

Visualizing Sequence-Governed Nucleotide Selectivities and Mutagenic Consequences through a Replicative Cycle: Processing of a Bulky Carcinogen N^2 -dG Lesion in a Y-Family DNA Polymerase[†]

Pingna Xu,[‡] Lida Oum,[§] Yuan-Cho Lee,[‡] Nicholas E. Geacintov,[§] and Suse Broyde^{*,‡}

[‡]Department of Biology, New York University, 1009 Silver Center, 100 Washington Square East, New York, New York 10003, and

[§]Department of Chemistry, New York University, 1001 Silver Center, 100 Washington Square East, New York, New York 10003

Received December 28, 2008; Revised Manuscript Received April 4, 2009

ABSTRACT: Understanding how DNA polymerases process lesions remains fundamental to determining the molecular origins of mutagenic translesion bypass. We have investigated how a benzo[*a*]pyrene-derived N^2 -dG adduct, 10*S*-(+)-*trans-anti*-[BP]- N^2 -dG ([BP]G*), is processed in Dpo4, the well-characterized Y-family bypass DNA polymerase. This polymerase has a slippage-prone spacious active site region. Experimental results in a 5'-C[BP]G*G-3' sequence context reveal significant selectivity for dGTP insertion that predominantly yields -1 deletion extension products. A less pronounced error-prone nonslippage pathway that leads to full extension products with insertion of A > C > G opposite the lesion is also observed. Molecular modeling and dynamics simulations follow the bypass of [BP]G* through an entire replication cycle for the first time in Dpo4, providing structural interpretations for the experimental observations. The preference for dGTP insertion is explained by a 5'-slippage pattern in which the unmodified G rather than G* is skipped, the incoming dGTP pairs with the C on the 5'-side of G*, and the -1 deletion is produced upon further primer extension which is more facile than nucleotide insertion. In addition, the simulations suggest that the [BP]G* may undergo an *anti/syn* conformational rearrangement during the stages of the replication cycle. In the minor nonslippage pathway, the nucleotide insertion preferences opposite the lesion are explained by relative distortions to the active site region. These structural insights, provided by the modeling and dynamics studies, augment kinetic and limited available crystallographic investigations with bulky lesions, by providing molecular explanations for lesion bypass activities over an entire replication cycle.

Understanding how DNA polymerases process lesions remains fundamental to determining the molecular origins of mutagenic translesion bypass. It is now widely appreciated that high-fidelity polymerases that are stalled by DNA lesions are replaced by translesion bypass polymerases that are capable of transiting the lesions (1–6). The structural differences between high-fidelity and translesion bypass polymerases that determine error-prone or error-free bypass of the lesions are of significant and fundamental interest since mutations can lead to cancer and

other diseases (7). To gain insights into the effects of polymerase structure on translesion bypass on a molecular level, we have utilized a model system employing a bulky mutagenic DNA lesion derived from benzo[*a*]pyrene (BP),¹ a widely studied environmental contaminant (7). Like many other polycyclic aromatic hydrocarbons, BP is metabolized in vivo to numerous metabolites (8–10), with the (+)-7(*R*),8(*S*),9(*S*),10(*R*)-benzo[*a*]pyrene diol epoxide [(+)-*anti*-BPDE] being a predominant tumorigenic derivative of BP (11). The major stable lesion formed when (+)-*anti*-BPDE reacts with DNA *in vitro* (12) or *in vivo* (13, 14) is the 10*S*-(+)-*trans-anti*-[BP]- N^2 -dG ([BP]G*) adduct (Figure 1A). As a representative high-fidelity polymerase, we investigated in a previous study (15) the A-family polymerase from the bacterium *Bacillus stearothermophilus*, specifically the fragment called BF (16). The model Y-family bypass polymerase from the archaeon bacterium *Sulfolobus solfataricus* (17), Dpo4, is investigated in this work. Dpo4 belongs to the DinB branch of the Y-family polymerases, of which human Pol κ is also a member (18).

A number of crystal structures in the forms of binary and ternary complexes, with and without DNA lesions, have been published for these two enzymes (e.g., refs 16 and (19–39)), and the availability of these structural data motivated our selection of

[†]This work was supported by the National Cancer Institute, National Institutes of Health, through Grants CA28038 (S.B.) and CA099194 (N.E.G.). Partial support for computational infrastructure and systems management was also provided by Grant CA75449 (S.B.). Computational resources supported by the NSF Partnerships for Advanced Computational Infrastructure are gratefully acknowledged.

^{*}To whom correspondence should be addressed. Telephone: (212) 998-8231. Fax: (212) 995-4015. E-mail: broyde@nyu.edu.

¹Abbreviations: PAH, polycyclic aromatic hydrocarbon; BP, benzo[*a*]pyrene; BPDE, 7,8-dihydroxy-9,10-epoxybenzo[*a*]pyrene diol epoxide; (+)-*anti*-BPDE, (+)-7(*R*),8(*S*),9(*S*),10(*R*)-BPDE; (-)-*anti*-BPDE, (-)-7(*S*),8(*R*),9(*R*),10(*S*)-BPDE; [BP]G*, (+)-*trans-anti*-[BP]- N^2 -deoxyguanosine; dNTP, 2'-deoxynucleoside triphosphate; Dpo4, *S. solfataricus* P2 DNA polymerase IV; CG, conjugate gradient; SD, steepest descent; MD, molecular dynamics; rmsd, root-mean-square deviation; QM, quantum mechanics; MM, molecular mechanics; AF, aminofluorene; [AF]G*, *N*-(2'-deoxyguanosin-8-yl)aminofluorene.

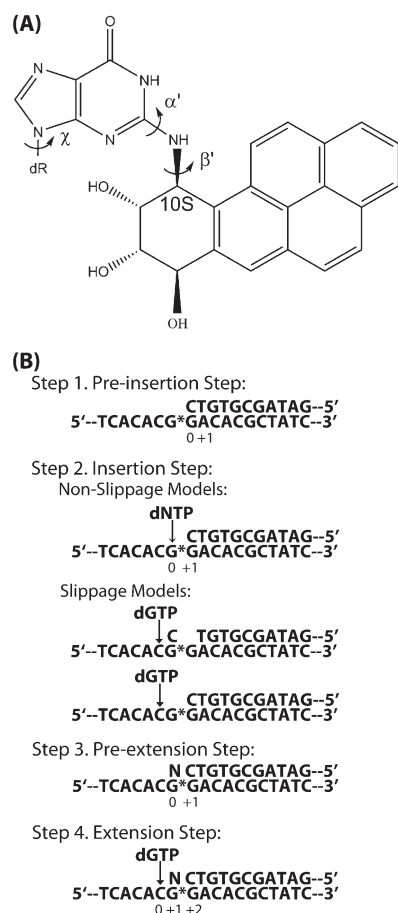


FIGURE 1: (A) Structure of the 10S-(+)-trans-anti-BP- N^2 -dG ([BP]G*, G*) adduct. Torsion angle definitions are as follows: χ , O4'-C1'-N9-C8; α' , N1-C2- N^2 -[BP]-C¹⁰; β' , C2- N^2 -[BP]-C¹⁰-C⁹. (B) DNA sequences for the G*-containing models. This base pair numbering pertains to the modeling studies. The numbering for the experimental work shown in Figures 2–4 is defined in Figure 2.

these two polymerases. In BF, proceeding from the binary complex (enzyme with primer–template DNA) to the ternary complex (enzyme with primer–template DNA and dNTP) involves a closing motion of the fingers domain, with an induced-fit mechanism (40, 41) that is used to select for the correct incoming dNTP. The structure of Dpo4 is significantly different from that of BF and has a preformed spacious and solvent-accessible active site that can simultaneously accommodate two templating bases (6, 25). Proceeding from the binary complex to the ternary complex does not involve a closing motion of the finger domain, and there is no analogous induced-fit mechanism for selecting the correct incoming dNTP (34). Furthermore, in BF, like in other high-fidelity polymerases, the minor groove side of the DNA duplex is tightly filled with enzyme amino acid residues and constitutes a scanning track that is crucial for high fidelity and processivity (20, 42, 43). In contrast, Dpo4 has a large open pocket on the major groove side and a smaller one on the minor groove side of the evolving DNA duplex, and there is no scanning track on the minor groove side. A recent review discusses these features and their connection to lesion processing (44).

Our objectives are to delineate the distinguishing features of translesion bypass involving the model [BP]G* bulky adduct at or near the primer–template junction within the active site of Dpo4; we wish to gain a deeper understanding of the effects of

base sequence context in which the lesions are embedded and to compare these phenomena with the case for BF. We have previously employed molecular modeling and dynamics simulations to follow these interactions through an entire replication cycle in a 5'-CG*G-3' sequence context in BF (15). In the work presented here, we use these computational approaches to investigate how Dpo4 processes the [BP]G* adduct through the same replication cycle (Figure 1B, steps 1–4) in the same sequence context; we again considered both binary and ternary complexes and all four partners opposite the [BP]G* adduct during the dNTP insertion and extension steps (Figure 1B, steps 2–4).

Our current experiments provide the foundation for the modeling studies and show that in the CG*G sequence context, Dpo4 bypasses the [BP]G* adduct relatively efficiently by predominantly utilizing a -1 deletion pathway that involves a misalignment of the primer strand at the active site (strand slippage). In the single-dNTP insertion assays with the lesion site next to be replicated, dGTP is predominantly selected, followed by dATP, and then dCTP. There is no observable dTTP insertion. Further extension beyond the lesion site is more facile than nucleotide incorporation. In striking contrast, BF was shown to be mainly blocked by this adduct: while nucleotide insertion opposite the adduct was observed with purines favored over pyrimidines, further extension was strongly inhibited (15, 45). In addition, the base 5'-flanking the adduct affects the nucleotide selectivities in Dpo4, but not in BF (45). Our current modeling studies provide detailed dynamic molecular structures of these various intermediates in the multistep process of replication past the [BP]G* adduct in Dpo4. The selectivity for dGTP is explained by a primer strand relocation or misalignment mechanism: an unusual 5'-slippage mechanism, in which the undamaged G rather than G* is skipped, and the incoming dGTP pairs with C, is suggested for the CG*G sequence context studied in this work. In a previous study with Dpo4 utilizing the 5'-TG*G-3' sequence context as a template, we proposed an analogous mechanism for explaining the preferred selectivity of insertion of dATP in single-nucleotide insertion assays (46). In contrast, a study of dNTP insertion opposite G* in the 5'-AG*C-3' sequence revealed little nucleotide selectivity under saturating conditions of dNTP concentrations (47). Primer strand slippage also accounts for the efficient translesion bypass that occurs mainly by the -1 deletion pathway in Dpo4. On the basis of experimental primer extension data reported here and relevant available experimental X-ray crystallographic data (25, 28, 34), we employed molecular dynamics simulation methods to analyze the structural features of the intermediate binary and ternary complexes over an entire replication cycle. This approach offers an integrated strategy for studying the dynamic properties of molecular interactions at the active sites of polymerases; it yields structural details about these intermediates that are not provided by the kinetics and have not yet been observed by crystallographic methods.

