Y-family DNA polymerases lack some of the mechanisms that replicative DNA polymerases employ to ensure fidelity, resulting in higher error rates during replication of undamaged DNA templates and the ability to bypass certain aberrant bases, such as those produced by exposure to carcinogens, including benzo[a]pyrene (BP). A tumorigenic metabolite of BP, (+)-anti-benzo[a]pyrene diol epoxide, attacks DNA to form the major 10S (+)-trans-anti-[BP]-N²-dG adduct, which has been shown to be mutagenic in a number of prokaryotic and eukaryotic systems. The 10S (+)-trans-anti-[BP]-N²-dG adduct can cause all three base substitution mutations, and the SOS response in Escherichia coli increases bypass of bulky adducts, suggesting that Y-family DNA polymerases are involved in the bypass of such lesions. Dpo4 belongs to the DinB branch of the Y-family, which also includes E. coli pol IV and eukaryotic pol κ. We carried out primer extension assays in conjunction with molecular modeling and molecular dynamics studies in order to elucidate the structure-function relationship involved in nucleotide incorporation opposite the bulky 10S (+)-trans-anti-[BP]-N²-dG adduct by Dpo4. Dpo4 is able to bypass the 10S (+)-trans-anti-[BP]-N²-dG adduct, albeit to a lesser extent than unmodified guanine, and the $V_{max}$ values for insertion of all four nucleotides opposite the adduct by Dpo4 are similar. Computational studies suggest that 10S (+)-trans-anti-[BP]-N²-dG can be accommodated in the active site of Dpo4 in either the anti or syn conformation due to the limited protein-DNA contacts and the open nature of both the minor and major groove sides of the nascent base pair, which can contribute to the promiscuous nucleotide incorporation opposite this lesion.

Members of the Y-family of DNA polymerases have been implicated in the bypass of DNA damage that poses blocks to high fidelity replicative A-family polymerases (1–9). The functional interplay between replicative and Y-family DNA polymerases, as well as the mechanism of damage bypass by both types of enzymes, is of considerable interest in understanding the origin of in vivo mutagenesis. While replicative DNA polymerases employ a number of mechanisms to ensure fidelity, including induced-fit geometric selection of the nascent base pair and minor groove scanning of the primer-template duplex region (reviewed in Refs. 10–16), Y-family DNA polymerases appear to lack such mechanisms, resulting in higher error rates during replication of undamaged DNA templates and the ability to bypass certain aberrant bases (1, 2, 5–8, 17–25). It is thought that the more permissive nature of Y-family polymerases has evolved, at least in part, to rescue replication machinery that is blocked by altered bases, such as those produced by exposure to bulky mutagens and carcinogens (26–29). Erroneous bypass of carcinogen-damaged bases can contribute to the carcinogenic process if a mutation occurs in a gene critical to cell cycle control (30, 31).

Chemical carcinogens are encountered throughout the environment, and the manner by which their tumorigenic effects are exerted is of much interest in understanding the origins of environmental mutagenesis and carcinogenesis. Polycyclic aromatic hydrocarbons (PAHs) are an omnipresent class of environmental carcinogens that are byproducts of the incomplete combustion of organic matter, and thus, are produced during volcanic eruptions, and are present in automobile, jet, and factory emissions, barbequed meats and fish, and cigarette smoke (32–36). Benzo[a]pyrene (BP) is a commonly studied PAH that has been implicated as a causative agent in lung cancer (37, 38); it is metabolically activated to a number of reactive intermediates in mammals (39), including the tumorigenic (+)-anti-benzo[a]pyrene diol epoxide (BPDE) (40). The electrophilic (+)-anti-BPDE attacks DNA predominantly at the exocyclic amino group of guanine to form the 10S (+)-trans-anti-[BP]-N²-dG ([±]ta[BP]G) adduct (41, 42)(shown in Fig. 1a). High resolution NMR solution studies have shown that [±]ta[BP]G adopts a normal anti conformation in duplex DNA opposite a dC partner, and the carcinogen moiety resides in the minor groove with its pyrenyl rings pointing toward the 5′-side of the damaged guanine (43, 44). A similar conformation, with the damaged guanine anti and hydrogen-bonded to its partner, is adopted when [±]ta[BP]G is placed opposite dA at a single-strand/double junction (45). In contrast, with no partner oppo-

The abbreviations used are: BP, benzo[a]pyrene; BPDE, (+)-anti-benzo[a]pyrene diol epoxide; DTT, dithiothreitol; SSA, solvent-exposed surface area; pol, polymerase; SD, steepest descent; r.m.s.d., root mean square deviations.
site the lesion at a single-strand/double-strand junction, the (+)-ta[BP]G adduct adopts a syn conformation with the carcinogen moiety stacked on top of the duplex region (46).

The (+)-ta[BP]G adduct is mutagenic in a number of prokaryotic (7, 47–61) and eukaryotic (4, 59, 62–64) systems. Specifically, this adduct can cause all three base substitution mutations, with their relative frequencies dependent on sequence context, as well as the host cell replication system by which the adduct is bypassed (47–53, 55, 56, 59–62, 65). In *Escherichia coli*, the replicative DNA polymerase III has been shown to participate in the predominant point mutation pathway, namely G→T transversions, at (+)-ta[BP]G adducts in a 5′-GG*A-3′ sequence context, where G* represents the adduct (60). However, induction of the SOS response in *E. coli* increases bypass of bulky adducts, including (+)-ta[BP]G, with a significant increase in both error-free and error-prone events (49, 51, 60, 66); this suggests that Y-family DNA polymerases, including pol IV and pol V, evolved to overcome replication hurdles in order to increase cellular survival (6, 60). In agreement, a recent study suggests two mutagenic pathways for G→T mutations in *E. coli* dictated by sequence context, one of which is dependent on DNA polymerase V, a Y-family DNA polymerase (67). Indeed, bulky adducts, such as (+)-ta[BP]G, predominantly block replicative DNA polymerases (47, 56–58, 68–70), but members of the Y-family, including pol κ (71), pol ζ (64), and pol η (63), bypass these adducts more readily than their replicative counterparts, albeit with reduced efficiency when compared with that in undamaged DNA (4, 7, 20, 63, 71). In addition, pol κ has been implicated in facilitating survival of mammalian cells exposed to metabolically activated benzo(a)pyrene (26), further supporting a physiological role for Y-family polymerases in lesion bypass.

