



# Molecular modeling benzo[a]pyrene N<sup>2</sup>-dG adducts in the two overlapping active sites of the Y-family DNA polymerase Dpo4

Sushil Chandani, Edward L. Loechler\*

Biology Department, Boston University, Boston, MA 02215, United States

Received 2 March 2006; received in revised form 3 May 2006; accepted 3 May 2006

Available online 9 May 2006

## Abstract

The potent, ubiquitous environmental mutagen/carcinogen benzo[a]pyrene (B[a]P) induces a single major adduct [+ta]-B[a]P-N<sup>2</sup>-dG, whose bypass in most cases results in either no mutation (dCTP insertion) or a G → T mutation (dATP insertion). Translesion synthesis (TLS) of [+ta]-B[a]P-N<sup>2</sup>-dG generally requires DNA polymerases (DNAPs) in the Y-family, which exist in cells to bypass DNA damage caused by chemicals and radiation. A molecular dynamics (MD) study is described with dCTP opposite [+ta]-B[a]P-N<sup>2</sup>-dG in Dpo4, which is the best studied Y-family DNAP from a structural point of view. Two orientations of B[a]P-N<sup>2</sup>-dG (BPmi5 and BPmi3) are considered, along with two orientations of the dCTP (AS1 and AS2), as outlined next. Based on NMR studies, the pyrene moiety of B[a]P-N<sup>2</sup>-dG is in the minor groove, when paired with dC, and can point toward either the base on the 5'-side (BPmi5) or the 3'-side (BPmi3). Based on published X-ray structures, Dpo4 appears to have two partially overlapping active sites. The architecture of active site 1 (AS1) is similar to all other families of DNAPs (e.g., the shape of the dNTP). Active site 2 (AS2), however, is non-canonical (e.g., the β- and γ-phosphates in AS2 are approximately where the α- and β-phosphates are in AS1). In the Dpo4 models generated herein, using the BPmi3 orientation the pyrene moiety of [+ta]-B[a]P-N<sup>2</sup>-dG points toward the duplex region of the DNA, and is accommodated without distortions in AS1, but with distortions in AS2. Considering the BPmi5 orientation, the pyrene moiety points toward the ss-region of DNA in Dpo4, and sits in a hole defined by the fingers and little fingers domain (“chimney”); BPmi5 is accommodated in AS2 without significant distortions, but poorly in AS1. In summary, when dCTP is paired with [+ta]-B[a]P-N<sup>2</sup>-dG in the two overlapping active sites in Dpo4, the pyrene in the BPmi3 orientation is accommodated better in active site 1 (AS1), while the pyrene in the BPmi5 orientation is accommodated better in AS2. Finally, we discuss why Y-family DNAPs might have two catalytic active sites.

© 2006 Elsevier Inc. All rights reserved.

**Keywords:** Cancer; Genotoxins; Mutations; Adducts; Lesion; DNA polymerases; Dpo4

## 1. Introduction

B[a]P is a potent mutagen and carcinogen, and is an example of a polycyclic aromatic hydrocarbon (PAH), which is a class of substances produced by incomplete combustion that are found ubiquitously in the environment [1–7]. Most carcinogens are

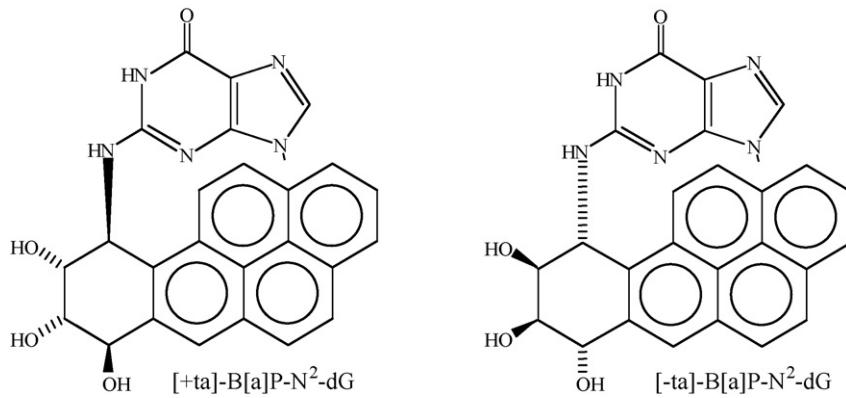
active by causing mutations, and mutations induced by PAHs in general, and B[a]P in particular, have been shown to cause the kinds of mutations that might be relevant to carcinogenesis (representative Refs. [8–12]), and may in fact be important in human cancer (e.g., [13] and references therein). B[a]P mutagenesis has been extensively studied, and mutational spectra with the biologically relevant metabolite (+)-anti-B[a]PDE have been determined in *Escherichia coli* [14,15] and in mammalian (CHO) cells ([16,17] and references therein).