MATERIALS AND METHODS

Chemicals, Enzymes, and Oligonucleotides. [γ -³²P]ATP (3000 ci/mmol) was purchased from Perkin-Elmer Life Sciences, Inc. (Boston, MA). The dNTPs (dATP, dCTP, dGTP, and dTTP) were purchased from New England Biolabs, Inc. (Beverly, MA). *S. solfataricus* DNA polymerase IV (Dpo4) was kindly provided by O. Rech Koblit and D. J. Patel. The site-specifically modified 43-mer template oligonucleotide strand, 5'-GAC TAC GTA CTG TCA CC G*GA CAC GCT ATC TGG CCA GAT

CCG C-3', was synthesized, purified, and verified as described previously (48).

Primer Extension Reactions. Running-start and standing-start primer extension reactions were carried out as described in more detail previously (15, 46). Briefly, in the running-start experiments, the 43-mer templates were annealed to a 32 P-labeled 22-mer primer (defined in Figure 2A), with the [BP]G* adduct (G*) positioned at template position 25 counted from the 3'-end of the template strand [CG*G template (Figure 2A)]. The running-start experiments were carried out under the following conditions: 4.5 nM template strands, 10 nM Dpo4, and 100 μ M dNTP in buffer solution [50 mM Tris-HCl (pH 8.0), 5 mM MgCl₂, 1 mM DTT, 50 μ g/mL BSA, and 4% glycerol]. The reaction was allowed to proceed at 37 °C for selected times before being terminated with 5 μ L of stop solution (20 mM EDTA, 95% formamide, 0.05% bromophenol blue, and 0.05% xylene cyanol). After the specified incubation times, the DNA samples were heated at 90 °C for 5 min, chilled on ice, and then applied to a 15–20% denaturing polyacrylamide gel containing 7 M urea. The replication products were visualized by autoradiography and quantitatively analyzed by a Storm 840 phosphorimager and Storm ImageQuant (GE Healthcare). In the standing-start single-dNTP insertion experiments, the 43-mer CG*G templates were annealed to a 32 P-labeled 24-mer primer (Figure 2D). The one-step dNTP insertion assays were carried out under the following conditions at 37 °C: 2 nM CG*G template–primer complex, 2 nM Dpo4, and each dNTP at 500 μ M, one at a time, with an incubation time of 5 min.

DNA Sequencing Experiments. A two-phase gel electrophoresis method originally used by Shibutani (49) was adapted and significantly modified to sequence the primer extension products derived from the running-start experiments. The principle of our modified approach is outlined in Figure 2. The main difference from the original Shibutani method is that we synthesized a radioactively labeled 22-mer primer strand. This internally labeled 32 P-labeled 22-mer primer strand was prepared by extending an 18-mer primer annealed to a 22-mer template strand (5'-end-labeled) that had the same sequence as the first 22-nucleotide sequence of the 43-mer template strand counted from the 3'-end (Figure 2A). The 18-mer was extended to a 22-mer by incubating the annealed 22-mer template–18-mer duplex (100 nM) in solution containing 500 nM BF₃, 2 μ M dTTP and [α - 32 P]dCTP, and 4 μ M dGTP at 37 °C for 30 min, thus generating the internally 32 P-labeled primer strand (the labeled C is underlined in Figure 2). These 22-mer primer oligonucleotides were then purified by denaturing gel electrophoresis, eluted, desalted, and annealed with either a 43-mer unmodified template strand (CGG) or a 43-mer template strand containing a single [BP]G* adduct [CG*G (Figure 2A)]. The 22-mer primer strand was then extended using Dpo4 as described above (4.5 nM template strand, 10 nM Dpo4, and each of the four dNTPs at 500 μ M at 37 °C). The primer extension products were then subjected to 7 M urea–20% denaturing polyacrylamide gel electrophoresis. The labeled 41–43-mer extended primer strands were then eluted from the gels, annealed with a 10-fold excess of unmodified and unlabeled 43-mer template strands CGG. These duplexes were then incubated with restriction enzymes *Tsp45I* and *AhoI* that sequence-specifically cleave the duplexes as shown in Figure 2B. The incubations of the duplexes with restriction enzymes *TspI* and *AhoI* (four units each) were conducted at 37 and 65 °C, respectively, as described by the supplier (New England BioLabs Inc.). The cleavage products

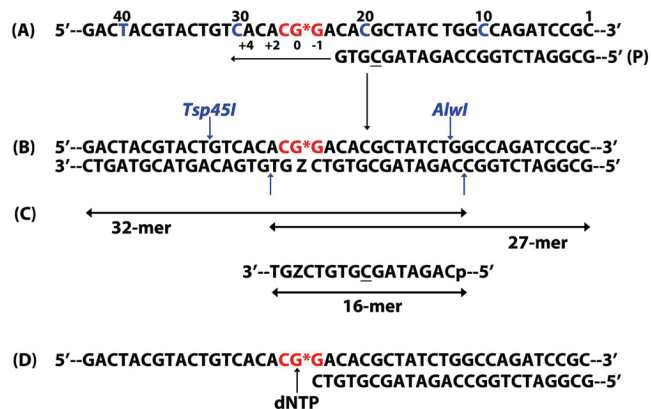


FIGURE 2: (A) Sequence of the 43-mer template containing the single [BP]G* adduct (G*) annealed to the complementary 22-mer primer (P) containing a 32 P-labeled C nucleotide (underlined). After termination of the primer extension reaction in panel A, the fully extended primers were isolated and purified by means of a 12% native gel and reannealed in buffer solution with a 10-fold excess of the unmodified 43-mer template strand (the G* is replaced by G) to yield the radioactively labeled newly synthesized primer strand annealed to the unmodified template strand shown in panel B. These duplexes were incubated with the restriction enzymes *Tsp45I* and *AhoI* that sequence-specifically cleave the extended primer strand at the sites shown. Successful cleavage by both enzymes yields a labeled 16-mer fragment, while single cuts by *Tsp45I* or *AhoI* fragments yield radioactively labeled 27-mer and 32-mer, respectively. This method is a modified version of the two-phase gel method described originally by Shibutani (49). (C) The lengths of the products generated by single- and double-cleavage events are shown by double-headed arrows. (D) Template–primer complex used in the single-dNTP insertion experiments shown in Figure 4A.

were separated by a two-phase gel system: the first 10 cm of the 82 cm plate consisted of a 20% denaturing polyacrylamide gel, whereas the remaining 72 cm consisted of a 20% native polyacrylamide gel. The gels were allowed to run overnight (~20 h) at 40 W and 2000 V. The migration patterns of the 16-mers were then compared to those of 16-mer standards 5'-ACGATAGCGTGTCTZGT, where Z is A, G, C, or T and Z corresponds to the nucleotide inserted opposite G* during the primer extension reaction (Figure 2). If the polymerase skips G* during the translesion synthesis (TLS), the 15-mer 5'-ACGATAGCGTGTCTG is observed which corresponds to a –1 deletion (Del) or frameshift mutation. The 15-mer Del products can be easily distinguished from the 16-mers that represent point mutations. The different 16-mers for which Z is A, G, C, or T can be distinguished only if their mobilities are sufficiently different from one another.

Molecular Modeling. Three Dpo4 crystal structures [Protein Data Bank (PDB) entries 1S0M, 2ASL, and 1JXL] (25, 28, 34) were used as the starting structures for molecular modeling, and the coordinates were obtained from the PDB (50). All molecular modeling was carried out using Insight 2000.1 (Accelrys, Inc., a subsidiary of Pharmacia, Inc.).

(i) **Remodeling the Crystal Structures.** An initial model for the unmodified Dpo4 ternary complex was constructed as described previously (51) on the basis of the crystal structure with PDB entry 1S0M (28) and provided a near-reaction ready active site region with well-coordinated Mg²⁺ ions (Figure S1 of the Supporting Information). The binary postinsertion complex utilized the structure from (34) (PDB entry 2ASL), which originally contained 7,8-dihydro-8-oxoguanine (oxoG); it was remodeled as follows: The oxoG was replaced with a G, the Ca²⁺ ion which is far from O3' of the primer terminus was removed,

and missing hydrogen atoms were added to the crystal structure by the LEaP module in AMBER (52). The slippage ternary complex based on the type II Dpo4 crystal structure (PDB entry 1JXL) (25) was remodeled as described by Xu et al. (46) and in the Supporting Information.

(ii) *Constructing the Initial Models Containing 10S-(+)-trans-anti-[BP]-N²-dG ([BP]G*)*. The remodeled crystal structures described above were used to build the initial models with [BP]G* for MD simulations. The DNA sequence was remodeled according to that used experimentally (Figure 1B). The BP moiety was covalently linked to N² of the guanine. According to the different positions of [BP]G* in Dpo4 at the four steps, four corresponding nonlippage DNA structures and one slippage DNA structure were built. For the nonlippage models, we investigated both *anti* and *syn* conformations of the glycosidic torsion χ of the BP-adducted dG* (Figure 1A) opposite all four partners. In addition, for G:A and G:G mismatches, *anti-anti*, *anti-syn*, and *syn-anti* conformations were investigated; we did not consider *syn-syn* conformations as these are rarely observed (53). A *syn*-dATP has been observed opposite a damaged template in a Dpo4 crystal structure (26). Torsion angles α' and β' at the linkage site between the guanine and the BP (Figure 1A) were rotated at 10° intervals within the low-energy domains (54), in combination to locate conformations with minimal collisions. Slippage models were constructed using the same strategy that is described by Xu et al. (46), detailed also in the Supporting Information. Two *anti*-[BP]G* models and one *syn*-[BP]G* model were located. The torsion angles for all initial models are given in Table S1 of the Supporting Information. The remodeled crystal structures also provided the initial models for the unmodified controls. All initial models are shown in Figures S2 and S3 of the Supporting Information.

MD Simulation Protocol. Partial charges for *anti/syn*-[BP]G*, *anti/syn*-dATP, *anti/syn*-dGTP, and *anti*-dCTP and *anti*-dTTP were computed previously (15, 55). The minimization, equilibration, and production were carried out using the same software and protocols described by Xu et al. (15).

Trajectory Analysis and Stability of the MD Simulations. To obtain ensemble average values for properties of interest, trajectories were collected for all systems and were analyzed using the PTRAJ and CARNAL modules of the AMBER package (52). For each system, the rmsds (root-mean-square deviations) of the whole structure and the active site (composed of all the residues within 5 Å of the nascent base pair) were calculated relative to the first production frame (Figure S4 of the Supporting Information). We found that the active sites of all systems are locally stable in the 1.0–2.5 ns time frame. The following analyses are based on these ranges unless otherwise stated. All structural figures were prepared using PyMOL (56).

RESULTS

*Primer Extension Assays Reveal Translesion Bypass in the CG*G Sequence Context*. A goal of this work was to evaluate the characteristics of translesion bypass of [BP]G* adducts by Dpo4 in the CG*G sequence context (Figure 2) and compare the results to the features obtained with the A-family polymerase BF investigated earlier in the same base sequence context (15). In the studies with BF, it was shown that this polymerase inserted dNTPs rather efficiently opposite the lesion but that further primer extension beyond the lesion was severely inhibited. Here we show that, in contrast, in the case of

Dpo4, the rate-limiting step in translesion synthesis (TLS) of the same lesion is the apparent insertion of a dNTP opposite the lesion, rather than the extension step as in the case of BF. We utilized the term “apparent” to highlight that, in our delineated slippage mechanism discussed below, the dNTP is actually inserted opposite the C on the 5'-side of G*. Another goal of this study was to compare the effects of base sequence in the CG*G sequence context with our earlier studies in the TG*G (46) and AG*C sequence contexts (47). We previously found that at 37 °C all four dNTPs are inserted with rather similar probabilities opposite the lesion in the AG*C sequence context under conditions of saturating dNTP concentration, and that dATP insertion is significantly favored in the TG*G sequence context. Here, in the CG*G sequence context, we have used a two-phase gel sequencing approach to determine that the apparent insertion of dGTP opposite [BP]G* is favored, and that –1 deletions are the dominant full translesion bypass products. An objective of this work was to elucidate the structural origins of such base sequence context effects. The experimental primer extension results were generated to set the stage for the modeling and computational components of this effort.