Previous studies have suggested that replicative DNA polymerases are predominantly blocked by (+)-ta[BP]G because the damaged base must adopt the less common syn conformation in order to be accommodated within the sterically constrained active site (72–74). However, little is known about the mechanism by which the low fidelity Y-family DNA polymerases accommodate bulky lesions, such as (+)-ta[BP]G, compared with high fidelity DNA polymerases. Dpo4 is a Y-family DNA polymerase from the crenarchaeon *Sulfolobus solfataricus* P2 that offers a unique opportunity to investigate the structure-function relationship involved in adduct bypass. Not only is the enzyme readily available for functional studies (1, 22), but high resolution crystal structures have also been solved of the ternary complex of this enzyme, including the primer-template and incoming nucleotide (2, 75–77).

The 40-kDa Dpo4 enzyme belongs to the DinB branch of the Y-family, which also includes *E. coli* pol IV and eukaryotic pol κ (78), and was isolated from *S. solfataricus* P2, a thermostable Archaea (1). Despite little sequence similarity, Dpo4 and other Y-family DNA polymerases display structural similarities to their replicative cousins, adopting shapes reminiscent of a right hand with palm, finger, and thumb domains (2, 22, 79–83). However, unlike the α-helical nature of replicative polymerase finger domains, a β-strand and adjacent loop in the finger domain of Dpo4 contact the face of the nascent base pair (2). In addition, Dpo4 possesses an extra C-terminal domain, dubbed the little finger, that is thought to be involved in increasing processivity (1, 2, 22). Dpo4 makes limited contacts with the
primer-template, and notably, the minor groove side of the primer-template is solvent-exposed (2) and lacks the minor groove-scanning tract present in replicative DNA polymerases (14, 15, 79, 80, 82–88). Limited protein-DNA interactions and relaxed fidelity mechanisms may help explain why Dpo4 is both distributive, incorporating 1–2 nucleotides per primer-template binding event, and relatively error-prone, with a mutation frequency of $8 \times 10^{-3}$ to $4 \times 10^{-3}$ when replicating undamaged DNA (1).

In the present work, we have carried out primer extension assays in conjunction with molecular modeling and molecular dynamics studies in order to elucidate the structure-function relationship involved in the incorporation of different nucleotides opposite the bulky (+)-ta[BP]G adduct catalyzed by Dpo4. Our experimental results reveal that this enzyme is able to incorporate all four bases opposite the damaged guanine at similar rates, and the simulations provide molecular explanations for the accommodation of each incoming nucleotide opposite this bulky carcinogen-DNA adduct.

**MATERIALS AND METHODS**

**Damaged Oligodeoxynucleotides and Primer Extension Template Construction** —The oligodeoxynucleotides were synthesized using automated methods based on phosphoramidite chemistry on a Biosearch Cyclone automated DNA synthesizer (Milligen-Biosearch Corp., San Rafael, CA) and were purified using high performance liquid chromatography (HPLC). The 11-mer oligonucleotide, 5′-TTA TAGG* CAC AC-3′ with the single (+)-ta[BP]G adduct at G*, was generated by a direct synthetic method and purified and characterized as described in detail elsewhere (71, 89–91). This oligodeoxynucleotide sequence was designed to prevent slippage-mediated bypass of the adduct (45). The 42-mers containing a single (+)-ta[BP]G were constructed by ligation methods (45, 65, 68); the modified 11-mers and flanking 12- and 19-mers were annealed with a complementary 38-mer template strand, and then the three oligonucleotides were ligated to one another using T4 DNA ligase (New England Biolabs, Inc., Beverly, MA) as described previously (47, 68, 71). The intact 42-mers were purified using 20% PAGE with 8 M urea, visualized using ethidium bromide and then excised from the gel. The sequence of the 42-mer template thus generated was: 5′-CAG TTA ATG TTA TAGG* CAC AC/G TCA TCT GCC CAG ATC GGC-3′, where G* represents (+)-ta[BP]G or unmodified guanine. The oligonucleotides were eluted from the gel slice using the PAGE purification kit (Bio-Rad), the ethidium bromide was removed by butanol extraction, and the samples were desalted using BioRad Spin6 columns. The unmodified control 42-mer template was constructed in a similar manner. Each template strand was end-labeled using T4 phosphoxygenase kinase (PNK) (New England Biolabs, Inc., Beverly, MA), and their purity was checked by running on a 20% polyacrylamide gel containing 8 M urea. Adduct presence was evident because the damaged strand migrated more slowly than the undamaged strand, and the presence of the (+)-ta[BP]G was verified by checking the fluorescence and UV absorption of the sample. Complementary primers of 18 and 21 bases long were purchased from Sigma Genosys (The Woodlands, TX), and were PAGE-purified as described above.