The major adduct formed by (+)-anti-B[a]PDE in cells is [+ta]-B[a]P-N<sup>2</sup>-dG (Fig. 1). Non-mutagenic translesion synthesis (TLS) of [+ta]-B[a]P-N<sup>2</sup>-dG requires both DNAP IV and DNAP V in *E. coli*, and a strong case can be made that they do insertion and extension, respectively [18–22]. The stereochemically related adduct [-ta]-B[a]P-N<sup>2</sup>-dG requires DNAP IV only [18]. (Although not strictly correct, we will refer to [+ta]- and [-ta]-B[a]P-N<sup>2</sup>-dG as being “mirror image”

**Abbreviations:** B[a]P, benzo[a]pyrene; (+)-anti-B[a]PDE, 7R,8S-dihydroxy-9S,10R-epoxy-7,8,9,10-tetrahydrobenzo[a]pyrene; [+ta]-B[a]P-N<sup>2</sup>-dG, the major adduct of (+)-anti-B[a]PDE, formed by trans addition of N<sup>2</sup>-dG to (+)-anti-B[a]PDE; [-ta]-B[a]P-N<sup>2</sup>-dG, the major adduct formed by trans addition of N<sup>2</sup>-dG to (-)-anti-B[a]PDE PAH, polycyclic aromatic hydrocarbon; TLS, translesion synthesis: the insertion of a base opposite a DNA adduct, as well as subsequent elongation; DNAP, DNA polymerase; AS1, active site 1; AS2, active site 2

\* Corresponding author. Tel.: +1 617 353 9259; fax: +1 617 353 6340.

E-mail address: loechler@bu.edu (E.L. Loechler).

Fig. 1. Structures of [+ta]- and [-ta]-B[a]P-N<sup>2</sup>-dG.

adducts.) G → T mutational TLS (dATP insertion) for [+ta]-B[a]P-N<sup>2</sup>-dG depends on DNAP V only [23], and [-ta]-B[a]P-N<sup>2</sup>-dG requires DNAP V, although we have not yet established whether DNAPs II or IV might have a role with the latter [18].

These findings have focused our attention on trying to understand how B[a]P-N<sup>2</sup>-dG adducts are bypassed by DNAPs IV and V, which are in the Y-family of DNA polymerases. Cells possess an unexpectedly large number of DNA polymerases; e.g., human cells have at least fifteen, while *E. coli* has at least five [24–27]. The cellular significance of some of these DNA polymerases (DNAPs) can be understood by noting that DNA is constantly subjected to insult by chemicals and radiation, and most lesions that are not removed by DNA repair block replicative DNA polymerases. To avoid such lethal blockage, cells possess a class of so-called “lesion-bypass DNA polymerases”, which conduct TLS [24–37]. Most of the DNAPs capable of TLS belong to the Y-family, which is a newly recognized superfamily of evolutionarily related DNAPs [24–37], where human cells have at least three Y-family members (DNAPs  $\eta$ ,  $\iota$  and  $\kappa$ ), while *E. coli* has two (DNAPs IV and V).

Y-family DNA polymerases range in size from ~350 aa to over 1000 aa, and share five conserved motifs (*I*–*V*) in their N-terminal ~200 amino acids, which includes all of the residues in the template/dNTP-binding pocket and polymerase catalytic active site [29]. They also share a number of important structural features, as revealed in X-ray crystallographic studies of the archaeal DNA polymerases Ddh [38] and Dpo4 [39–47], yeast DNAP  $\eta$  [48], and human DNAPs  $\iota$  [49] and  $\kappa$  [50]. As with all DNA polymerases, Y-family members have thumb, palm and fingers domains. However, their fingers and thumb are relatively small or “stubby”, resulting in greater solvent accessibility to the template/dNTP-binding pocket [35], which is more spacious — presumably to accommodate the bypass of bulky and/or deformed DNA lesions. Furthermore, Y-family polymerases grip DNA with an additional domain, which has been called the “wrist” [38], the “polymerase-associated domain (PAD)” [48] or the “little finger domain” [39].

Cells are believed to have multiple lesion-bypass DNA polymerases, because of the structural and conformational diversity of DNA lesions, where the lesion-bypass pathway (be it error-free or mutagenic) is likely to depend on lesion

structure/conformation, which in turn dictates what DNAP is involved in the TLS event. For example, DNAP  $\eta$  accurately bypasses several C- and T-containing cyclopyrimidine dimers (CPDs) [51–56], while it principally inserts dATP opposite B[a]P adducts [57–60], which would lead to G → T mutations in cells. In contrast, DNAP  $\kappa$  bypasses benzo[a]pyrene adducts accurately (dCTP insertion) *in vitro* [57,61–65], which may be biologically relevant, given that DNAP  $\kappa$  appears to be responsible for minimizing B[a]P-induced mutations in a mammalian cell [64,65].