The distribution of primer extension products of different lengths generated by Dpo4 using 43-mer templates containing a single [BP]G* residue (Figure 1A) depends on base sequence context (45), temperature, activity of the Dpo4 polymerase sample, primer–template to Dpo4 ratios, and polymerase and DNA concentrations. The time course of a typical running-start primer extension reaction is shown in Figure 3A, and an example of a profile of extension products formed after a 15 min incubation is shown in Figure 3B. Additional product profiles (for the 12 and 30 min time points) are shown in Figure S5 of the Supporting Information. Since the templates were gel-purified before use, and the purity of the templates was better than 98% (45), the fast phase cannot be attributed to the presence of unmodified template strand contamination. In these experiments, the 22-mer primer strand (labeled P in Figure 3A) was present in ~25% excess to maximize the extent of formation of products that were needed for subsequent sequencing experiments. Approximately 65% of the 22-mer strands initially present were rapidly extended within the first 3 min of incubation time; another 8% were converted from 3 to 15 min, and only 2% of the strands were converted during the last 15 min of incubation (Figure 3C). Thus, the depletion of primer band P occurs in two phases, a rapid one within the first 3 min time interval, followed by a significantly slower phase within the next 12 min. Within the final 15 min incubation time, the concentrations of primer strands remain practically constant. Similar two-phase kinetics observed upon primer extension catalyzed by Dpo4 have been reported and discussed previously (38). Such biphasic kinetics most likely represent differing conformational states of the lesion within the polymerase active site that interconvert slowly; they are consistent with the existence of a productive ternary complex conformation that can lead to nucleotide incorporation, and a nonproductive state that is reaction-incapable (38). Such two-state conformations have been crystallographically observed with Dpo4 for several bulky lesions (28, 38, 39). The ~25% residual 22-mer strand levels (Figure 3C) represent the single-stranded primers that are in excess. As shown elsewhere (45), residual primer strand levels are not observed when the initial primer: template ratio is ≤ 1.0 . The thermodynamic stabilities of the 22-mer primer–43-mer template duplexes are lower than the stabilities of any of the partially or fully extended and longer

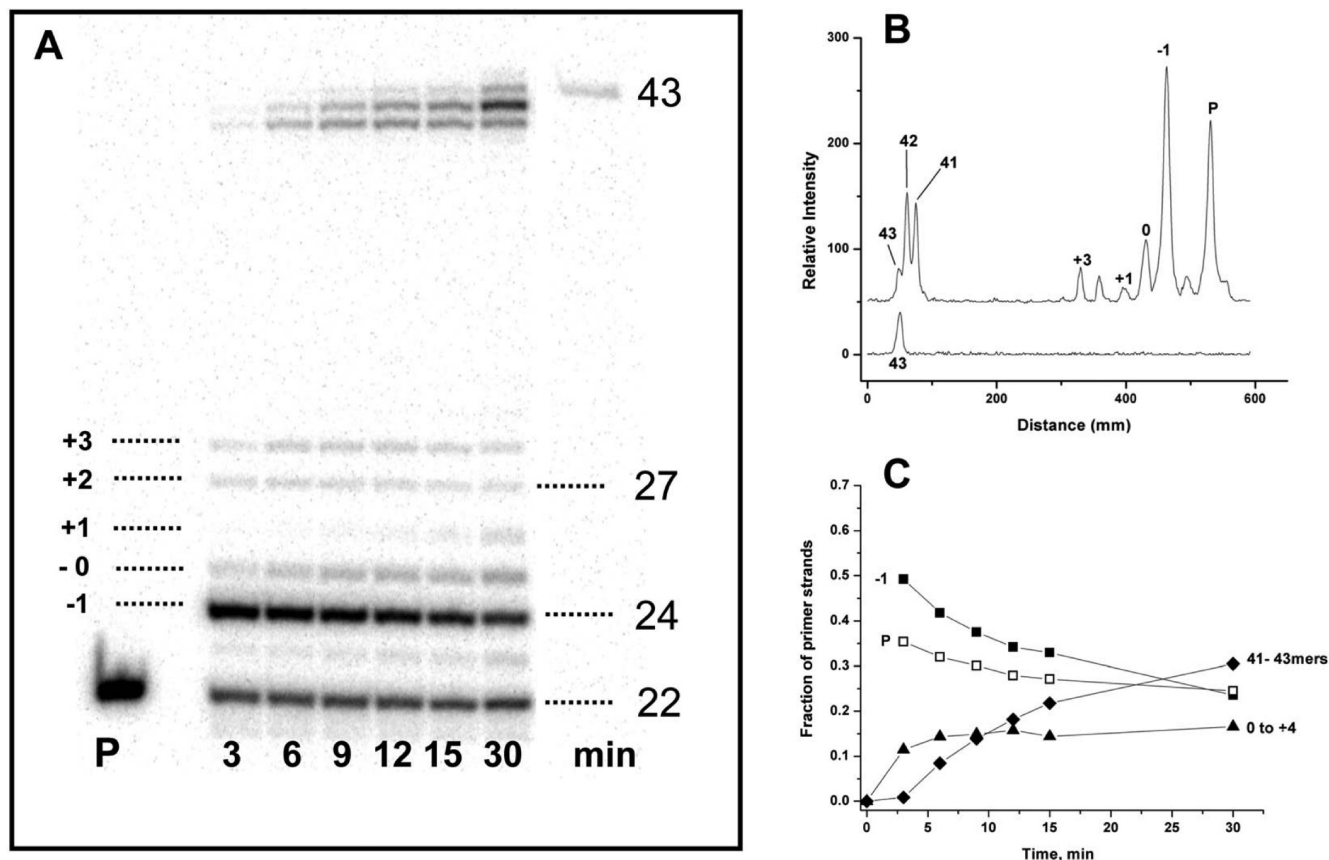


FIGURE 3: (A) Running-start replication of CG*G 43-mer template catalyzed by Dpo4. The primer strands (22-mers) that contained internally labeled C residue were annealed to 43-mer templates (4.5 nM) with the G* lesion at position 25 (labeled 0). The buffer solutions also contained dATP, dGTP, dTTP, and dCTP (100 μ M each) and 10 nM Dpo4. The incubation times were 3, 6, 9, 12, 15, and 30 min at 37 °C. The 22-mer primer strand and a 32 P-end-labeled 43-mer (corresponding to a fully extended primer) are shown as markers in the left- and right-most lanes, respectively. The partially extended primer band marked -1 corresponds to a 24-mer; the band at 0 is a 25-mer with its 3'-terminal base opposite G*, while +1 is a 26-mer, +2 a 27-mer, and +3 a 28-mer primer extension product. (B) Densitometry trace of the 15 min lane in panel A. (C) Time course of the products indicated in panel B: (□) 22-mer unextended primer strands (P), (■) partially extended 24-mer (-1) primer strands, (▲) sum of 0 \rightarrow +4 extension products, and (◆) sum of 41-mer \rightarrow 43-mer extension products.

primer-template duplexes. Therefore, it is unlikely that the excess 22-mer primers can efficiently displace any of the extended primer strands during the 30 min incubation time, thus accounting for the slow depletion of the 22-mer primer strands during the last 15 min (Figure 3C).

The most abundant and persistent extension product is the 24-mer (labeled -1 in Figure 3A) which reaches a level of ~50% within the first 3 min and then slowly decreases to 25% after 30 min (Figure 3C). The persistence of the -1 band during the entire 30 min incubation period indicates that extension from the 22-mer to the 24-mer primer strand is significantly faster than the apparent insertion of a nucleotide opposite position 0 that defines the position of an extended 25-mer. This is due to the slow -1 to 0 nucleotide incorporation step. In contrast, when the incubation temperature is increased to 55 °C (Figure S6 of the Supporting Information), the -1 band intensity decreases substantially after a 15 min incubation time and beyond, although the 22-mer primer band levels exhibit little change. This is consistent with an acceleration of the bypass of the [BP]G* lesions and higher reaction rates at the higher temperature (57), but without a significant depletion of the 22-mer primer strand for the reasons noted above.

Since the -1 band is the strongest among all primer extension bands beyond the 22-mer primer band, we conclude that the -1 \rightarrow 0 dNTP step is the slowest in the series of steps that ultimately lead to 41-42-mer nucleotide extension products for

incubation times of up to 15 min, and a faint 43-mer, fully extended primer band can be observed at the 30 min time point (Figure 3A,B). The 41- and 42-mer bands observed stem from a -1 deletion TLS step, as discussed further below. Faint bands are also observed up to the +1, +2, and +3 primer extension products, indicating that the adduct disturbs the primer-template structure at least up to template position +3. The time evolution of selected primer extension products, presented in Figure 3C, shows the slow accumulation of 41-43-mer extension products. The product profiles for the 12 and 30 min incubation time intervals are shown in the Supporting Information (Figure S5). At short incubation times, the 41-mer is dominant, with smaller amounts of 42-mers and 43-mers. However, at the longest reaction times, the 42-mer products are dominant, with smaller amounts of 41-mers, and even smaller relative amounts of 43-mers.

The single-step standing-start primer extension assays (Figure 4A) reveal that the order of dNTP insertion in the slowest -1 \rightarrow 0 step (opposite the [BP]G* lesion) is dGTP \gg dATP $>$ dCTP \gg dTTP.