**Primer Extension Assays** —The 21-mer primers, which extend up to 1 base before the adduct, were 5′-end labeled using 1 unit of T4PNK (New England BioLabs, Beverly, MA) and 0.01 mM cGMP (γ,γ′-PITP) (3,000 Ci/mmol) (PerkinElmer Life Sciences, Wellesley, MA) in 70 mM Tris-HCl (pH 7.6), 10 mM MgCl2, and 5 mM DTT. The 10-μM reactions were preincubated for 1 h at 37 °C, the volume was brought up to 20 μl with water, and the solutions were passed through Micro Bio-Spin P6 Chromatography Columns to remove unincorporated radioactivity. The labeled primer and unlabeled 42-mer DNA templates were annealed in a 1:1.25 ratio by heating to 80 °C and cooling slowly to 4 °C. In the standing-start assay, the primer was extended to an 8-mer and the reaction stopped with 10 μl of stop solution (80% formamide, 5 mM Tris-boric acid, 1 mM EDTA, 0.1% xylene cyanol, 0.1% bromphenol blue). The reaction-stop mixtures were heated to 90 °C for 10 min prior to analysis on a 20% polyacrylamide gel containing 8 M urea. Gels were dried and exposed to a phosphorimager plate prior to data analysis with ImageQuant software (Amersham Biosciences).

**Running-start primer extension assays were also carried out (results not shown), using similar methods, but utilizing a 16-mer primer that extended up to four bases before the adduct. The labeled primer and unlabeled 42-mer DNA templates were annealed in a 1:1.25 ratio by heating to 80 °C and cooling slowly to 4 °C. Running-start primer extension assays with Dpo4 were carried out in a 40-μl final volume, containing 40 mM Tris-HCl (pH 8.0), 5 mM MgCl2, 10 mM DTT, 250 μM dATP, 250 μM dGTP, 250 μM dCTP, 250 μM dTTP, 5 μM [32P]dNTP, and 10 μM Dpo4. Running-start primer extension assays using T7 DNA polymerase were also carried out in 40 μl volume, containing 40 mM Tris-HCl (pH 7.5), 20 mM MgCl2, 50 mM NaCl, 2000 μM dNTPs (Promega Corp., Inc., Madison, WI), 2.5 mM 3′-PPrimertemplate and 10 nM T7 DNA polymerase. Reactions were incubated at 37 °C, and 10-μl aliquots were removed and stopped with 10 μl of stop solution (80% formamide, 5 mM Tris-boric acid, 1 mM EDTA, 0.1% xylene cyanol, 0.1% bromphenol blue) after 15, 30, and 60 min. Control reactions were carried out in an analogous fashion for each polymerase with undamaged DNA template for 60 min, except that 125 μM dNTPs was used. The reaction-stop mixtures were heated to 90 °C for 10 min and electrophoresed on a 20% polyacrylamide gel containing 8 M urea. Gels were dried and exposed to a phosphorimager plate prior to data analysis with ImageQuant software (Amersham Biosciences).**

**Molecular Modeling and Molecular Dynamics Simulations of Dpo4 Ternary Complexes Containing the (+)-trans-anti-BP/NC-DG Adduct** —Parameters for the (+)-ta[BP]G, dNTP, and Mg2+ residues developed previously were employed in the present work (72–74). The 1.7-Å resolution I crystal structure of the Dpo4 ternary complex (2) was used as the starting structure for molecular models (PDB ID: 1JX4) and the coordinates were obtained from the Protein Data Bank (92). The Builder module of InsightII (Accelrys, Inc., San Diego, CA) was used for molecular modeling, including remodeling the sequence in the crystal structure to correspond with that used in the primer extension assays. The catalytic Mg2+ ion in the active site was unresolved in the crystal structure, likely because of the lack of hydroxyl group on the 3′-terminus of the primer, and it was modeled in using its predicted coordination site (2). The (+)-ta[BP]G residue was modeled in both anti and syn conformations and each anti dNTP was modeled opposite the adduct; an unmodified control simulation, with an undamaged anti guanine template and dCTP partner was also created in parallel. In addition, systems were constructed where the anti (+)-ta[BP]G was modeled opposite syn dATP and syn dGTP, since a Dpo4 structure with a damaged template opposite a syn incoming purine has been previously observed in a crystal structure (75). Therefore, a total of eleven simulations were carried out. New England Biolabs, Inc., Beverly, MA), and their purity was checked by running on a 20% polyacrylamide gel containing 8 M urea. Adduct presence was evident because the damaged strand migrated more slowly than the undamaged strand, and the presence of the (+)-ta[BP]G was verified by checking the fluorescence and UV absorption of the sample. Complementary primers of 18 and 21 bases long were purchased from Sigma Genosys (The Woodlands, TX), and were PAGE-purified as described above.

**Primer Extension Assays** —The 21-mer primers, which extend up to 1 base before the adduct, were 5′-end labeled using 1 unit of T4PNK (New England BioLabs, Beverly, MA) and 0.01 mM cGMP (γ,γ′-PITP) (3,000 Ci/mmol) (PerkinElmer Life Sciences, Wellesley, MA) in 70 mM Tris-HCl (pH 7.6), 10 mM MgCl2, and 5 mM DTT. The 10-μM reactions were preincubated for 1 h at 37 °C, the volume was brought up to 20 μl with water, and the solutions were passed through Micro Bio-Spin P6 Chromatography Columns to remove unincorporated radioactivity. The labeled primer and unlabeled 42-mer DNA templates were annealed in a 1:1.25 ratio by heating to 80 °C and cooling slowly to 4 °C. In the standing-start assay, the primer was extended to an 8-mer and the reaction stopped with 10 μl of stop solution (80% formamide, 5 mM Tris-boric acid, 1 mM EDTA, 0.1% xylene cyanol, 0.1% bromphenol blue). The reaction-stop mixtures were heated to 90 °C for 10 min prior to analysis on a 20% polyacrylamide gel containing 8 M urea. Gels were dried and exposed to a phosphorimager plate prior to data analysis with ImageQuant software (Amersham Biosciences).
To begin equilibration, 30 ps of molecular dynamics (MD) was carried out at 10 K with 25 kcal/mol-Å restraints on the solute. The systems were then heated from 10 to 310 K over 25 ps, again holding the solute fixed with 25 kcal/mol-Å restraints, and the system was held at 310 K for 60 ps to mix the waters and ions. The remainder of the simulations were carried out at 10 K with 25 kcal/mol-Å restraints; and 45 ps with 0.1 kcal/mol-Å restraints. Following equilibration, unrestrained MD was carried out for 2.5 ns.