*E. coli* DNAP IV and hDNAP  $\kappa$  both insert dCTP opposite [+ta]-B[a]P-N<sup>2</sup>-dG (see above, [18–22]), and this and a similarity of insertion opposite five other DNA lesions suggested to us that they are functional orthologs [66]. Recent work on a similar mutation in both DNAPs IV and  $\kappa$  strongly reinforces this conclusion [22]. *E. coli* DNAP V and hDNAP  $\eta$  both insert dATP opposite [+ta]-B[a]P-N<sup>2</sup>-dG (see above), and also show similarity of insertions opposite a variety of other DNA lesions, suggesting that they are functional orthologs [66]. We used homology modeling to build structures and noted that DNAPs IV and  $\kappa$  have similar amino acids in their active sites – as do DNAPs V and  $\eta$  – that might explain how they are functional orthologs (see Section 4.3). Our models of DNAPs IV and V were built from X-ray coordinates for Dpo4, which is the best studied Y-family DNAP.

The dNTP can be in different shapes in the active site of Dpo4 [44]: a “chair-like” shape, which is similar to the shape of dNTPs in the active sites of other families of DNAPs, or a “goat tail-like” shape, which is non-canonical. Herein, we argue that both dNTP shapes are likely to be catalytically active, and, accordingly, we refer to these dNTP shapes as being in distinct, but overlapping, active site 1 (AS1) and active site 2 (AS2), respectively. Furthermore, we describe a molecular modeling study with B[a]P-N<sup>2</sup>-dG adducts in AS1 and AS2 of Dpo4. A 5'-TGC sequence relevant to our recent mutagenesis studies was chosen [18], and [+ta]- and [-ta]-B[a]P-N<sup>2</sup>-dG were both studied. B[a]P-N<sup>2</sup>-dG adducts can form more-or-less normal adduct-G:C base pairs if the B[a]P moiety is in the minor groove in either of two orientations: with the B[a]P pointing toward the base pair on the 5'-side (BPmi5) or the 3'-side (BPmi3). In NMR studies, [+ta]-B[a]P-N<sup>2</sup>-dG adopts the BPmi5 conformation [67,68], while the mirror image adduct

[–ta]-B[a]P-N<sup>2</sup>-dG adopts the BPmi3 conformation [69], where both show more-or-less normal adduct-G:C base pairing. Accordingly, we docked both the BPmi5 and BPmi3 conformations for both [+ta]- and [–ta]-B[a]P-N<sup>2</sup>-dG in AS1 and AS2 of Dpo4, and molecular dynamics (MD) was conducted at 300 K. The refined structures show that AS1 accommodates BPmi3 better, while AS2 accommodates BPmi5 better. The potential implications of these observations are discussed.

## 2. Methodology

Graphics were visualized using InsightII (Accelrys Inc. [70]) on a SGI O2. Molecular modeling and molecular dynamics were done with CHARMM 27 [71–73] and calculations were done either on an SGI Origin200 or one of Boston University's IBM Power4 systems (p690 or p655). X-ray structures were obtained from the RSCB Protein Data Bank [74]. Dpo4 coordinates from 1JX4 [39] were used for structures in active site 1 (AS1), where the ddATP was converted to dCTP using the approach we have used in the past [66]. Of Dpo4 structures, 1JX4 has the best resolution (1.70 Å), and overall has the most normal structure of its DNA, where other AS1 structures have lesions, bulges and/or gaps (filled by non-DNA-containing components) in or near the active site. The primer-O3'-to-P $\alpha$ -ddATP distance is close to reaction-ready in 1JX4 (~3.5 Å), while it is larger (~4.41 Å or greater) in other AS1 structures (Table S1). Dpo4 coordinates from the A-peptide of 1RYS [40] were used for structures in active site 2 (AS2), where dATP was converted to dCTP. Of Dpo4 structures in the AS2 category, 1RYS-A was chosen, because it is most different from AS1 with the base of its dNTP in the [n + 1] position, presumably because it has a lesion (TT-CPD) in the template active site, whereas AS2 structures without lesions in the template active site have the base in the [n] position (see Section 3). AS2 of lesion-free 1JXL [35], was also used as a guide, although none of its coordinates were included in the starting structures. We note that the assignment of O1 $\alpha$  and O2 $\alpha$  in the X-ray structures for AS1 and AS2 were the reverse of the assignments in CHARMM, but when reporting on our modeled structures we have conformed to X-ray convention. Likewise O2 $\gamma$  and O3 $\gamma$  are reversed in AS1.