Sequencing of Translesion Bypass Products (41-43-mers) Reveals the Prevalence of -1 Deletion Products. In the sequencing experiments, the 41-43-mer primer extension products [the bands shown at the top of the gel (Figure 3A), 30 min time point] were eluted from the gel. The samples were then analyzed by the two-phase gel sequencing approach after the extended primer-template duplexes had been cleaved

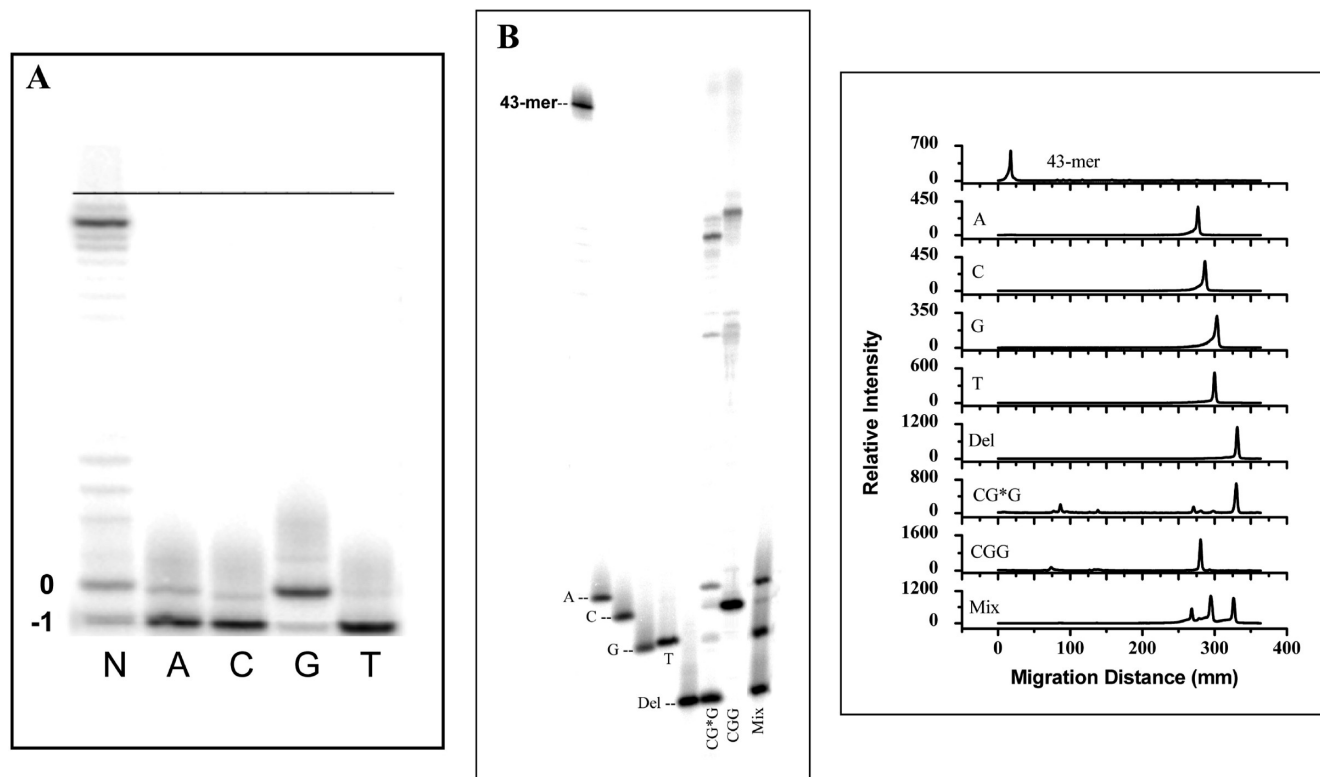


FIGURE 4: (A) Single-dNTP insertion opposite G* ($-1 \rightarrow 0$ step) catalyzed by Dpo4. Incorporation of each individual dNTP opposite G*, where A is dATP, C is dCTP, G is dGTP, T is dTTP, and N is a mixture of all four nucleotides (4 dNTPs). The dNTP concentrations were 500 μ M for each dNTP, 5 nM for the 24-mer primer–template complex, and 2 nM for Dpo4 at 37 $^{\circ}$ C; the reaction time was 5 min. (B) Two-phase gel analysis of the primer extension products (41–43-mers) catalyzed by Dpo4 using unmodified CGG and modified CG*G 43-mer template strands. Migration of oligonucleotide standards 3'-TGXCTGTGCGATAGAC-5' where X is A, C, G, T, or Del (-1 , missing X, a 15-mer). Primer strands of unknown composition derived from the replication of the 43-mer template strands (mixture of 41–43-mer primer extension products) with CGG or CG*G (lanes labeled CGG and CG*G, respectively). Lane 'Mix': mixture of all five oligonucleotide standards. (C) Densitometry tracings of the two-phase gel analysis of the primer extension products catalyzed by Dpo4 using unmodified CGG and modified CG*G 43-mer template strands in panel B.

by restriction enzymes as described in Materials and Methods. The central 16-nucleotide portion of the extended primer strands (restriction enzyme fragments) thus isolated contains the nucleotides inserted opposite the [BP]G* adduct (if any) and the neighboring nucleotides (45).

A typical gel comparing the electrophoretic mobilities of the restriction enzyme fragments to those of the 16-mer standards where Z is A, C, G, or T, or the -1 deletion product (15-mer) is shown in Figure 4B, and the densitometry tracings of each lane are shown in Figure 4C [there were no bands shorter than the 15-mer shown, thus ruling out the existence of -2 deletion products (45)]. The 43-mer band shown at the top of the gel in Figure 4B (left corner) is an oligonucleotide standard with the same sequence as that of a normal full extension 43-mer product. The lanes labeled A, C, G, and T are size marker 16-mer sequences (Figure 2C) where Z is A, C, G, and T, respectively. These sequences are markers for the 16-mer restriction fragments derived from the 41–43-mer primer extension bands that arise from the translesion bypass of G* in the 43-mer CG*G template strand. The band labeled Del represents the 15-mer primer with Z missing altogether and thus is a standard for a -1 deletion primer extension product. The lane labeled "Mix" contains all five standards. It is evident that the 16-mer sequences for which Z is G or T cannot be distinguished.

The lane labeled CGG in Figure 4B shows a control experiment conducted with an unmodified template strand CGG, while the lane marked CG*G shows the restriction products obtained from the duplex shown in Figure 2B. The lower bands (27-mers)

represent the unsuccessful cleavage catalyzed by restriction enzyme *Tsp45I* and are observed in both the case of the modified and the unmodified CGG and CG*G duplexes ($\leq 20\%$ of the total). Weaker 32-mer bands can be observed in both lanes that correspond to single cleavage events by the *Afl* restriction enzyme site. The CGG lane in Figure 4B shows a dominant single band that comigrates with the 16-mer for which Z is C, as expected for this control sample. However, in the CG*G lane, the dominant product is a band that comigrates with the 15-mer deletion standard (Del). Thus, the dominant TLS product catalyzed by Dpo4 in the case of the [BP]G* adduct in the CG*G sequence is a -1 frameshift mechanism leading to the (-1) Del product (57% of the bypass products). This was confirmed by Sanger sequencing experiments (Figure S7 of the Supporting Information) of the individual 41- and 42-mer primer extension products (Figures S8 and S9, respectively, of the Supporting Information). These results clearly show that the sequences of the 41- and 42-mers are identical in that both are missing a nucleotide that would have been inserted opposite the [BP]G* adduct. Only the lengths of these two extended primer strands are different. The 43-mer strand was also sequenced by the Sanger method, but the sequencing lane patterns were more complex (Figures S10 and S11 of the Supporting Information) than in the case of the 42- or 41-mers. Two series of bands are recognized. The weaker series corresponds to a -1 deletion and is probably due to the -1 deletion 42-mer product that could not be entirely separated from the 43-mer during the gel excision of the 43-mer band. The stronger series of bands (Figure S11 of the

Supporting Information) corresponded to the 43-mer extension product with an A inserted opposite the lesion G* during TLS. These Sanger sequencing results are in qualitative agreement with the two-phase gel sequencing results of the mix of 41-, 42-, and 43-mer extension products (Figure 4B,C). A quantitative analysis of the data in Figure 4C demonstrates that the different Z nucleotides (Figure 2) are inserted “opposite” the [BP]G* adduct in running-start experiments in which Z is A (10%), C (6%), and G or T (5%). The single-step insertion experiments suggest that the insertion of dGTP is significantly more efficient than that of dTTP (Figure 4A); we therefore hypothesize that the 16-mer for which Z is G dominates over products for which Z is T.

The single-step nucleotide insertion experiments shown in Figure 4A cannot distinguish between a mechanism that involves the pairing of dGTP opposite [BP]G*, followed by nucleotide incorporation, and one involving the pairing of the dGTP with the 5'-flanking C template residue, followed by nucleotide incorporation. However, the two-phase gel sequencing experiments depicted in Figure 4B indicate that the latter slipped frameshift or primer relocation mechanism is dominant, although the dGTP:[BP]G* paired intermediate cannot be excluded since G is found to be incorporated opposite the adduct (Figure 4B,C). However, a transient misalignment mechanism (58, 59) in which the dGTP pairs with the 5'-flanking template C, followed by realignment and then further primer extension from a dGTP:[BP]G* mismatch, can also account for the observed substitution errors opposite the lesion (58).

Molecular Modeling and Dynamics Simulations: Connecting Structure to Function over an Entire Replication Cycle. Our previous modeling studies suggested that a change in the glycosidic bond torsion angle at the [BP]G* site from a normal *anti* conformation with the BP moiety on the minor groove side of the evolving duplex to a *syn* conformation with BP on the major groove side may play a role in translesion bypass catalyzed by BF (15) and that both *anti* and *syn* conformations are feasible at the insertion step in a different sequence context in Dpo4 (47). In addition, both *anti* and *syn* conformations have been recently observed with a bulky adduct in Dpo4 (38). In the work presented here, we again employed molecular modeling and dynamics simulations to interpret, on a molecular level, the results of Dpo4-catalyzed translesion bypass experiments of the [BP]G* adduct in a 5'-CG*G-3' sequence context, investigating both *anti* and *syn* conformations of [BP]G*. The experimental observations indicate the prevalence of relatively easy bypass of [BP]G*

via a -1 deletion pathway, and a less prevalent nonslippage pathway involving incorporation of dNTP opposite the [BP]G* adduct with dATP > dCTP > dGTP.

To analyze possible intermediates that can account for these results, we constructed different models that could explain the experimental observations. We define slippage and nonslippage models. In the nonslippage models, the incoming dNTP pairs with [BP]G* and the 3'-terminal base pairs with the unmodified template strand base G in CG*G. We investigated two types of slippage models: (a) one in which the incoming dGTP pairs with C in CG*G and the 3'-terminal primer base C pairs with G* instead of with the unmodified G that is skipped altogether (Figure 5A) and (b) another in which dGTP pairs in the same manner as in model (a), but G* is skipped (Figure 5B) (46). All slippage and nonslippage models constructed contained the dG* lesions in either the *anti* or the *syn* conformations. Specifically, as illustrated in Figure 1B, the models included (1) a binary complex (DNA and Dpo4 only) with the 3'-terminal primer base C opposite G in CG*G, termed the “pre-insertion step”; (2) the incoming dNTP paired with either [BP]G* (non-slippage models), or with C in the CG*G sequence context in either pairing scheme A or B of Figure 5 (slippage models), termed the “insertion step”; (3) the terminal 3'-primer base paired with [BP]G* in the absence of any dNTP (binary complex “pre-extension step”); and (4) same as (3), but with a dNTP opposite the template C in CG*G (ternary complex “extension step”). These are the four consecutive steps as the [BP]G* adduct threads through the Dpo4 active site and correspond to steps 1–4, respectively, of our earlier study with the replicative A-family polymerase BF (15).

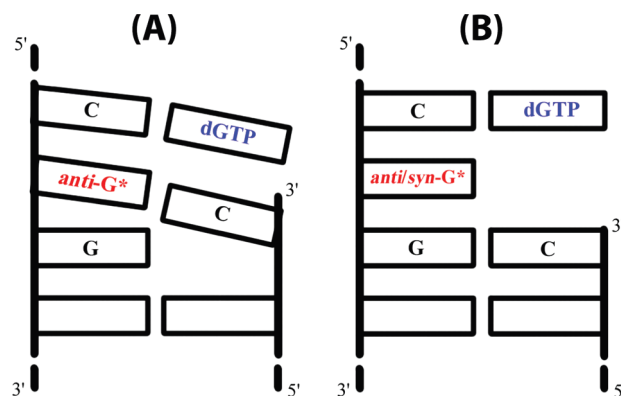


FIGURE 5: Patterns of (A) 5'-strand slippage and (B) 3'-strand slippage.