The coordinates of the simulated systems were saved every 1 ps for detailed analysis of the time-dependent structural features, and the structural characteristics analyzed include: root mean square deviations (r.m.s.d.) of the system, the active site region of the system and individual residues, interatomic distances, hydrogen bonding occupancies, hydrogen bonding distances, and angles and torsion angles. The cut-off for hydrogen bonding heavy-atom to heavy-atom distance was 3.4 Å and that for the hydrogen bonding angle was 140°. Plots of r.m.s.d. demonstrating the stability of the simulations are shown in Supplementary Materials Fig. S1. Solvent-accessible surface area was calculated using the Connolly algorithm (102) with a probe diameter of 1.4 Å in InsightII (Accelrys, Inc.).

RESULTS AND DISCUSSION

**Primer Extension Assays Reveal Similar V_{max} Values for each Nucleotide Opposite (±)-trans-anti-[BP]-N2-dG by Dpo4—** Running-start primer extension experiments (data not shown) indicate that primer extension catalyzed by Dpo4 is slowed one base prior to (±)-ta[BP]G on the template strand, but a nucleotide can be incorporated opposite the adduct, and the damage can be bypassed to a significant extent. In contrast, a replicative DNA polymerase, such as T7, produces little full-length extension products under the same experimental conditions.

Standing-start primer extension assays were carried out using a primer that extended up to the modified guanine, with the adduct (or unmodified guanine in the control reactions) as the next-to-be-replicated base, in order to determine the V_{max} for incorporation of each nucleotide opposite (±)-ta[BP]G by Dpo4. The control standing-start primer extension experiments, as well as previously published work (1), showed that dC is overwhelmingly inserted most often opposite unmodified guanine. Examples of standing-start primer extension experiments at two temperatures (37 and 55 °C) are shown in Fig. 2a. In these
examples the dNTP concentration was 2 mM in each case. All four 2'-deoxyribonucleotide triphosphates are effectively incorporated opposite the adduct, especially at 55 °C. However, the scenario modeled in our structural studies involves full occupancy of the enzyme active site by the incoming dNTP, which represents conditions of substrate saturation. This is relevant to $V_{\text{max}}$, the rate of incorporation in a ternary complex that has a dNTP at its active site. Accordingly, corresponding $V_{\text{max}}$ values for the insertion of each nucleotide opposite the damaged guanine were obtained experimentally in order to elucidate the biochemical behavior of the enzyme under saturating conditions. We determined $V_{\text{max}}$ by evaluating the rates of incorporation of the different dNTPs at sufficiently high nucleotide concentrations such that further increases caused no further changes in the rates of incorporation, thus identifying $V_{\text{max}}$ in each case. Interestingly, Dpo4 is highly promiscuous in standing-start primer extension assays opposite (+)-ta[BP]G.

Fig. 2b shows the $V_{\text{max}}$ values and standard deviations for the incorporation of each nucleotide opposite the damaged base. All four nucleotides are incorporated opposite (+)-ta[BP]G with comparable $V_{\text{max}}$ values. This indicates that, once the active site of Dpo4 is occupied by one of the incoming nucleotides opposite the damaged guanine, they are all incorporated to a similar extent. The $V_{\text{max}}$ values determined in this study for the damaged DNA were approximately one order of magnitude lower than those for unmodified DNA (1, 2).

The Pyrenyl Moiety of (+)-trans-anti [BP]:N2-dG Can Be Accommodated on the Minor Groove Side of the Nascent Base Pair, with the Adduct in the anti Conformation within the Dpo4 Active Site—Unlike high fidelity DNA polymerases (15, 79, 80, 87, 88, 104), Dpo4 does not forge critical interactions with the Active Site Accommodated on the Minor Groove Side of the Nascent Base pair, as shown by the distance between these domains (Fig. S4 in Supplementary Material). For example, the Lys-78 Cα (palm) to Lys-275 Cα (little finger) distance was 12.3 ± 0.7 Å and 16.5 ± 0.8 Å in the unmodified and anti (+)-ta[BP]G:dCTP simulations, respectively. The BP moiety in the minor groove is positioned between protein residues and the phosphodiester backbone of the template DNA strand. This causes the template phosphodiester backbone to shift toward the major groove compared with the unmodified system when the two systems are aligned via the palm, fingers and thumb domains, as shown in Fig. 4a. If the BP moiety were present in conjunction with the DNA in the conformation adopted by the unmodified system, it would clash with the nascent base pair on the template and incoming bases in each simulation, and the (+)-ta[BP]G template base, the incoming dGTP, and dATP, anti (+)-ta[BP]G: dGTP and anti (+)-ta[BP]G:dCTP structures. In addition, two simulations were carried out with the damaged base in the anti conformation and a purine incoming nucleotide in the syn conformation: anti (+)-ta[BP]G:syn dATP and anti (+)-ta[BP]G: syn dGTP. These structures were created because an anti damaged base was observed opposite a syn incoming purine nucleotide in a previously solved crystal structure (75). The simulated structures, after potential energy minimization, equilibration, and 2.5 ns of unrestrained molecular dynamics, are shown in Figs. S1 and S2 in the Supplementary Material show the time-dependent analyses of the r.m.s.d. and torsion angles governing the orientation of the damaged guanine and carcinogen moiety, respectively, in the simulations during the 2.5 ns of unrestrained molecular dynamics.