The bases in the X-ray structure for 1JX4 or 1RYS-A were changed to the 5-TGC sequence context (5'-CCTG\*CAG-GAATTCACTG/5'-CAGTGAATTCTG) that we have been using in our most recent mutagenesis studies [18]. Missing hydrogens were added using Biopolymers in InsightII. A minimized octamer containing [+ta]-B[a]P-N<sup>2</sup>-dG in the BPmi3 conformation in the identical 5'-TGC sequence was superimposed on the DNA in the X-ray structure. Highest priority was given to positioning the dC paired with [+ta]-B[a]P-N<sup>2</sup>-dG so that it superimposed closely with the dCTP in the Dpo4 structure. The template base in the Dpo4 structure was removed and replaced with [+ta]-B[a]P-N<sup>2</sup>-dG, as follows. (AS1) The O5'-C5' bond was rotated on [+ta]-B[a]P-N<sup>2</sup>-dG to decrease the distance between the phosphate and the O3' of the 5'-T in the Dpo4 structure from 1.98 to 1.70 Å, which was

judged adequate given a typical bond length of 1.58 Å. O3' of [+ta]-B[a]P-N<sup>2</sup>-dG was 1.64 Å from the phosphate on the 3'-C, which was deemed satisfactory. Finally, the adduct bond was rotated ~30° in a direction that minimized van der Waals contact between the B[a]P moiety and the protein. (AS2) Dihedral angles in the sugar-phosphate backbone were rotated in both the 3'-C and next base (i.e., 3'-3'-A), such that the distance was proper (i.e., 1.58 Å) from the O3' of the 3'-C to the phosphorous of [+ta]-B[a]P-N<sup>2</sup>-dG; the position of the base in the 3'-C was not changed, which resulted in a slight elongation of the C1'-N1 distance from 1.45 to 1.70 Å. Dihedral angles in the sugar-phosphate backbone were also rotated in the 5'-T to give a final distance of 1.70 Å between the O5' of [+ta]-B[a]P-N<sup>2</sup>-dG and the phosphorous of the 3'-T. The adduct bond was rotated slightly in order to minimize van der Waals contact between the B[a]P moiety and the protein. Similar changes were carried out on all eight possible combinations involving both adducts (+ta- and [–ta]-B[a]P-N<sup>2</sup>-dG), both conformations (BPmi5 and BPmi3), and both active sites (AS1 and AS2). To the 183 water molecules and the ions in the 1JX4 X-ray structure, two layers of TIP3P water were added (using the Soak utility in InsightII) around the DNA not in contact with protein, as the surface of this DNA is recessed in comparison to the surface of the protein. Finally, three more layers of water were added around the protein and the hydrated DNA, which gave a smooth, coherent solvent shell with 2265 total waters. Water was added similarly in all other structures. Water alone was minimized (2000 steps of ABNR) to allow it to settle into the protein, and then the entire ensemble was minimized (200 steps of ABNR) to allow the bases in the new DNA sequence to adapt to the rest of the structure. Molecular dynamics (MD) trajectories were conducted with the Verlet algorithm in CHARMM, using the velocity Verlet integration method (time step 1 fs) with electrostatic, van der Waals interactions and non-bonded Lennard-Jones interactions smoothed to zero at 11 Å. The SHAKE algorithm was used to constrain bonds with hydrogen atoms using a tolerance of 0.0005 Å. A full charge of –1.0 was included on phosphates. Following heating to 300 K over 20 ps using the Verlet algorithm, MD was conducted at 300 K for 240 ps with constraints on the distances of atoms chelated to each magnesium in the active site (30 kcal/mol), after which MD was continued without constraints for an additional 1000 ps. Key atom-to-atom distances over the final 1000 ps are shown in Figure S2 (Supplemental materials). Prior to viewing structures, they were minimized (200 steps of ABNR) to remove distortions due to the elevated temperatures.

We must comment on why it did not make sense to choose either X-ray structure containing a B[a]P adduct [39] as a starting structure. (1) These structures contain a minor N<sup>6</sup>-dA adduct, which places the B[a]P moiety on both the wrong base and the wrong atom, since we are interested in understanding the major mutagenic adduct, which has B[a]P attached at N<sup>2</sup>-dG. (2) The dNTP is not paired with the dA moiety of the adduct in either of the two B[a]P-containing peptides in 1SOM. (3) The B[a]P moiety of the N<sup>6</sup>-dA adduct is either in the developing major groove (1SOM-A) or is intercalated (1SOM-B), which contrasts with B[a]P-N<sup>2</sup>-dG paired with dCTP, in

Table 1

Comparisons of Dpo4 X-ray structures, which appear to sort into two active sites (AS1 and AS2)