(A) *anti*-[BP]G*-5'-slippage

(B) *anti*-[BP]G*-3'-slippage

(C) *syn*-[BP]G*

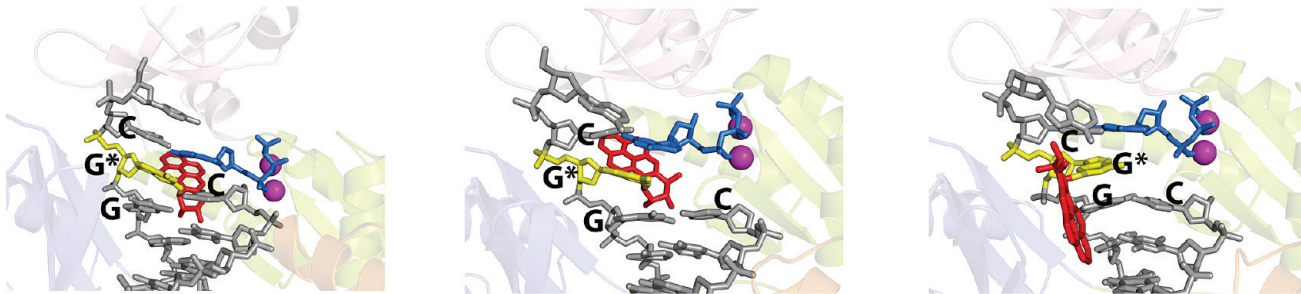
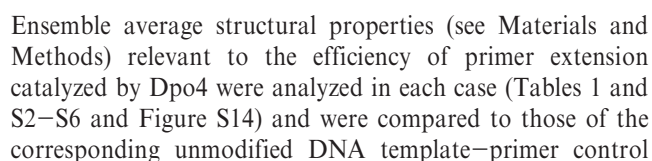


FIGURE 6: Simulated (A) *anti*-[BP]G* 5'-slippage, (B) *anti*-[BP]G* 3'-slippage, and (C) *syn*-[BP]G* slippage systems after 2.5 ns of MD. Color code: gray, DNA strands; magenta, Mg^{2+} ions (sphere representation); yellow, modified dG; red, BP moiety; blue, dGTP. The CG*G template sequence and the primer terminal C are indicated in each panel. The primer terminal C pairs with G* in panel A but with G in panels B and C. Note the shorter O3'–P α distance in panel A vs that in panels B and C (Table 1). Stereoviews of these adducted structures and the unmodified control are given in Figures S12 and S13 of the Supporting Information.



complexes. The structural factors specifically considered are (1) the number and occupancies of the hydrogen bonds in the [BP]G*-containing base pair and [BP]G*'s neighboring base pairs (Table S2) [hydrogen bond occupancies are the percent of time during the stable region of the MD trajectory (from 1.0 to 2.5 ns) that the hydrogen bond is present according to the following criteria: a distance of ≤ 3.3 Å between heavy atoms (donor and acceptor) and a hydrogen bonding donor–hydrogen–acceptor angle of $180 \pm 45^\circ$]; (2) the presence of hydrogen bonds between amino acid residues and incoming dNTP that stabilize

Table 1: P α –O3' Distances That Are Near Reaction-Ready in the 5'-Slippage Pattern, but Not in the 3'-Slippage Pattern

slippage model	ensemble average of the P α –O3' distance (Å)
<i>anti</i> -[BP]G* 5'-slippage	3.7 ± 0.2
<i>anti</i> -[BP]G* 3'-slippage	7.5 ± 0.6
<i>syn</i> -[BP]G* 3'-slippage	6.9 ± 1.0

nucleotide binding in the ternary complexes (Table S3); (3) the C1'–C1' distance (~ 10.8 Å) of the 0 and +1 base pairs (Table S4); (4) the frequency of sampling a near reaction-ready distance (3.1–3.5 Å) between P α of the dNTP and O3' of the 3'-terminal primer base in ternary complexes (Table S4); (5) the attack angle in the ternary complexes, formed by O (primer 3'-end), P α (dNTP), and O3 α (dNTP) [for in-line attack, this angle should be close to 180° (Table S4)]; and (6) the distance between the two Mg²⁺ ions (Table S4) and the quality of their octahedral coordination (Table S5) in the ternary complexes. These criteria are based on a well-organized polymerase crystal structure (60) and QM/MM calculations (61).

Slippage Models Explain the Strong dGTP Selectivity in the Predominant –1 Deletion Pathway for Dpo4. The experimental results indicate that dGTP insertion opposite [BP]G* and –1 deletion products are dominant in the 42-mers and fully extended 43-mer primer strands. To account for these observations, we constructed two *anti*-dG* adduct models

Table 2: Composite Evaluation of Structural Distortions in the Nonslippage Pathway^a

		base pair hydrogen bonding geometry ^b	stacking interaction ^c	protein– dNTP interaction ^d	0 base pair C1'–C1' ^e	+1 base pair C1'–C1' ^f	P α –O3' ^g	Mg ²⁺ ions: distance and coordination ^h	composite score
pre-insertion step	<i>anti</i> -[BP]G*	0	0		0	0			0
	<i>syn</i> -[BP]G*	0	0		0	0			0
insertion step	<i>anti</i> -[BP]G*: <i>anti</i> -dATP	–1	0	–1	–1	0	0	–1	–4
	<i>anti</i> -[BP]G*: <i>syn</i> -dATP	–1	0	0	0	0	0	0	–1
	<i>syn</i> -[BP]G*: <i>anti</i> -dATP	–2	–1	–1	0	0	0	–1	–5
	<i>anti</i> -[BP]G*: <i>anti</i> -dGTP	–2	0	0	–1	0	0	0	–3
	<i>anti</i> -[BP]G*: <i>syn</i> -dGTP	–1	0	–2	–1	0	0	0	–4
	<i>syn</i> -[BP]G*: <i>anti</i> -dGTP	–2	–1	–1	0	0	0	0	–4
	<i>anti</i> -[BP]G*:dCTP	–2	0	0	0	0	0	0	–2
	<i>syn</i> -[BP]G*:dCTP	–3	–1	0	0	0	0	0	–4
	<i>anti</i> -[BP]G*:dTTP	–2	0	–2	0	0	0	0	–4
	<i>syn</i> -[BP]G*:dTTP	–3	–1	0	–1	0	0	0	–5
pre-extension step	<i>anti</i> -[BP]G*: <i>anti</i> -dA	–3	–1		–1	0			–5
	<i>anti</i> -[BP]G*: <i>syn</i> -dA	–1	0		–1	0			–2
	<i>syn</i> -[BP]G*: <i>anti</i> -dA	–3	–1		–1	0			–5
	<i>anti</i> -[BP]G*: <i>anti</i> -dG	–2	0		0	0			–2
	<i>anti</i> -[BP]G*: <i>syn</i> -dG	–1	0		–1	0			–2
	<i>syn</i> -[BP]G*: <i>anti</i> -dG	–3	–1		–1	0			–5
	<i>anti</i> -[BP]G*:dC	0	0		0	0			0
	<i>syn</i> -[BP]G*:dC	–2	–1		0	0			–3
	<i>anti</i> -[BP]G*:dT	0	0		0	0			0
	<i>syn</i> -[BP]G*:dT	–2	–1		–1	0			–4
extension step	<i>anti</i> -[BP]G*: <i>anti</i> -dA	–1	–1	0	0	–1	0	0	–3
	<i>anti</i> -[BP]G*: <i>syn</i> -dA	–1	0	–1	0	0	0	0	–2
	<i>syn</i> -[BP]G*: <i>anti</i> -dA	–2	–1	0	0	0	0	0	–3
	<i>anti</i> -[BP]G*: <i>anti</i> -dG	–3	–1	–1	0	0	0	0	–5
	<i>anti</i> -[BP]G*: <i>syn</i> -dG	–1	–1	0	0	0	–5	0	–7
	<i>syn</i> -[BP]G*: <i>anti</i> -dG	–1	–1	0	0	0	0	0	–2
	<i>anti</i> -[BP]G*:dC	–2	0	0	–1	0	–5	0	–8
	<i>syn</i> -[BP]G*:dC	–2	–1	0	0	0	0	0	–3
	<i>anti</i> -[BP]G*:dT	0	0	–1	0	0	0	0	–1
	<i>syn</i> -[BP]G*:dT	–1	–1	0	0	–1	0	0	–3

^a The unmodified Watson–Crick paired controls were assigned a base score of 0 for each structural property. ^b A score of –1 was applied for a decrease in hydrogen bond occupancy compared to the control by more than 50%; in addition, in the case of mismatches, a score of –1 was applied for each hydrogen bond that was lacking compared to the three in the unmodified control. See Table S2 and Results in the Supporting Information for details. ^c A score of –1 was applied for a severe disruption of the base stacking interactions at and near the lesion site. See Figure S13 and Results in the Supporting Information for details. ^d A score of –1 was applied for a decrease in occupancy compared to the control by more than 30%. See Table S3 for details. ^e C1'–C1' distance of the 0 base pair (Figure 1B). A score of –1 was applied when the distance deviated from the control by more than 1 Å. See Table S4 for details. ^f C1'–C1' distance of the +1 base pair (Figure 1B). A score of –1 was applied when the distance deviated from the control by more than 1 Å. See Table S4 for details. ^g A score of –5 was applied for a zero occupancy in the near reaction-ready distance of 3.1–3.5 Å. See Table S4 for details. ^h For Mg²⁺ ion coordination, a score of –1 was applied for an increase in any of the coordination distances compared to the control by more than 0.5 Å. See Table S5 for details. For the Mg²⁺–Mg²⁺ ion distance, no disruptions in the distance between the two Mg²⁺ ions were observed; an increase in the distance compared to the control by more than 0.5 Å is considered to be a disruption. See Table S4 for details.

(*anti*-[BP]G* 5'-slippage and *anti*-[BP]G* 3'-slippage) (Figure 5A, B) and one *syn*-dG* (*syn*-[BP]G*) model. As described in detail below, we found that the 5'-slippage pattern with the unmodified G skipped (Figure 5A) favors insertion of dGTP opposite the C 5' to the adduct.

The *anti*-[BP]G* 5'-slippage system starts with a structure in which the primer terminal C pairs with G*, thus skipping the unmodified G in the CG*G sequence (Figure 5A); during the MD, this C mostly remains paired with G* and is not observed to pair with G (Table S6 and Figure 6A). The P α -O3' distance has an ensemble average value of 3.7 ± 0.2 Å (Tables 1 and S6). The C1'-C1' distance, the Mg²⁺-Mg²⁺ distance, and the Mg²⁺ coordination are near normal (Table S6).

Both the *anti*- and *syn*-[BP]G* 3'-slippage systems start with structures in which the 3'-terminal primer C pairs with the unmodified G in CG*G (Figure 5B); during the MD, this C remains paired with G and is never observed to pair with G* (Table S6 and Figure 6B,C). However, the P α -O3' distance is never found to be within the nearly reaction-ready range in either of these two systems: the ensemble average values are 7.5 ± 0.6 and 6.9 ± 1.0 Å for the *anti*- and *syn*-[BP]G* 3'-slippage systems, respectively (Tables 1 and S6).

Since the P α -O3' ensemble average distance is near the reaction-ready range of 3.7 ± 0.2 Å in the *anti*-[BP]G* 5'-slippage system, we conclude that the 5'-slippage pattern with G instead of G* skipped clearly favors insertion of dGTP opposite the C in the CG*G sequence. The incorporation of dGTP into the primer strand with subsequent primer extension from this Watson-Crick C:G pair yields the -1 deletion products.