The anti (+)-ta[BP]G forms stable hydrogen bonds with anti dCTP, dATP, and dGTP during their respective simulations, as exemplified in Fig. S2b. In addition, the nascent base pairs in these simulations remain coplanar during the simulation, as exemplified in Fig. 3a. In contrast, hydrogen bonds between anti (+)-ta[BP]G and partner anti dGTP do not form because the identical hydrogen bonding edges of each base repel each other; this repulsion forces the dGTP base moieties out of the plane of the damaged guanine template, as shown in Fig. 3a. Table I lists the solvent-exposed surface area (SSA) of the template and incoming bases in each simulation, and the (+)-ta[BP]G residue in the anti (+)-ta[BP]G:dGTP simulation has the highest SSA of the anti (+)-ta[BP]G systems. This indicates that the anti (+)-ta[BP]G:dGTP pair is the most solvent-exposed nascent base pair. This probably results from the unfavorable interactions between the hydrogen bonding faces of the two guanine residues in the anti conformation. In lieu of interaction with the anti (+)-ta[BP]G template base, the incoming dGTP participates in hydrogen bonding interactions with the dC9 residue on the 3′ side of the modified guanine (see Fig. 1c), forming a pseudo-slipped structure, as shown in Fig. S3b. The dC9 base interacts both with the skewed dGTP base and its normal partner G90. Misaligned bases within the active site of Dpo4 have been observed previously in x-ray crystal structures (2, 77), and this enzyme is capable of misincorporating bases via a mechanism that is mediated by misaligned incoming nucleotides (18, 103). Therefore, this type of slipped structure within Dpo4 may be adequate for successful nucleotide incorporation. Such a slipped structure could persist through elongation, resulting in a frameshift mutation. Alternatively, this slipped structure could realign and be extended by the polymerase, resulting in a point mutation.

Although the (+)-ta[BP]G adduct can be accommodated by Dpo4 in the anti conformation, placing the bulky BP moiety in the minor groove results in crowding between the pyrenyl rings and the protein and primer-template DNA. This wedging of the BP moiety into the minor groove clef causes the little finger domain (2) to move away from the fingers and palm domains, as shown by the distance between these domains (Fig. S4 in Supplementary Material). For example, the Lys-78 Cα (palm) to Lys-275 Cα (little finger) distance was 12.3 ± 0.7 Å and 16.5 ± 0.8 Å in the unmodified and anti (+)-ta[BP]G:dCTP simulations, respectively. The BP moiety in the minor groove is positioned between protein residues and the phosphodiester backbone of the template DNA strand. This causes the template phosphodiester backbone to shift toward the major groove compared with the unmodified system when the two systems are aligned via the palm, fingers and thumb domains, as shown in Fig. 4a. If the BP moiety were present in conjunction with the DNA in the conformation adopted by the unmodified system, it would clash with the +1 template base, as well as Lys-78 and Met-76. The shift in position of the template backbone caused by the BP moiety is likely the cause of the opening of the little finger, which is not observed in the simulation with the unmodified DNA.

In addition to widening the minor groove cleft in the Dpo4 protein, the presence of the bulky BP moieties on the minor groove side of the nascent base pair causes rearrangement of several residues near the active site to allow for the accommodation of the carcinogen moiety, as shown in Fig. 4b. Most notably, Lys-78, which contacts the active site region DNA on its minor groove side in the unmodified system, is rotated ~180° in the anti (+)-ta[BP]G:dGTP simulation to allow room for the BP moiety. Due to the rotation and motion of Lys-78, Glu-79 is also significantly moved in the anti (+)-ta[BP]G:dCTP system compared with its position in the unmodified structure. In addition, Met-76 contacts the nascent base pair on its minor groove side in the unmodified system and moves toward the 5′-side of the modified base in order to allow the BP moiety to reside in the minor groove. It is likely that the reorganization caused by the presence of the BP moiety on the minor groove side of the nascent base pair, including the opening of the little finger and rearrangement of active site region residues, results in the poor bypass of (+)-ta[BP] by Dpo4 in

---

2. R. Woodgate, personal communication.
vitro. However, the ability of Dpo4 to accommodate the bulky BP moiety in the minor groove without a major distortion of the protein likely plays an important role in the increased ability of this enzyme to bypass (\(=ta\))-[BP]G in contrast to the minimal bypass by replicative T7 DNA polymerase. It also helps explain the ability of Dpo4 to incorporate cytosine opposite the bulky lesion, albeit to a reduced extent as compared with that opposite unmodified guanine.

Incoming Purine Nucleotide Triphosphates in the syn Conformation Can Be Accommodated Opposite anti (\(+\))-[BP]-N\(^2\)-dG Adduct by Dpo4—Recent work by Ling et al. (75) has shown that Dpo4 can bypass a cyclobutane pyrimidine dimer (CPD) and can stably accommodate a syn ddATP opposite the 5'-thymine base moiety of the lesion. Similarly, our simulations reveal that a syn purine nucleotide can be accommodated opposite the (\(+\))-[BP]G adduct in the anti conforma-
All values are for the structures at 2.5 ns except for anti (+)-ta[BP]G: dCTP, which is at 1.75 ns, representative of the structure predominant throughout the simulation. For this case, values at 2.5 ns are 305.4 Å² and 145.2 Å² for the template and incomer, respectively.

Table I

| Template base   | dNTP   | Solvent-exposed surface area | dNTP residue |
|-----------------|--------|------------------------------|--------------|
|                 |        | Å²                           |              |
| Unmod           | dCTP   | 35.7                         | 73.2         |
| anti (+)-ta[BP]G| dCTP   | 199.8                        | 80.3         |
|                 | dATP   | 184.9                        | 92           |
|                 | dGTP   | 240.7                        | 88.2         |
|                 | dTTP   | 180.8                        | 83.5         |
| syn ta[BP]G     | dATP   | 203.4                        | 94.4         |
|                 | syn dGTP | 254.5                    | 105.3        |
| syn (+)-ta[BP]G | dCTP   | 258.5                        | 82.3         |
|                 | dATP   | 202.8                        | 94.3         |
|                 | dGTP   | 208.2                        | 102.8        |
|                 | dTTP   | 193.7                        | 68.6         |