AS1	Template	Nucleotide	P1	P2	P3	P4	
1SOM-B	T	dATP		α	β	γ	
1N56-AB	T	dATP		α	β	γ	
1SOO-AB	A	dTTP		α	β	γ	
1SON	G	dCTP		α	β	γ	
1S10	G	dCTP		α	β	γ	
2AGQ	T	dATP		α	β	γ	
2ATL-AB	G	dCTP		α	β	γ	
2C22	C	dGTP		α	β	γ	
2ASD-AB	GO	dCTP		α	β	γ	
2C2E	GO	dCTP		α	β	γ	
2C2R	GO	dCTP		α	β	γ	
1RYS-B <sup>a</sup>	3'-TT-CPD	dATPs		α	β	γ	
AS2	Template	Nucleotide	P1	P2	P3	P4	
1JXL	G	ddCTP		α	β	γ	
2AGP-A	T	dGTP		α	β	γ	
2AGP-B	–	dATP		α	β	γ	
2C28	C	ddGTP		α	β	γ	
1RYR	5'-TT-CPD	dATP		α	β	γ	
1RYS-A	3'-TT-CPD	dATPs		α	β	γ	
Unusual	Template	Nucleotide	P1	P2	P3	P4	Other
1N48	T	dATP	α	β		γ	
1JX4	T	ddADP		α	β		
1S9F	T	ddADP		α	β		
1S97-ABCD	T	dATP		α	γ	β	
2C2D	GOs	dATP		α	γ	β	
1SOM-A	T	dATPs	γ	α	β		
2BRO	C	dGTP	γ	α		β	
2BQR	T	dATP	α	β		γ	
2BQU	T	dATP	α	β		γ	

AS1 is a canonical DNA polymerase active site, while AS2 is unusual (see text). “1S97-ABCD” (as an example) indicates the RCSB Protein Data Bank designation “1S97”, while “ABCD” indicates the four subunits in the crystal unit cell. The “s” in “dATPs” and “GOs” indicates the *syn*-glycosidic orientation, while “GO” stands for 8-oxoguanine.

<sup>a</sup> The α-phosphate in 1RYS-B is in a slightly different position (Fig. 3A).

which the B[a]P moiety must be in the minor groove. Thus (regrettably), the 1S0M structures do not reveal anything of interest about insertion opposite B[a]P-N<sup>2</sup>-dG and are not directly helpful.

### 3. Results

#### 3.1. Dpo4 appears to have two overlapping active sites (AS1 and AS2)

Twenty-four Dpo4 structures are currently in the RCSB Protein Data Bank, including with: undamaged DNA [39: 1JX4 & 1JXL], a TT-CPD [40: 1RYR & 1RYS], a B[a]P adduct [41: 1S0M], an AP site [42: 1N48, 1N56, 1S0N, 1SOO and 1S10], a G:T mismatch [43: 1S97 and 1S9F; 44: 2AGO, 2AGP and 2AGQ], 1, N<sup>2</sup>-ethenoguanine [45: 2BQ3, 2BQR, 2BQU and 2BRO] and 8-oxoguanine [46: 2C22, 2C28, 2C2D, 2C2E and 2C2R; 47: 2ASD, 2ASJ, 2ASL, 2ATL and 2AUO]. Some unit cells contain more than one monomer (designated, e.g., 1RYS-A versus 1RYS-B), resulting in thirty-four distinct structures (Table 1). The structures may contain a dNTP, a ddNTP or a ddNDP, but for simplicity we will refer to these as “dNTPs” when discussing them generically.

dNTPs in the active site of Dpo4 can adopt different shapes, which have been called “chair-like” and “goat tail-like” [44]. Herein, we argue that these shapes represent dNTPs in distinct, but overlapping active sites, where each is catalytically active (see Section 4.1). Accordingly, we refer to them as active site 1 (AS1, “chair-like” shape) and active site 2 (AS2, “goat tail-like” shape).

AS1 (Fig. 2) is canonical in that dNTP positioning is virtually identical to dNTP positioning in the active sites of other DNAP families, while active site 2 (AS2, Fig. 2) has novel dNTP positioning. As an example, the dNTP from 1SOM-B (AS1) was superimposed on 1RYS-A (AS2) based on aligning amino acids in Dpo4, and a top view reveals that the sugar-triphosphates arch in different ways (Fig. 2A). In a side