Structural Distortions throughout the Replicative Cycle in Nonslippage Models Explain Nucleotide Insertion Preferences opposite [BP]G in This Minor Pathway.* The experimental results indicate that besides the predominant -1 deletion products, the fully extended primers also contain products without deletion mutations: 10% G*:A mismatch, 6% G*:C pair, and 5% G*:G mismatch. To gain insight into the underlying mechanisms, we have constructed nonslippage models at the pre-insertion step, insertion step, pre-extension step, and extension step (Figure 1B), with all four partners opposite the [BP]G* adduct in steps 2–4. Distortions of the structural properties for all the nonslippage models were scored according to the criteria described above and are given in Table 2: a more negative score indicates a more distorted structure. Our current best understanding of reaction-ready polymerase active site geometry was utilized in selecting the properties evaluated and the scoring criteria employed. A well-organized active site of a recent crystal structure (60) and QM/MM calculations reveal (61) features that closely resemble our unmodified models, suggesting that our criteria for evaluating distortions are reasonable; although the relationship between the scores and the biological outcome is likely nonlinear, the reasonably good correlation between the experimental data and our scoring results indicates that this approach is useful.

All *anti* structures place the BP lesion on the minor groove side of the evolving duplex while the BP rings are on the major groove side in the *syn* structures. At step 1, the pre-insertion step, both *anti*- and *syn*-[BP]G* models have a composite score of 0, suggesting that the adduct in the single-strand overhang causes negligible distortion to the structure. At step 2, the insertion step, the *anti*-[BP]G*:*syn*-dATP system (composite score of -1) appears to be most favorable for nucleotide incorporation, followed by the *anti*-[BP]G*:dCTP system

(composite score of -2) and then the *anti*-[BP]G*:*anti*-dGTP system (composite score of -3). The other systems appear to be too distorted for the nucleotidyl transfer reaction to occur. At step 3, the pre-extension step, the *anti*-[BP]G*:dC and *anti*-[BP]G*:dT structures, both with a composite score of 0, appear to be most feasible, followed by the *anti*-[BP]G*:*syn*-dA, *anti*-[BP]G*:*anti*-dG, and *anti*-[BP]G*:*syn*-dG systems (all with a composite score of -2), and then the *syn*-[BP]G*:dC system (composite score of -3). In the other systems, the structures are distorted to a more significant extent and appear to be unfeasible. At step 4, the extension step, the *anti*-[BP]G*:dT system (composite score of -1) appears to be most favorable for primer extension beyond the lesion site, followed by the *anti*-[BP]G*:*syn*-dA and *syn*-[BP]G*:*anti*-dG systems (both with a composite score of -2), and then the *anti*-[BP]G*:*anti*-dA, *syn*-[BP]G*:*anti*-dA, *syn*-[BP]G*:dC, and *syn*-[BP]G*:dT systems (all with a composite score of -3). The remaining systems appear to be unfavorable for a successful nucleotidyl transfer reaction beyond the lesion site. The least distorted structures for each step are given in Figure 7. A full description of the structural properties of each of these systems is provided in the Supporting Information.

DISCUSSION

5'-Slippage with Unmodified G Skipped Can Account for Base Sequence Context Effects and the Predominant -1 Deletion Mutation. The 5'-slippage pattern indicated by our modeling studies, in which the unadducted G instead of the modified G* in the 5'-CG*G-3' sequence context is skipped (Figure 5A), allows a favorable active site alignment (Tables 1 and S6), which supports insertion of dGTP opposite the C base flanking [BP]G* on its 5'-side. This is consistent with a previous study utilizing the TG*G sequence in the Dpo4 polymerase which revealed strong dATP selectivity; in that study, the same 5'-slippage pattern was suggested to favor dATP insertion opposite the T base flanking [BP]G* on its 5'-side (45, 46). Therefore, the slippage mechanism we have proposed may be applicable to sequences with a repetitive G*G sequence theme as in CG*G/TG*G. Both CG*G and TG*G are prone to slippage because of the tandem GG motif; the mechanism accounts for sequence-dependent nucleotide selectivity which relies on the identity of both the 3'- and 5'-bases flanking G*. In contrast to our findings that dGTP is strongly selected over the other nucleotides in the CG*G sequence context, Perlow-Poehnelt et al. reported that all four dNTPs are promiscuously incorporated opposite G* in an AG*C sequence context under conditions of saturating dNTP concentration (47). The nonrepetitive AG*C sequence context does not support the current 5'-slippage pattern, and such slippage does not appear to occur at the nucleotide insertion step in this sequence context.

Our 5'-slippage mechanism differs from the one described by Bauer et al. (39) that was based on crystallographic and primer extension studies with the damaged base in a different sequence context (5'-TG*A-3'), without contiguous guanines and with the G* in the post-insertion rather than the insertion position; in this case, it is the adducted base that is skipped. We note that the mechanism proposed by our simulations does not preclude the possibility that G* may be skipped; the looping out of the adduct G* is a general mechanism that can occur in all sequence contexts, and it allows for a relocation of the terminal 3'-primer nucleotide from the next-to-be-replicated G* to the

5'-flanking unmodified template base. Our results suggest that the exact primer frameshift misalignment mechanism can also depend on the nucleotide flanking the adduct G* on its 3'-side, as well as the base on its 5'-side. The misalignment mechanism in which the G on the 3'-side of G* is skipped and the terminal 3'-primer nucleotide relocates toward the 5' direction appears to be favored for the 5'-CG*G sequence considered here and for the TG*G sequence investigated previously (46). Various strand slippage mechanisms as mutational intermediates have been extensively studied by Ripley (59) and by Kunkel (58, 62). The relevance of these types of mechanisms to *in vivo* mutagenesis is a topic of considerable recent interest (63).

To explain that the deletion product is the predominant one among the fully extended primers, we infer that following dGTP insertion opposite the C, extension beyond this unmodified Watson–Crick C:G pair is facile despite the presence of a skipped base on the template strand. Primer extension beyond a G*:N base pair, which would yield full extension products, appears to be less favored in the experiments reported here, most likely because any G*:N base pair is more disturbing than a Watson–Crick C:G pair even if the primer strand is misaligned because the G that is 3' to the lesion G* is skipped.

Varying Preferences for Partners opposite the [BP]G Adduct in the Nonslippage Pathway Are Explained by Relative Distortions of the Models.* The strong preference for dGTP insertion occurs primarily through the predominant -1 deletion pathway, with the minor contribution by G*:G mismatch mutations; the preference for the other nucleotides then follows the order dATP > dCTP \gg dTTP. The sequencing analyses of the fully extended primers reveal that among the products without a -1 deletion, the G*:A mismatch is the most prevalent, followed by the G*:C pair and the G*:G mismatch, while productive G*:T mismatches are least likely. The composite scores in Table 2 reflect the extent of structural distortion for each nonslippage G* system. There are multiple conformational possibilities for each pair starting from step 2, the nucleotide insertion step. For each pairing scheme, the system with the least distortion is characterized by the least negative score. At step 2, the *anti*-[BP]G*:*syn*-dATP mismatch (composite score of -1) is modestly distorted with one of three hydrogen bonding interactions between G* and A disturbed. For the *anti*-[BP]G*:dCTP case, distortion is greater (composite score of -2) with two of the three Watson–Crick hydrogen bonds between G* and C disturbed. This is consistent with the experimental result in which there is greater insertion of dATP opposite the G* than insertion of dCTP. Neither of the two [BP]G*:dTTP systems (scores of -4 and -5) appears to be favorable for nucleotide incorporation. In the *anti*-[BP]G*:dTTP system, both the interactions between dTTP and its partner G* and the interactions stabilizing dTTP binding are disrupted. In the *syn*-[BP]G*:dTTP system, the nascent G:T base pair at the active site is severely distorted: the dTTP forms no hydrogen bonds with G*, the C1'–C1' distance of this base pair is significantly enlarged, and the G* stacks poorly with its neighboring base. These modeling results are consistent with the experimental observations that explain dTTP insertion is not detectable (Figure 4).

Conformational Pathways through the Replicative Cycle May Utilize anti–syn Transitions of [BP]G.* The scoring system in Table 2 can be used to examine the feasibility of the conformational pathways for all four steps of the replicative cycle we investigated. Figure 7 summarizes the pathways that are most preferred according to Table 2. We note that in certain cases, a

conformational change from *anti* to *syn* around the glycosidic bond of the dG* favors the progression through the cycle. Such an *anti*–*syn* conformational transition was observed in a BF crystal structure containing the bulky [AF]G* adduct; the AF-modified dG* was in a *syn* conformation at the preinsertion site, but after one round of nucleotide incorporation within the crystal, the modified dG* rotated to the *anti* conformation that placed the AF residue into the open major groove side in the post-insertion site, while the *syn* conformation would have placed the lesion in the crowded minor groove scanning track (21). In addition, recently, both *anti* and *syn* conformations of the glycosidic bond in the bulky adduct *N*²-CH₂(2-naphthyl)G have been observed crystallographically in Dpo4 (38) and had also been predicted for [BP]G* in Dpo4 (47). Furthermore, we mention in this connection that during replication the mobile polymerase machinery allows flexibilities that are not manifested in the static snapshots of crystal structures. In our cases, such *anti*–*syn* rearrangements appear to be necessary to avoid highly distorted structures. For example, in the *anti*-[BP]G*:dC system at step 4, the extension step, the BP residue pushes the incoming dGTP away from its normal position (Figures S12, #37, and S13, #37), resulting in a poor P α (dGTP)–O3' (primer end) distance (Table S4); however, the *syn*-[BP]G* conformation appears to be feasible for dGTP insertion (Table S4). Similarly, both *anti*-[BP]G*:*anti*-dG and *anti*-[BP]G*:*syn*-dG systems are much more distorted than the *syn*-[BP]G*:*anti*-dG system at step 4, the extension step, and the *syn*-[BP]G* conformation appears to be feasible for the nucleotidyl transfer reaction. Thus, bypass of the G*:C pair and the G*:G mismatch may involve an *anti*–*syn* conformational change of the BP-modified dG*, while the G*:A mismatch does not. This would make bypass of the G*:A mismatch easier than bypass of the G*:C pair or the G*:G mismatch, as observed. As discussed earlier, neither of the two [BP]G*:dTTP systems at step 2 (the insertion step) appears to be favorable for nucleotide incorporation, explaining why virtually no product containing a G*:T mismatch is detected (Figure 4). However, our results in Table 2 do indicate that if any dTTP were incorporated, extension would be facile.