Y-Family DNA Polymerase Accommodation of a Bulky DNA Adduct

The Bulky (+)-trans-anti-[BP]-N²-dG Adduct Causes Less Disruption of Protein Structure When It Adopts the syn Conformation, with the BP Moiety Residing on the Major Groove Side of the Nascent Base Pair, within the Active Site of Dpo4—Dpo4 has a large open pocket on the major groove side of the primer-template and nascent base pair, reminiscent of that seen in high fidelity DNA polymerases (72, 79, 80, 83, 86), and the bulky BP moiety is easily accommodated within this region when the template (+)-ta[BP]G adopts the syn conformation. Therefore, the (+)-ta[BP]G was also modeled in the syn conformation, with the BP moiety in the major groove, opposite each of the four nucleotides, to create the syn (+)-ta[BP]G:dCTP, syn (+)-ta[BP]G:dATP, syn (+)-ta[BP]G:dGTP and syn (+)-ta[BP]G:dTTP structures. In contrast to the systems in which the adduct adopted the anti conformation, placing the BP moiety in the major groove, as in the syn (+)-ta[BP]G:dNTP systems, does not cause the little finger domain to open up, as seen in Fig. 4c and Fig. 5d. Also, the residues located on the minor groove side of the primer-template of the syn (+)-ta[BP]G structures are in positions very similar to those in the unmodified system, since this region is not disrupted by the presence of the bulky carcinogen moiety. The major groove side of the nascent base pair provides a more spacious cavity in which the BP moiety can reside without nearly as significant a disruption of the Dpo4 structure. Fig. 4c shows a comparison of the cavity size on the minor and major groove sides of the primer-template/nascent base pair in the Dpo4 enzyme.

Although the BP moiety has more room in the major groove, certain residues in the active site region still undergo modest rearrangement as compared with their positions in the unmodified DNA-enzyme complex in order to accommodate the pyrenyl moiety. Primarily, Leu-293 is shifted toward the 3'-side of the template strand in order to avoid steric crowding with the saturated ring of the BP moiety, and Arg-332 is reoriented in the syn (+)-ta[BP]G simulations to hydrogen bond with the phosphodiester backbone of the modified guanine residue. When (+)-ta[BP]G adopts the syn conformation, it hydrogen bonds to both dCTP and dATP within the active site of Dpo4 during their respective simulations, as shown in Fig. S3d. One stable hydrogen bond forms between the nascent base pair in the syn (+)-ta[BP]G:dATP simulation; the dATP incoming base is hydrogen bonded to N7 of the template through its N6 atom for 81% of the simulation. The dCTP residue is hydrogen bonded to the syn (+)-ta[BP]G for 64% of the simulation; during the last ~250 ps, this interaction is disrupted through a motion of the dCTP base moiety toward the major groove side of the nascent base pair. Prior to this shift, the dCTP base is also involved in a hydrogen bond with Lys-78 for 51% of the simulation.
ta[BP]G simulations, as seen in Table I. Regardless of the specific partner dNTP, availability of the syn domain of the modified guanine within the active site of Dpo4 provides another avenue through which nucleotide incorporation can occur without major disruption of the protein structure.

Summary and Conclusions—The Dpo4 enzyme is a good model system for elucidating structure-function relationships because of the availability of both high resolution crystal structures of its ternary complex with DNA for structural studies (2) and purified enzyme for in vitro functional assays (1). While
running-start primer extension assays showed that Dpo4 exhibits a slowed incorporation of dNTPs opposite (+)-ta[BP]G, full bypass of the damaged base is observed as well. Standing-start primer extension assays showed that the enzyme is promiscuous opposite the adduct, inserting all four bases with comparable $V_{\text{max}}$ values (Fig. 2).

Molecular modeling and molecular dynamics simulations of (+)-ta[BP]G opposite each of the four dNTPs reveal that the adduct can adopt either the anti or syn conformation within the Dpo4 active site. This finding highlights structural differences between Dpo4 and the more stringent replicative DNA polymerases, specifically T7 DNA polymerase, where (+)-ta[BP]G must adopt the syn conformation to be accommodated within the active site of the ternary complex without a major perturbation of the protein structure (72–74). In the anti conformation, (+)-ta[BP]G can be accommodated opposite each of the four nucleotides within the active site of Dpo4, including all four dNTPs in the anti conformation and the purine dNTPs in the syn conformation. This could explain the promiscuity of this enzyme opposite the damaged base. However, dGTP opposite anti (+)-ta[BP]G becomes distorted because of the abutting identical hydrogen bonding edges of the incoming dNTP and template base. In fact, both anti and syn dGTP opposite the anti adduct forgo hydrogen bonding with the template base in favor of that with $dC_5$, on the 3' side of the adduct. Despite the lack of hydrogen bonds with the templating base, $dG$ is incorporated opposite the adduct with a $V_{\text{max}}$ comparable to that of the other three bases. It is likely that slipped structures are acceptable substrates for Dpo4 (2, 18, 77, 103).

In addition, (+)-ta[BP]G can be accommodated in the syn conformation opposite all four anti dNTPs. Dpo4 can exploit the solvent-exposed regions on both the minor and major groove sides of the nascent base pair to accommodate the bulky carcinogen moiety of (+)-ta[BP]G, effectively giving the enzyme at least two options for accommodating the damaged base, in either the anti or syn conformation. It is possible that the syn conformation is less accessible, and predominance of the adduct anti conformation within the active site could result in the significant blockage of Dpo4 by (+)-ta[BP]G. When opposite dC at a primer-template junction, (+)-ta[BP]G adopts the anti conformation (45), suggesting that this conformation may be favored when the adduct has a partner base, or at least the non-mutagenic partner.