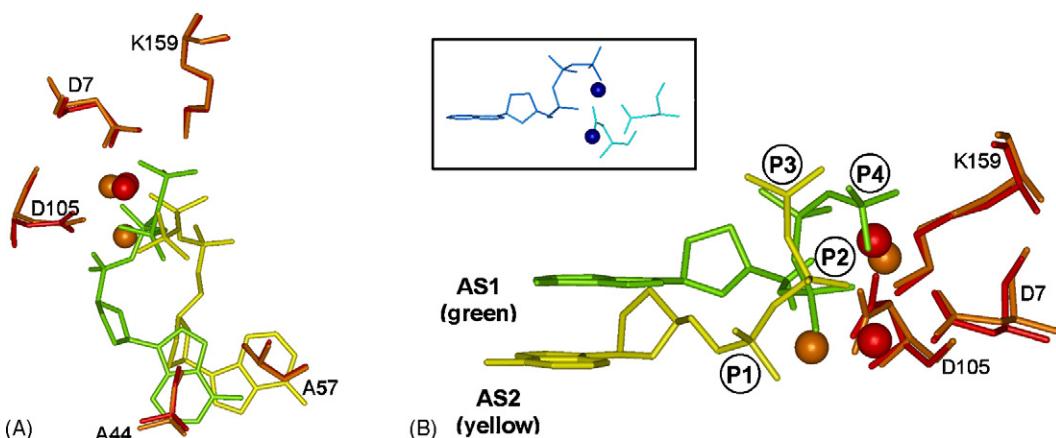


Fig. 2. Side view (left) and top view (right) of active site 1 (AS1, green) vs. active site 2 (AS2, yellow) in Dpo4. Key amino acids from 1SOM-B (red), which has dATP in AS1, were superimposed on the same amino acids in 1RYS-A (brown), which has dATP in AS2. The spheres are divalent cations for AS1 (red) and AS2 (brown). The insert shows that the catalytic active site of T7 DNAP resembles AS1 (green) and not AS2 (yellow). Coordinates are from references cited in the text.

view, 1SOM-B (AS1, green, Fig. 2B) shows the expected dNTP orientation based on comparisons to other DNAPs (e.g., T7 DNAP, insert Fig. 2), which is why the 1SOM-B orientation can be thought of as canonical. In 1RYS-A (AS2, yellow, Fig. 2B), the dNTP is lower down with the base approximately in the  $[n + 1]$  position. Based on this superimposition, there appear to be four potential phosphate-binding sites (P1–P4 in Fig. 2B), where AS1 uses P2–P4, while AS2 uses P1–P3. Table 1 summarizes the Dpo4 X-ray structures with dNTPs in AS1 versus AS2, as well as several other dNTP orientations that almost certainly cannot be catalytically active, given that the phosphates are not oriented such that catalysis would be possible. Fig. 3A shows multiple AS1 structures superimposed. (The AS1 structure in yellow is slightly non-conforming and is considered in Section 4.) Fig. 3B shows multiple AS2 structures, in which the base of the dNTP appears to be able to adopt multiple orientations and can lie anywhere from the  $[n]$  to the  $[n + 1]$  position. AS2 structures without a lesion in the template position have the base of the dNTP in the  $[n]$  position (Fig. 3B: 1JXL [green], 2AGP-A [pink], 2AGP-B [not shown], and 2C28 [blue]), while AS2 structures with a lesion in the template (TT-CPD) have the base in the  $[n + 1]$  position (Fig. 3B: 1RYS-A [yellow]) or between  $[n]$  and  $[n + 1]$  (Fig. 3B: 1RYR [red]).

### 3.2. Modeling $[+ta]$ - and $[-ta]$ -B[a]P-N<sup>2</sup>-dG in AS1 and AS2 of Dpo4

In our molecular biological studies, we are currently working on both the G → T mutagenesis pathway (dATP insertion) and the non-mutagenic pathway (dCTP insertion) [18]. To support the latter effort, we initiated a modeling project to study dCTP opposite  $[+ta]$ - and  $[-ta]$ -B[a]P-N<sup>2</sup>-dG in the BPmi5 and BPmi3 conformations. We have begun by docking  $[+ta]$ - and  $[-ta]$ -B[a]P-N<sup>2</sup>-dG in Dpo4 initially, because it is the best understood Y-family DNAP with a sizeable collection of X-ray structures. Initial structures were generated and refined using molecular dynamics (Section 2). Even though we cannot be certain whether AS2 is a viable second active site, we must assume it might be (see Section 4.1) and take it into account in our modeling work. MD runs were conducted on all eight possible combinations involving both adducts ( $[+ta]$ - and  $[-ta]$ -B[a]P-N<sup>2</sup>-dG), both conformations (BPmi5 and BPmi3), and both active sites (AS1 and AS2), as well as for the unadducted controls (AS1 and AS2), for a total of ten structures. Distances between certain key atoms over the final 1000 ps of MD are plotted in Figure S2 with average values (and standard deviations) given in Table S2 (Supplemental materials). In Section 2 we discuss three reasons why it