Lesion Bypass in Dpo4 Is Facilitated by Slippage, while Nonslippage-Prone BF Predominantly Blocks the Progress of the Polymerase past the Lesion. In this work, our experimental studies have revealed that Dpo4 allows primer extension past the [BP]G* lesion site to a significant extent; bypass of the [BP]G* in the current CG*G sequence context is achieved through a major -1 deletion pathway and a minor nonslippage pathway. Single-nucleotide insertion assays in the CG*G sequence context, together with a previous study in a TG*G sequence context, reveal strong nucleotide selectivities: dGTP is strongly favored in CG*G, and dATP is significantly preferred in the TG*G sequence contexts. In this predominant -1 deletion pathway, insertion of the nucleotide opposite the base on the 5'-side of the adduct involves the skipping of the template strand base G (Figures 5A and 6A), thus introducing a significant structural disturbance that hinders extension from the $+1$, $+2$, and $+3$ sites (Figure 3A,B). In addition, the $0 \rightarrow +1$ step in Dpo4 is faster than the slowest $-1 \rightarrow 0$ step (Figure 3A). However, processing of the [BP]G* lesion in the CG*G sequence context by an A-family high-fidelity DNA polymerase, the *Bacillus* fragment (BF), is strikingly different, as shown in our earlier work (15). In this polymerase, insertion of the nucleotide opposite the adduct is allowed with the preferential insertion of purines (dATP and dGTP), but further extension is strongly inhibited with

Table 3: Replicative Activities past the [BP]G* Adduct by Dpo4 and BF

	Dpo4	BF
running-start primer extension	relatively facile bypass through a predominant -1 deletion pathway and a less prevalent nonslippage pathway; further extension beyond the adduct more facile than the nucleotide incorporation opposite the lesion	mainly blocked with a small amount of bypass; nucleotide insertion opposite the adduct allowed, but further extension strongly inhibited
standing-start single-dNTP insertion	strong base sequence context effect: dGTP preferred in CG*G and dATP preferred in TG*G	no base sequence context effect: purines favored over pyrimidines in both CG*G and TG*G

only $\sim 1\%$ full extension observed (15). In addition, the template base flanking the adduct on the 5'-side did not alter the selectivity of insertion of purine dNTP opposite the [BP]G* adduct when catalyzed by BF (45). The different replicating activities of these two polymerases for this adduct are summarized in Table 3.

These differences result from the structural distinctions of the two enzymes, especially in the active site region. Dpo4 has a spacious active site which can accommodate two templating bases simultaneously and thus can readily allow slippage to occur (25). The observed nucleotide selectivities in Dpo4 result from a realignment or relocation of the 3'-end primer base so that dGTP or dATP pairs with template C or T, respectively, flanking the [BP]G* adduct on the 5'-side; the subsequent nucleotidyl transfer reaction in such slippage structures involves extension from a normal Watson–Crick C:G or T:A pair. The minor nonslippage pathway in Dpo4 is more facile than the analogous primer extension in BF (15). At step 1 (Figure 1B), both *anti*- and *syn*-dG* cause negligible distortion to the structure. From step 2 to 4, differing dG* *anti/syn* conformational preferences were deduced from the simulations depending on the partner base opposite the [BP]G* adduct (Table 2); Dpo4 can more easily sample a reaction-favorable glycosidic conformation of the dG* as well as the partner since a conformational *anti-syn* rearrangement would take place more readily in the spacious active site of Dpo4 than in BF.

Significantly different from Dpo4, BF has a much more confined active site with strand slippage much less likely to occur; this can account for the absence of sequence context-governed nucleotide selectivity in the experiments. In addition, the minor groove side of the DNA in BF is densely packed with amino acid residues, which constitute a scanning tract to ensure replication with high fidelity and processivity; any factor that disrupts this minor groove scanning tract inhibits the enzyme activity. The blocking of TLS activity of BF by [BP]G* is explained by the *anti* dG* conformation with BP on the crowded minor groove side, as shown in a crystal structure (24) and our modeling studies (15). The observed limited bypass can be explained by a model in which the [BP]G* adduct assumes the *syn* dG* conformation, placing BP on the open major groove side. The preference for the pairing of purine dNTPs opposite the adduct is attributed to a better accommodation of these dNTPs opposite the adduct in the *syn*-[BP]G* conformation which has a pyrimidine-like shape in the active site and allows for TLS (15). The *syn*-[BP]G* causes more structural distortion at the extension steps than the insertion step, explaining why further extension beyond the lesion site is more inhibited than nucleotide incorporation opposite the G* adduct in BF.

Overall, the slippage-prone structure of Dpo4 explains its predominant lesion bypass and sequence-governed nucleotide selectivity properties in the CG*G template sequence context. By contrast, the absence of sequence-governed nucleotide selectivity in BF and its strong lesion blocking properties stem from its

spatially tight and constrained active site that does not favor strand slippage.

CONCLUSIONS

Our modeling studies combined with primer extension experiments have provided new structural insights into the selective incorporation of dNTP opposite the lesion G* and the subsequent primer extension beyond the lesion over an entire Dpo4 replicative cycle in the CG*G sequence context. The selectivity of incorporation of dGTP results primarily from an unusual 5'-slippage mechanism in which the unmodified G rather than the G* is skipped (Figure 5A). Furthermore, the modeling provides structural understanding of the effect of the lesion on translesion synthesis, since the dNTP has two alternate slippage mechanism choices (skipping G or G*) that cannot be distinguished by kinetic experiments. In addition, our simulations suggest that the G* may undergo an *anti-syn* conformational transition during the replicative cycle which could be particularly feasible in the spacious Dpo4 active site region. By contrast, the poor extension following nucleotide incorporation opposite the lesion as well as the absence of sequence-governed nucleotide selectivity in the high-fidelity polymerase BF stems from its nonslippage-prone structural properties. By combining insights gained from simulations with experimental investigations, we achieve enhanced understanding of the structural origins for the functional observations. Simulating the lesion bypass through an entire replication cycle provides molecular views of the replicative stages in lesion processing that have not yet been achieved in crystal structures and are not kinetically observable.

ACKNOWLEDGMENT

We thank the reviewers for valuable suggestions.

SUPPORTING INFORMATION AVAILABLE

Details of molecular modeling protocols for the slippage models, detailed results for the molecular dynamics (MD) simulations of the nonslippage models, torsion angles of [BP]G* for all initial models, and detailed analyses of trajectory ensembles. Experimental results comparing (1) extension product profiles at 12 and 30 min of the results shown in Figure 3A, (2) running-start primer extension experiments at 37 and 55 °C, and (3) Sanger sequencing data of the 41-, 42-, and 43-mer extension products are also given. This material is available free of charge via the Internet at <http://pubs.acs.org>.

REFERENCES

- Pages, V., and Fuchs, R. P. (2002) How DNA lesions are turned into mutations within cells?. *Oncogene* 21, 8957–8966.