It is plausible that the more open and permissive nature of the Dpo4 active site facilitates accommodation of aberrant and/or damaged base pairs within the active site region, possibly with less-constrained alignments than in high fidelity replicative DNA polymerases. This effect may work in conjunction with the anti/syn structural opportunities within the active site to produce the nucleotide incorporation promiscuity observed in standing-start primer extension experiments with the Dpo4 enzyme. In fact, (+)-ta[BP]G has been observed in both anti and syn conformations at a primer-template junction in solution (45, 46). Thus, a combination of structural factors in the Dpo4 enzyme, involving both the spacious active site and the open minor groove together, may provide the flexibility to admit the different incoming nucleotides with similar rates.

As pointed out by an insightful reviewer, a number of hypotheses can be considered to account for the comparable incorporation rates of all four nucleotides opposite (+)-ta[BP]G. These include (1) a single orientation of (+)-ta[BP]G, for which all dNTPs have equal affinity; (2) different orientations of (+)-ta[BP]G, each with its preference for a certain dNTP or dNTPs; (3) a “non-instructional” conformation of (+)-ta[BP]G, allowing equivalent accommodation of each dNTP. Our results suggest multiple conformations of (+)-ta[BP]G can facilitate nucleotide incorporation within the spacious active site of Dpo4, and all four dNTPs can be accommodated opposite one or more of these conformations. In addition, we find that purine-incoming nucleotides can also be accommodated in more than one conformation (anti and syn) opposite the damaged guanine. Thus, through a number of combinations of template and incomer conformations, together with the inherent flexibility and roominess of the active site, our results suggest how Dpo4 can incorporate each of the four dNTPs opposite (+)-ta[BP]G at comparable rates in an instructional, template-directed manner.

Y-family DNA polymerases are structurally distinct from their replicative cousins and it has been proposed that they may be involved in bypassing DNA damage in vivo (27, 78, 103, 105). Because of their possible involvement in damage processing, it is of great interest to understand how these polymerases handle DNA adducts, including bulky lesions such as those derived from BPDE. A recent crystal structure has revealed the accommodation of a BPDE-adenine adduct, linked to $N^6$ of da, within Dpo4. Two structures were resolved, including a proposed blocking structure with the BP moiety intercalated between adjacent base pairs, similar to that seen by high resolution NMR in duplex DNA (106), and a proposed bypass structure, with the BP moiety solvent-exposed in the major groove (76). The major groove structure showed that the enzyme can coerce the adduct into adopting a conformation that is disfavored in duplex DNA (106, 107) and on the nucleoside level (108).

It appears that the increased bypass ability and promiscuous
nature of nucleotide incorporation opposite (+ta)[BP]G by DpO4 can be attributed to the open nature of the active site and the ability to accommodate bulky moieties in the minor groove, albeit to a lesser degree than in the major groove. It is possible that this structural theme is also echoed in other Y-family DNA polymerases, including a DpO4 DinB family relative, pol δ, which has been shown to primarily incorporate D¢ opposite (+ta)[BP]G (14). Perhaps pol δ has a pocket on the minor groove side of the nascent base pair that allows for a less disruptive accommodation of the bulky adduct in the anti conformation thus facilitating a greater extent of translesion bypass, tailored to insert D¢ opposite the lesion. Elucidation of specific function–structure relationships involved in DNA adduct bypass by Y-family DNA polymerases may help lend insight into the mechanisms by which these adducts are processed in vivo, including mutagenic routes that may lead to the initiation and progression of carcinogenesis.

Acknowledgments—We thank Dr. Roger Woodgate for providing purified DpO4 enzyme for biochemical assays. Simulations were run using computer time granted from the National Partnership for Advanced Computational Infrastructure (NPACI) at the University of Michigan Center for Advanced Computing, and support at this institution is greatly appreciated.