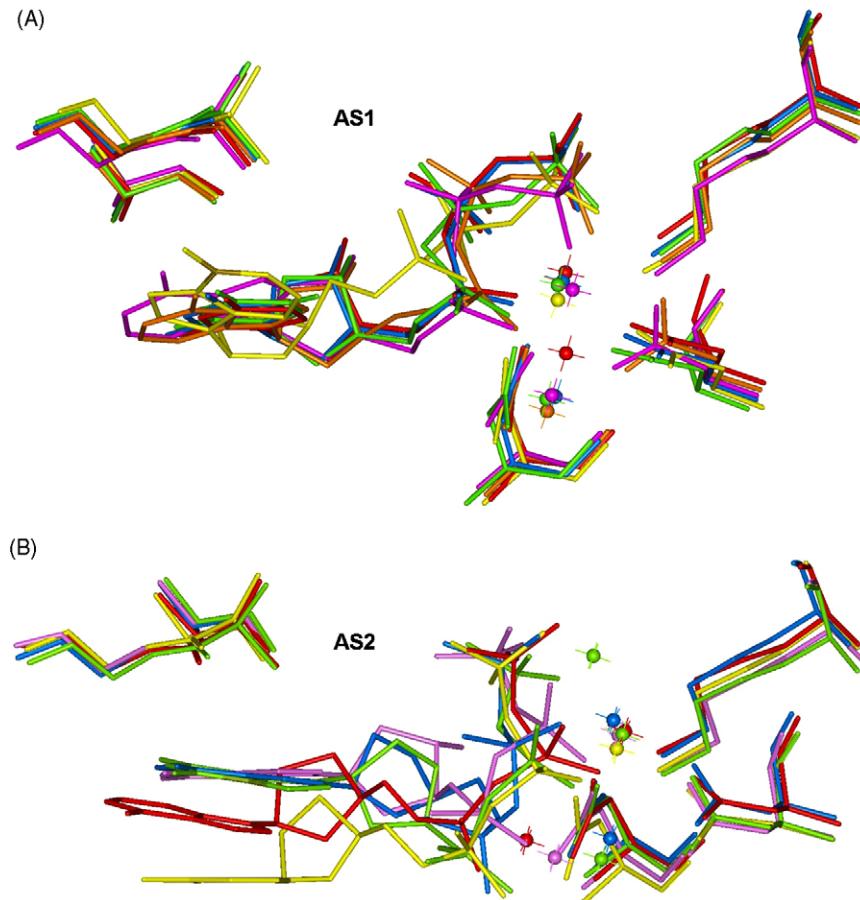


Fig. 3. Superimposing various dNTPs representative examples of AS1 (panel A) and AS2 (panel B). AS1: 1SOM-B (brown); 1N56-AB (purple); aS0O-AB (green); 1SON (red); 1S10 (blue); 1RYS-B (yellow). AS2 1JXL: (green); 2AGP-A (pink); 2C28 (blue); 1RYR (red); 1RYS-A (yellow). Structural superimposition depended on aligning the same the amino acids as in Fig. 2A (i.e., A44, A57, D7, D105 and K159).

Table 2

Comparing the modeled structures for [+ta]- and [−ta]-B[a]P-N<sup>2</sup>-dG in the BPmi5 and BPmi3 conformations in Dpo4 for AS1 and AS2

Conformation	B[a]P-N <sup>2</sup> -dG	AS1	AS2
Control	No adduct	4.35 Å/134.4° 3/7 None	3.99 Å/135.6° 3/4 None
BPmi5	[+ta]-	5.59 Å/126.3° 4/7 [n + 1] not base paired	4.00 Å/150.3° 2/4 [n + 1] slight propeller twist
BPmi5	[−ta]-	4.56 Å/140.4° 5/6 dCTP buckled, V32 through B[a]P, [n + 1] not base paired	3.59 Å/150.2 2/3 None
BPmi3	[+ta]-	4.25 Å/154.1° 2/7 [n + 1] distorted base pair	5.57 Å/103.8°* 0/2 [n + 1] not base paired, [n + 1] badly buckled
BPmi3	[−ta]-	4.38 Å/87.9° 4/6 [n + 1] slight buckle	3.98 Å/171.6° 1/3 [n] not base paired, [n] bases not co-planar, [n + 1] not base paired

The first entry (e.g., “4.35 Å/134.4°”) indicates the proximity (distance) of the atoms that form the phosphoester bond (i.e., O3' of the primer terminus and P $\alpha$  of dCTP) and the angle between the nucleophilic atom (O3' of the primer terminus), the electrophilic atom (P $\alpha$  of dCTP) and the leaving group atom (O3 $\alpha$  of dCTP). The second entry (e.g., “3/7”) indicates the number of favorable interactions between the protein and the dCTP that tend to dampen negative charge on the oxygens attached to “P $\alpha$ /P $\beta$ ”, where decreased negative charge on P $\alpha$  makes it more electrophilic, while decreased negative charge on P $\beta$  makes it a better leaving group. The numbers are from Table S3, which is derived from data in Table S2. The third entry (e.g., “none”), indicates structural elements that were suboptimal based on visually inspecting the final structures following MD trajectories.

did not make sense to use either X-ray structure containing a B[a]P adduct [37] as a starting structure.