2. McCulloch, S. D., Kokoska, R. J., Chilkova, O., Welch, C. M., Johansson, E., Burgers, P. M., and Kunkel, T. A. (2004) Enzymatic switching for efficient and accurate translesion DNA replication. *Nucleic Acids Res.* 32, 4665–4675.
3. Friedberg, E. C., Lehmann, A. R., and Fuchs, R. P. (2005) Trading places: How do DNA polymerases switch during translesion DNA synthesis? *Mol. Cell* 18, 499–505.
4. Prakash, S., Johnson, R. E., and Prakash, L. (2005) Eukaryotic translesion synthesis DNA polymerases: Specificity of structure and function. *Annu. Rev. Biochem.* 74, 317–353.
5. Lehmann, A. R. (2006) Translesion synthesis in mammalian cells. *Exp. Cell Res.* 312, 2673–2676.
6. Yang, W., and Woodgate, R. (2007) What a difference a decade makes: Insights into translesion DNA synthesis. *Proc. Natl. Acad. Sci. U.S.A.* 104, 15591–15598.
7. Luch, A. (2005) Nature and nurture: Lessons from chemical carcinogenesis. *Nat. Rev. Cancer* 5, 113–125.
8. Conney, A. H. (1982) Induction of microsomal enzymes by foreign chemicals and carcinogenesis by polycyclic aromatic hydrocarbons: G. H. A. Clowes Memorial Lecture. *Cancer Res.* 42, 4875–4917.
9. Cavalieri, E. L., and Rogan, E. G. (1995) Central role of radical cations in metabolic activation of polycyclic aromatic hydrocarbons. *Xenobiotica* 25, 677–688.
10. Burczynski, M. E., Harvey, R. G., and Penning, T. M. (1998) Expression and characterization of four recombinant human dihydrodiol dehydrogenase isoforms: Oxidation of trans-7, 8-dihydroxy-7,8-dihydrobenzo[a]pyrene to the activated o-quinone metabolite benzo[a]pyrene-7,8-dione. *Biochemistry* 37, 6781–6790.
11. Buening, M. K., Wislocki, P. G., Levin, W., Yagi, H., Thakker, D. R., Akagi, H., Koreeda, M., Jerina, D. M., and Conney, A. H. (1978) Tumorigenicity of the optical enantiomers of the diastereomeric benzo[a]pyrene 7,8-diol-9,10-epoxides in newborn mice: Exceptional activity of (+)-7 β ,8 α -dihydroxy-9 α ,10 α -epoxy-7,8,9,10-tetrahydrobenzo[a]pyrene. *Proc. Natl. Acad. Sci. U.S.A.* 75, 5358–5361.
12. Cheng, S. C., Hilton, B. D., Roman, J. M., and Dipple, A. (1989) DNA adducts from carcinogenic and noncarcinogenic enantiomers of benzo[a]pyrene dihydrodiol epoxide. *Chem. Res. Toxicol.* 2, 334–340.
13. Weinstein, I. B., Jeffrey, A. M., Jennette, K. W., Blobstein, S. H., Harvey, R. G., Harris, C., Autrup, H., Kasai, H., and Nakanishi, K. (1976) Benzo[a]pyrene diol epoxides as intermediates in nucleic acid binding *in vitro* and *in vivo*. *Science* 193, 592–595.
14. Koreeda, M., Moore, P. D., Wislocki, P. G., Levin, W., Yagi, H., and Jerina, D. M. (1978) Binding of benzo[a]pyrene 7,8-diol-9,10-epoxides to DNA, RNA, and protein of mouse skin occurs with high stereoselectivity. *Science* 199, 778–781.
15. Xu, P., Oum, L., Beese, L. S., Geacintov, N. E., and Broyde, S. (2007) Following an environmental carcinogen N²-dG adduct through replication: Elucidating blockage and bypass in a high-fidelity DNA polymerase. *Nucleic Acids Res.* 35, 4275–4288.
16. Kiefer, J. R., Mao, C., Hansen, C. J., Basehore, S. L., Hogrefe, H. H., Braman, J. C., and Beese, L. S. (1997) Crystal structure of a thermostable *Bacillus* DNA polymerase I large fragment at 2.1 Å resolution. *Structure* 5, 95–108.
17. Boudsocq, F., Iwai, S., Hanaoka, F., and Woodgate, R. (2001) *Sulfolobus solfataricus* P2 DNA polymerase IV (Dpo4): An archaeal DinB-like DNA polymerase with lesion-bypass properties akin to eukaryotic pol η . *Nucleic Acids Res.* 29, 4607–4616.
18. Ohashi, E., Ogi, T., Kusumoto, R., Iwai, S., Masutani, C., Hanaoka, F., and Ohmori, H. (2000) Error-prone bypass of certain DNA lesions by the human DNA polymerase kappa. *Genes Dev.* 14, 1589–1594.
19. Kiefer, J. R., Mao, C., Braman, J. C., and Beese, L. S. (1998) Visualizing DNA replication in a catalytically active *Bacillus* DNA polymerase crystal. *Nature* 391, 304–307.
20. Johnson, S. J., Taylor, J. S., and Beese, L. S. (2003) Processive DNA synthesis observed in a polymerase crystal suggests a mechanism for the prevention of frameshift mutations. *Proc. Natl. Acad. Sci. U.S.A.* 100, 3895–3900.
21. Hsu, G. W., Kiefer, J. R., Burnouf, D., Becherel, O. J., Fuchs, R. P., and Beese, L. S. (2004) Observing translesion synthesis of an aromatic amine DNA adduct by a high-fidelity DNA polymerase. *J. Biol. Chem.* 279, 50280–50285.
22. Hsu, G. W., Ober, M., Carell, T., and Beese, L. S. (2004) Error-prone replication of oxidatively damaged DNA by a high-fidelity DNA polymerase. *Nature* 431, 217–221.
23. Johnson, S. J., and Beese, L. S. (2004) Structures of mismatch replication errors observed in a DNA polymerase. *Cell* 116, 803–816.
24. Hsu, G. W., Huang, X., Luneva, N. P., Geacintov, N. E., and Beese, L. S. (2005) Structure of a high fidelity DNA polymerase bound to a benzo[a]pyrene adduct that blocks replication. *J. Biol. Chem.* 280, 3764–3770.
25. Ling, H., Boudsocq, F., Woodgate, R., and Yang, W. (2001) Crystal structure of a Y-family DNA polymerase in action: A mechanism for error-prone and lesion-bypass replication. *Cell* 107, 91–102.
26. Ling, H., Boudsocq, F., Plosky, B. S., Woodgate, R., and Yang, W. (2003) Replication of a cis-syn thymine dimer at atomic resolution. *Nature* 424, 1083–1087.
27. Ling, H., Boudsocq, F., Woodgate, R., and Yang, W. (2004) Snapshots of replication through an abasic lesion: Structural basis for base substitutions and frameshifts. *Mol. Cell* 13, 751–762.
28. Ling, H., Sayer, J. M., Plosky, B. S., Yagi, H., Boudsocq, F., Woodgate, R., Jerina, D. M., and Yang, W. (2004) Crystal structure of a benzo[a]pyrene diol epoxide adduct in a ternary complex with a DNA polymerase. *Proc. Natl. Acad. Sci. U.S.A.* 101, 2265–2269.
29. Trincao, J., Johnson, R. E., Wolffe, W. T., Escalante, C. R., Prakash, S., Prakash, L., and Aggarwal, A. K. (2004) Dpo4 is hindered in extending a G-T mismatch by a reverse wobble. *Nat. Struct. Mol. Biol.* 11, 457–462.
30. Boudsocq, F., Kokoska, R. J., Plosky, B. S., Vaisman, A., Ling, H., Kunkel, T. A., Yang, W., and Woodgate, R. (2004) Investigating the role of the little finger domain of Y-family DNA polymerases in low fidelity synthesis and translesion replication. *J. Biol. Chem.* 279, 32932–32940.
31. Vaisman, A., Ling, H., Woodgate, R., and Yang, W. (2005) Fidelity of Dpo4: Effect of metal ions, nucleotide selection and pyrophosphorolysis. *EMBO J.* 24, 2957–2967.
32. Zang, H., Goodenough, A. K., Choi, J. Y., Irimia, A., Loukachevitch, L. V., Kozekov, I. D., Angel, K. C., Rizzo, C. J., Egli, M., and Guengerich, F. P. (2005) DNA adduct bypass polymerization by *Sulfolobus solfataricus* DNA polymerase Dpo4: Analysis and crystal structures of multiple base pair substitution and frameshift products with the adduct 1,N²-ethenoguanine. *J. Biol. Chem.* 280, 29750–29764.
33. Zang, H., Irimia, A., Choi, J. Y., Angel, K. C., Loukachevitch, L. V., Egli, M., and Guengerich, F. P. (2006) Efficient and high fidelity incorporation of dCTP opposite 7,8-dihydro-8-oxodeoxyguanosine by *Sulfolobus solfataricus* DNA polymerase Dpo4. *J. Biol. Chem.* 281, 2358–2372.
34. Rechkoblit, O., Malinina, L., Cheng, Y., Kuryavyi, V., Broyde, S., Geacintov, N. E., and Patel, D. J. (2006) Stepwise Translocation of Dpo4 Polymerase during Error-Free Bypass of an oxoG Lesion. *PLoS Biol.* 4, 1–18.
35. Eoff, R. L., Irimia, A., Angel, K. C., Egli, M., and Guengerich, F. P. (2007) Hydrogen bonding of 7,8-dihydro-8-oxodeoxyguanosine with a charged residue in the little finger domain determines miscoding events in *Sulfolobus solfataricus* DNA polymerase Dpo4. *J. Biol. Chem.* 282, 19831–19843.
36. Wang, Y., Musser, S. K., Saleh, S., Marnett, L. J., Egli, M., and Stone, M. P. (2008) Insertion of dNTPs opposite the 1,N²-propanodeoxyguanosine adduct by *Sulfolobus solfataricus* P2 DNA polymerase IV. *Biochemistry* 47, 7322–7334.
37. Irimia, A., Eoff, R. L., Pallan, P. S., Guengerich, F. P., and Egli, M. (2007) Structure and activity of Y-class DNA polymerase DPO4 from *Sulfolobus solfataricus* with templates containing the hydrophobic thymine analog 2,4-difluorotoluene. *J. Biol. Chem.* 282, 36421–36433.
38. Zhang, H., Eoff, R. L., Kozekov, I. D., Rizzo, C. J., Egli, M., and Guengerich, F. P. (2009) Versatility of Y-family *Sulfolobus solfataricus* DNA polymerase Dpo4 in translesion synthesis past bulky N²-alkylguanine adducts. *J. Biol. Chem.* 284, 3563–3576.
39. Bauer, J., Xing, G., Yagi, H., Sayer, J. M., Jerina, D. M., and Ling, H. (2007) A structural gap in Dpo4 supports mutagenic bypass of a major benzo[a]pyrene dG adduct in DNA through template misalignment. *Proc. Natl. Acad. Sci. U.S.A.* 104, 14905–14910.
40. Wong, I., Patel, S. S., and Johnson, K. A. (1991) An induced-fit kinetic mechanism for DNA replication fidelity: Direct measurement by single-turnover kinetics. *Biochemistry* 30, 526–537.
41. Koshland, D. E. (1958) Application of a Theory of Enzyme Specificity to Protein Synthesis. *Proc. Natl. Acad. Sci. U.S.A.* 44, 98–104.
42. Doublié, S., and Ellenberger, T. (1998) The mechanism of action of T7 DNA polymerase. *Curr. Opin. Struct. Biol.* 8, 704–712.
43. Morales, J., and Kool, E. T. (1999) Minor groove interactions between polymerase and DNA: More essential to replication than Watson-Crick hydrogen bonds. *J. Am. Chem. Soc.* 121, 2323–2324.
44. Broyde, S., Wang, L., Rechkoblit, O., Geacintov, N. E., and Patel, D. J. (2008) Lesion processing: High-fidelity versus lesion-bypass DNA polymerases. *Trends Biochem. Sci.* 33, 209–219.

45. Oum, L. (2007) Base-sequence and temperature effects on in vitro studies of translesion synthesis past anti-BPDE- N^2 -dG adducts, New York University, New York.
46. Xu, P., Oum, L., Geacintov, N. E., and Broyde, S. (2008) Nucleotide selectivity opposite a benzo[*a*]pyrene-derived N^2 -dG adduct in a Y-family DNA polymerase: A 5'-slippage mechanism. *Biochemistry* 47, 2701–2709.
47. Perlow-Poehnelt, R. A., Likhterov, I., Scicchitano, D. A., Geacintov, N. E., and Broyde, S. (2004) The spacious active site of a Y-family DNA polymerase facilitates promiscuous nucleotide incorporation opposite a bulky carcinogen-DNA adduct: Elucidating the structure-function relationship through experimental and computational approaches. *J. Biol. Chem.* 279, 36951–36961.
48. Zhuang, P., Kolbanovskiy, A., Amin, S., and Geacintov, N. E. (2001) Base sequence dependence of in vitro translesional DNA replication past a bulky lesion catalyzed by the exo- Klenow fragment of Pol I. *Biochemistry* 40, 6660–6669.
49. Shibutani, S. (1993) Quantitation of base substitutions and deletions induced by chemical mutagens during DNA synthesis in vitro. *Chem. Res. Toxicol.* 6, 625–629.
50. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000) The Protein Data Bank. *Nucleic Acids Res.* 28, 235–242.
51. Wang, L., and Broyde, S. (2006) 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). *Nucleic Acids Res.* 34, 785–795.
52. Case, D. A., Pearlman, D. A., Caldwell, J. W., Cheatham, T. E. I., Wang, J., Ross, W. S., Simmerling, C. L., Darden, T. A., Merz, K. M., Stanton, R. V., Cheng, A. L., Vincent, J. J., Crowley, M., Tsui, V., Gohlke, H., Radmer, R. J., Duan, Y., Pitera, J., Massova, I., Seibel, G. L., Singh, U. C., Weiner, P. K., and Kollman, P. A. (2002) *AMBER*, version 7, University of California, San Francisco.
53. Saenger, W. (1988) *Principles of Nucleic Acid Structure*, pp 556, Springer, Berlin.
54. Xie, X. M., Geacintov, N. E., and Broyde, S. (1999) Stereochemical origin of opposite orientations in DNA adducts derived from enantiomeric *anti*-benzo[*a*]pyrene diol epoxides with different tumorigenic potentials. *Biochemistry* 38, 2956–2968.
55. Perlow, R. A., and Broyde, S. (2002) Toward understanding the mutagenicity of an environmental carcinogen: Structural insights into nucleotide incorporation preferences. *J. Mol. Biol.* 322, 291–309.
56. Delano, W. L. (2002) *The PyMOL Molecular Graphics System*, DeLano Scientific, San Carlos, CA.
57. Perlow-Poehnelt, R. A., Likhterov, I., Wang, L., Scicchitano, D. A., Geacintov, N. E., and Broyde, S. (2007) Increased flexibility enhances misincorporation: Temperature effects on nucleotide incorporation opposite a bulky carcinogen-DNA adduct by a Y-family DNA polymerase. *J. Biol. Chem.* 282, 1397–1408.
58. Kunkel, T. A. (1990) Misalignment-mediated DNA synthesis errors. *Biochemistry* 29, 8003–8011.
59. Ripley, L. S. (1990) Frameshift mutation: Determinants of specificity. *Annu. Rev. Genet.* 24, 189–213.
60. Batra, V. K., Beard, W. A., Shock, D. D., Krahn, J. M., Pedersen, L. C., and Wilson, S. H. (2006) Magnesium-induced assembly of a complete DNA polymerase catalytic complex. *Structure* 14, 757–766.
61. Wang, L., Yu, X., Hu, P., Broyde, S., and Zhang, Y. (2007) A water-mediated and substrate-assisted catalytic mechanism for *Sulfolobus solfataricus* DNA polymerase IV. *J. Am. Chem. Soc.* 129, 4731–4737.
62. Kunkel, T. A. (1992) DNA replication fidelity. *J. Biol. Chem.* 267, 18251–18254.
63. Seo, K. Y., Nagalingam, A., Miri, S., Yin, J., Chandani, S., Kolbanovskiy, A., Shastry, A., and Loechler, E. L. (2006) Mirror image stereoisomers of the major benzo[*a*]pyrene N^2 -dG adduct are bypassed by different lesion-bypass DNA polymerases in *E. coli*. *DNA Repair* 5, 515–522.