to three water molecules for structures where d_a and d_b values are large and far from those for an ideal two-metal-ion geometry (from reference [58]) represented by the solid bars. The extra water is expelled as d_a and d_b are constrained to approach their ideal two-metal-ion values.

5) **Figure 5:** The average catalytic site geometry obtained from 5ns classical simulations for the 8oxo**G:C** (a) and 8oxo**G:A** (b) systems. The primary reactive distances for catalysis, O3′-P_α (d_a) and O3′-MG2 (d_b) are shown. Corresponding free energy landscapes are shown in figures (c) and (d) respectively with distances d_a and d_b given in Å. The location of the average simulation geometry is marked by X in these plots. The scale for energies (shown on the right for each free energy plot) is in units of k_BT ~0.6 kcal/mol (for a T value of 300 K). The errors in free energies are estimated to be ~ ± 0.9 k_BT (0.54 kcal/mol).

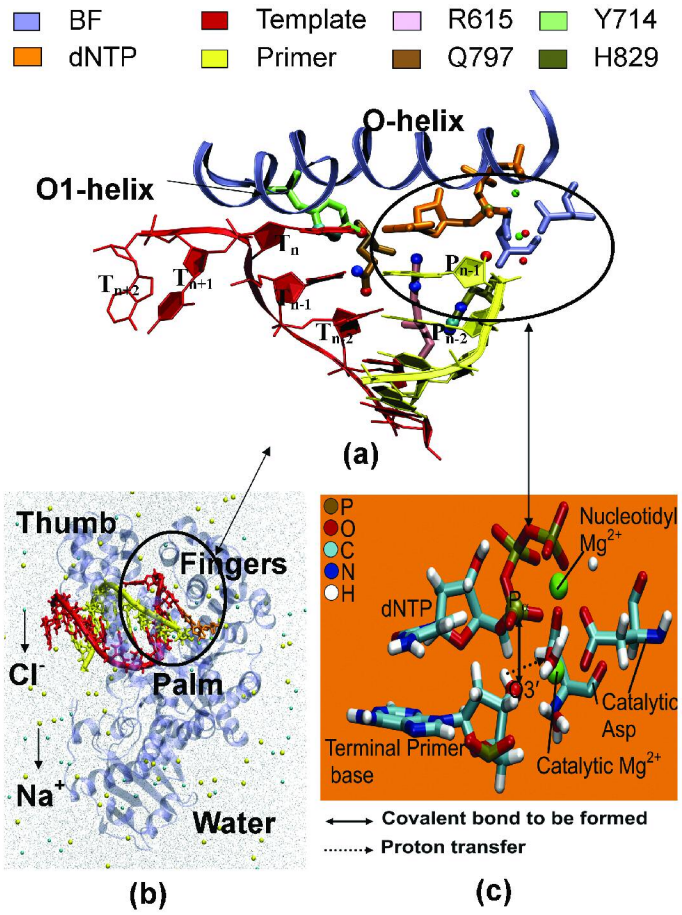
Residue	Effect of mutation*	Residue-DNA interaction
R615	• 23-fold decrease in DNA binding. • 2-fold decrease in fidelity • 300-fold decrease in catalytic activity	• Contacts dNTP sugar (O4′) and **A**(P_{n-1}) base (N3)
R714	• Decrease in replication fidelity	• Contacts **G**(T_n) base (N2 hydrogen)
Q797	• 40-fold decrease in DNA binding • 3-10 fold increase in fidelity	• Contacts **T**(T_{n-1}) base (O2) and **C**(T_{n-2}) sugar (O4′)
H829	• Increase in replication fidelity • Decrease in DNA binding	• Contacts – **A**(P_{n-1}) sugar (O_4) and **G** (P_{n-2}) base (N_3)

*** from references [76] and [78]**

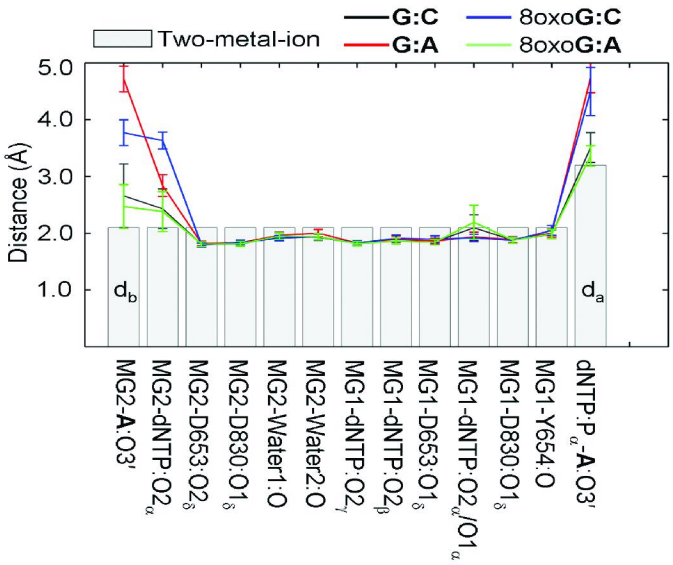
Table I: Summary of mutation of key polymerase residues as inferred from the interactions seen for the **G:C** simulations, and the effects on **a G:A** mismatch and oxidatively damaged **G** at the insertion site on these interactions. Notation for each DNA base is as follows: Each base is denoted either as a primer (P) or a template (T) base with the subscript n denoting the nascent base pair with the template **G/**8oxo**G** and the incoming dNTP. Subscripts n-1,n-2 represent bases further down in the DNA duplex.