References

1. Boudsocq, F., Iwai, S., Hanaoka, F., and Woodgate, R. (2001) Nucleic Acids Res. 29, 4607–4616
2. Ling, H., Boudsocq, F., Woodgate, R., and Yang, W. (2001) Cell 107, 91–102
3. Suzuki, N., Ohashi, E., Kolbanovskiy, A., Geacintov, N. E., Grollman, A. P., Ohmori, H., and Shibutani, S. (2001) Biochemistry 40, 6100–6106
4. Rechkoblit, O., Zhang, Y., Guo, D., Wang, Z., Amin, S., Krasinsky, J., Lowene, N., and Geacintov, N. E. (2001) J. Biol. Chem. 277, 30488–30494
5. Bessen, A., and Fuchs, R. P. (2000) EMBO J. 19, 3887–3897
6. Napolitano, S., Janel-Bintz, R., Wagner, J., and Fuchs, R. P. (2000) EMBO J. 19, 6259–6270
7. Pages, V., and Fuchs, R. P. (2002) Oncogene 21, 8957–8966
8. Kool, E. T. (1998) Biopolymers 48, 3–17
9. Kool, E. T. (2002) Annu. Rev. Biochem. 71, 191–219
10. Kool, E. T., Morales, J. C., and Guckian, K. M. (2000) Angew Chem. Int. Ed. Engl. 39, 999–1008
11. Stetzer, T. (1999) J. Biol. Chem. 274, 17385–17388
12. Stetzer, T., Rychlewski, R., and Kuncz, K. (2000) J. Biol. Chem. 275, 4972–4979
13. Double, S., and Ellenberger, T. (1998)Curr. Opin. Struct. Biol. 8, 704–712
14. Double, S., Sawaya, M. R., and Ellenberger, T. (1999) Structure Fold Des 7, 821–834
15. Vaisman, A., Frank, E. G., McDonald, J. P., Tissier, A., and Woodgate, R. (2002) Mutat. Res. 510, 9–22
16. Kukowska, R., Bebenek, K., Boudsocq, F., Woodgate, R., and Kunkel, T. A. (2002) J. Biol. Chem. 277, 19833–19838
17. Gonzalez, M., and Woodgate, R. (2002) Bioessays 24, 141–148
18. Frank, E. G., Sayer, J. M., Kroth, H., Ohashi, E., Ohmori, H., Jerina, D. M., and Woodgate, R. (2002) Nucleic Acids Res. 30, 5284–5292
19. Duvachelle, J. B., Blance, L., Fuchs, R. P., and Cordonnier, A. M. (2002) Nucleic Acids Res. 30, 2061–2067
20. Boudsocq, F., Ling, H., Yang, W., and Woodgate, R. (2002) DNA Repair (Amst) 1, 343–358
21. Vaisman, A., and Woodgate, R. (2001) EMBO J. 20, 6520–6529
22. Baynton, K., and Fuchs, R. P. (2000) Trends Biochem. Sci. 25, 74–79
23. Woodgate, R. (1999) Genes Dev. 13, 2191–2195
24. Ogi, T., Shinkai, Y., Tanaka, K., and Ohmori, H. (2002) Proc. Natl. Acad. Sci. U.S.A. 99, 15548–15553
25. Goodman, M. F. (2002) Annu. Rev. Biochem. 71, 17–50
26. Ohmori, H., Friedberg, E. C., Fuchs, R. P., Goodman, M. F., Hanaoka, F., Hinkle, D., Kunkel, T. A., Lawrence, C. W., Lewnez, Z., Nomih, L., Prakash, L., Prakash, S., Tod, T., Walker, G. C., Wang, Z., and Woodgate, R. (2001) Mol. Cell 7, 8–7
27. Goodman, M., and Tippins, B. (2000) Nat. Rev. Mol. Cell Biol. 1, 101–109
28. Weinberg, R. A. (2000) Science 287, 62–69
29. Scherberka-P., P., Bebenek, K., and Kunkel, T. A. (2003) Sci. Aging Knowl Edge (2003), RES
30. Perriu, J., Liska, P., Hansus, C., Theinap, A., and Felix, G. (1993) in Proc. 13th Int. Symp. PolyNucl. Aromatic Hydrocarbons (Garrigues, P., and Lamotte, M., eds) pp. 337–346, Gordon and Breach Science Publishers, Langhorne, PA
31. Grimmer, G. (1993) in Proc. 13th Int. Symp. PolyNucl. Aromatic Hydrocarbons (Garrigues, P., and Lamotte, M., eds) pp. 31–41, Gordon and Breach Science Publishers, Langhorne, PA
32. Grimmer, G., Jacob, J., Dettbarn, G., and Naucke, K. W. (1997) Int. Arch. Occup. Environ. Health 69, 231–239
33. Cook, J. H., Hewett, C. L., and Hieger, I. (1933) J. Chem. Soc. 385–405
91. Mao, B., Xu, J., Li, B., Margulis, L. A., Smirnov, S., Ya, N. Q., Courtney, S. H., and Geacintov, N. E. (1995) Carcinogenesis 16, 357–365
92. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000) Nucleic Acids Res. 28, 235–242
93. Xie, X. M., Geacintov, N. E., and Broyde, S. (1999) Biochemistry 38, 2956–2968
94. Pearlman, D. A., Case, D. A., Caldwell, J. W., Ross, W. S., Cheatham, T. E., DeBolt, S., Ferguson, D. M., Seibel, G. L., and Kollman, P. A. (1995) Comp. Phys. Comm. 91, 1–41
95. Hingerty, B., Ritchie, R., Ferrell, T., and Turner, J. (1985) Biopolymers 24, 427–439
96. Mezei, M. (1997) J. Comp. Chem. 18, 812–815
97. Jorgensen, W., Chandrasekhar, J., Madura, J., Impey, R., and Klein, M. (1983) J. Chem. Phys. 79, 926–935
98. Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K., Ferguson, D., Spellmeyer, D., Fox, T., Caldwell, J., and Kollman, P. A. (1995) J. Am. Chem. Soc. 117, 5179–5197
99. Cheatham, T., Cieplak, P., and Kollman, P. (1999) J. Biomolec. Struct. Dynam. 16, 845–862
100. Darden, T. A., York, D., and Pedersen, L. (1993) J. Chem. Phys. 98, 10089–10092
101. Harvey, S. C., Tan, R. K.-Z., and Cheatham, T. E. (1996) J. Comp. Chem. 19, 726–740
102. Connolly, M. L. (1983) Science 221, 709–713
103. Kokoska, R. J., McCullough, S. D., and Kunkel, T. A. (2003) J. Biol. Chem. 30, 30
104. Osheroff, W. P., Beard, W. A., Yin, S., Wilson, S. H., and Kunkel, T. A. (2000) J. Biol. Chem. 275, 28033–28038
105. Kannouche, P., Fernandez de Henestrosa, A. R., Coull, B., Vidal, A. E., Gray, C., Zicha, D., Woodgate, R., and Lehmann, A. R. (2002) EMBO J. 21, 6246–6256
106. Volk, D. E., Thiruvanantham, V., Rice, J. S., Luxon, B. A., Shah, J. H., Yagi, H., Sayer, J. M., Yeh, H. J., Jerina, D. M., and Gorenstein, D. G. (2003) Biochemistry 42, 1410–1420
107. Geacintov, N., Cosman, M., Hingerty, B., Amin, S., Broyde, S., and Patel, D. (1997) Chem. Res. Toxicol. 10, 111–146
108. Tan, J., Geacintov, N. E., and Broyde, S. (2000) Chem. Res. Toxicol. 13, 811–822
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
Rebecca A. Perlow-Poehnelt, Ilya Likhterov, David A. Scicchitano, Nicholas E. Geacintov and Suse Broyde

J. Biol. Chem. 2004, 279:36951-36961.
doi: 10.1074/jbc.M404332200 originally published online June 21, 2004

Access the most updated version of this article at doi: 10.1074/jbc.M404332200

Alerts:
• When this article is cited
• When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

Supplemental material:
http://www.jbc.org/content/suppl/2004/07/08/M404332200.DC1

This article cites 104 references, 19 of which can be accessed free at
http://www.jbc.org/content/279/35/36951.full.html#ref-list-1