Residue-Base Interaction	**G:C** kcal/mol	**G:A** kcal/mol	8oxo**G:C** kcal/mol	8oxo**G:A** kcal/mol
R615-dCTP/dATP	-34.6 (4.8)	-18.9 (2.7)	-34.4 (4.5)	-33.2 (3.6)
R615-**A**(P_{n-1})	-17.6 (2.8)	-11.9 (1.9)	-18.8 (3.0)	-15.7 (2.4)
Y714-dCTP/dATP	-2.2 (0.9)	-1.6 (0.8)	-4.6 (0.7)	-4.0 (0.9)
Y714-**G**/8oxo**G**(T_n)	-8.9 (1.2)	-6.8 (1.2)	-2.7 (0.8)	-3.9 (1.2)
Q797 –**T**(T_{n-1})	-5.8 (0.8)	-5.2 (1.0)	-2.9 (1.3)	-3.7 (1.2)
H829 –**A**(P_{n-1})	-9.3 (1.3)	-5.6 (1.0)	-3.3 (1.2)	-2.5 (1.0)
H829 –**G**(P_{n-2})	-4.9 (2.4)	-8.1 (1.7)	-5.5 (2.1)	-5.9 (1.5)

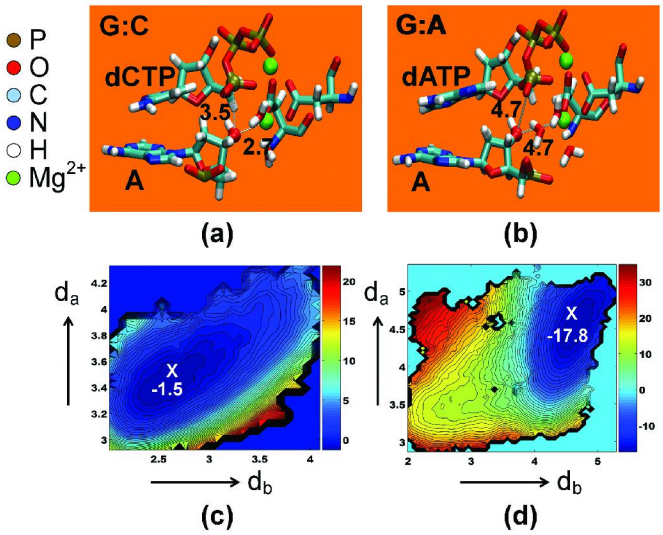
Table II: Nonbonded (van der Waals + electrostatic) interaction energy (in kcal/mol) between four highly conserved polymerase residues at the active site (R615, Y714, Q797 and H829) and the DNA for all four model systems (**G:C**, **G:A**, 8oxo**G:C** and 8oxo**G:A**), averaged over the last 2ns of our unconstrained MD trajectories. The standard deviations are given in parenthesis.



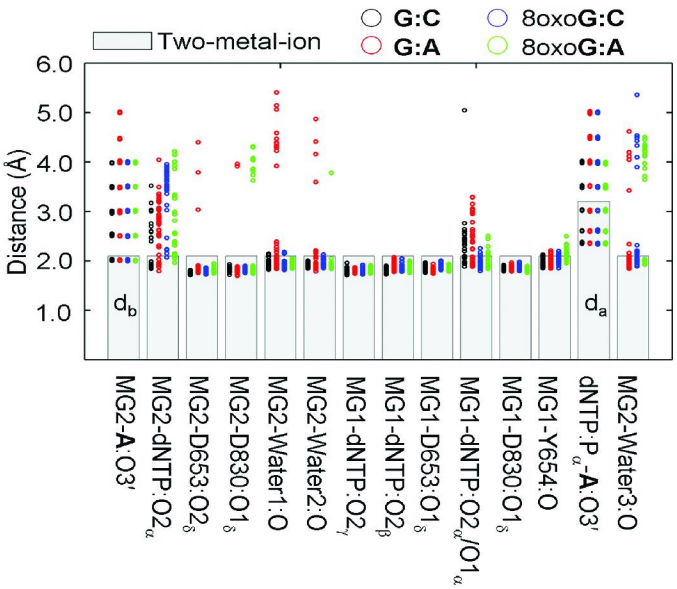
215x279mm (300 x 300 DPI)



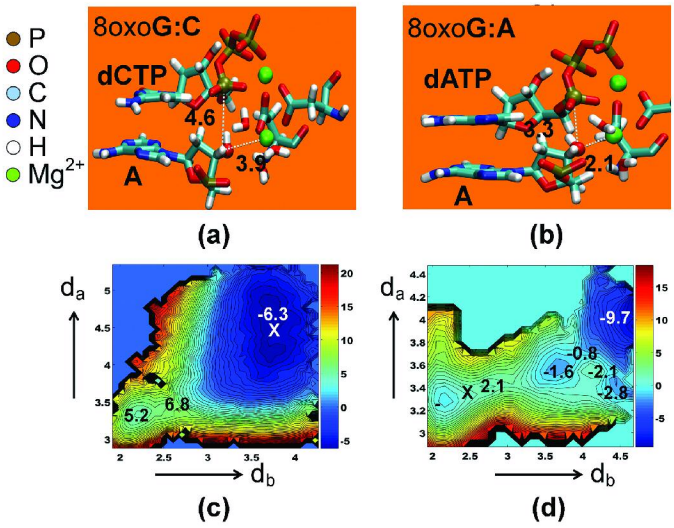
215x279mm (300 x 300 DPI)



215x279mm (300 x 300 DPI)



215x279mm (300 x 300 DPI)



215x279mm (300 x 300 DPI)