Two methods were used to assess the reasonableness of the final structures: (1) visual inspection and (2) the monitoring of various parameters that are likely to be revealing about catalysis. Regarding the latter, five kinds of parameters were considered: (i) hydrogen bonding of the G:C base pair between the adduct-dG and the dCTP; (ii) the C1'-C1' distance in the adduct-dG:dCTP base pair; (iii) proximity (distance) of the atoms that must form the phosphoester bond (i.e., O3' of the primer terminus and P $\alpha$  of dCTP); (iv) the angle between the nucleophilic atom (O3' of the primer terminus), the electrophilic atom (P $\alpha$  of dCTP) and the leaving group atom (O3 $\gamma$  of dCTP), which ideally would be 180°; (v) numerous interatomic distances that are likely to maintain a catalytically viable active site, as revealed in the X-ray structures and as discussed in Section 4.1. Figures in Supplemental materials show plots of these distance parameters graphed over the final 1000 ps of MD, while average values for distances during the MD trajectories are listed in Table S3. Table 2 summarizes our observations with a focus on flaws in the structures that might preclude catalysis (analyzed in Section 4).

#### 4. Discussion

Numerous findings show that Y-family DNAPs are involved in mutagenic and non-mutagenic TLS with B[a]P-N<sup>2</sup>-dG adducts (Section 1). Since Dpo4 is by far the best understood Y-family member, it is a solid starting point to begin to investigate

how Y-Family DNAP structure relates to TLS with B[a]P-N<sup>2</sup>-dG. Before analyzing our modeling results, we evaluate whether the non-canonical active site AS2 has the properties necessary to be catalytically active.

##### 4.1. Is the non-canonical site AS2 likely to be catalytically active?

Since the dNTP in AS1 is chair-like, and, thus, resembles the shape of dNTPs in the active sites of DNAPs in all other families, it is reasonable to assume that AS1 is catalytically active. Although the goat tail-like orientation of the dNTP in AS2 may be a non-productive complex, a variety of arguments suggest that it might be a true active site. [Distances associated with the key interactions mentioned below for AS1 and AS2 are given in Table S1 (Supplemental materials).]

- (1) In the catalytic step of a DNA polymerase, the O3' of the primer attacks P $\alpha$ , so these atoms should be close to each other. In AS2, the O3'-P $\alpha$  distance is under 4.0 Å in 4/5 structures (Table S1, range: 3.66–4.69 Å). For reference, van der Waals radii predict a distance of ~3.3 Å between non-bonded O/P atoms, while the CHARMM force field allows a distance of no less than ~3.8 Å without a resultant major energy penalty. Even the best AS1 structures have greater O3'-P $\alpha$  distances (4.41–4.60 Å, Table S1). [We note that 1JX4 has ddATP, but it is AS1-like and has a close O3'-P $\alpha$  distance of ~3.5 Å.] On balance the O3'-P $\alpha$  distance in AS2 structures appears to be closer to reaction-ready on average

than in AS1 structures, adding credence to the notion that AS2 is likely to be a true active site.

(2) To improve reaction rates in DNA polymerases,  $P\alpha$  electrophilicity is enhanced by dampening the negative charge on the oxygens attached to  $P\alpha$  through chelation with divalent cations. In the best AS1 structures, one non-bonded oxygen on  $P\alpha$  ( $O1\alpha$ ) is chelated by two metals (Table S1, Figs. 2A, 3A, and 4A), while several other oxygens ( $O2\alpha$  and  $O5'$ ) are close to metals (Table S1, Figs. 2A, 3A and 4A). (The positioning of atom  $O1\alpha$ ,  $O2\alpha$ , etc., can be found in Figure S1 of Supplemental materials.) In AS2, a divalent cation is in position to dampen two oxygens on  $P1\alpha$  ( $O1\alpha$  and  $O5'$ , Table S1). Based on distances (Figures S1 and S2), we estimate that on average AS1 has 2.75 interactions with oxygens on  $P\alpha$  (Table S2), while AS2 has 1.83. Several water molecules in AS2 X-ray structures appear to be in position to dampen negative

charge on the second non-bonded oxygen of  $P\alpha$  ( $O2\alpha$ , not shown).

- (3) In the DNA polymerase catalytic step, divalent ion chelation to oxygens on  $P\beta$  dampen its negative charge, thus effectively lowering the  $pK_a$  of the  $\beta$ -phosphate leaving group and facilitating bond breaking. In AS1 and AS2, both divalent cations and a collection hydrogen bonding residues interact with oxygens attached to  $P\beta$  (Table S1). On average, AS1 has 8.25 interactions with oxygens on  $P\beta$  (Table S2), while AS2 has 6.33. In AS1, R51 along with the backbone hydrogens of Y10 and Y11 appear important (Fig. 4A). In AS2 the  $\varepsilon$ -amino group of K159 is well positioned to hydrogen bond with both non-bonded oxygens of  $P\beta$  (Fig. 4B), and this lysine is conserved in Y-family DNAPs [66].
- (4) In all DNAPs, two positionally conserved aspartate residues chelate two divalent ions. In Dpo4, D7 and D105 are these

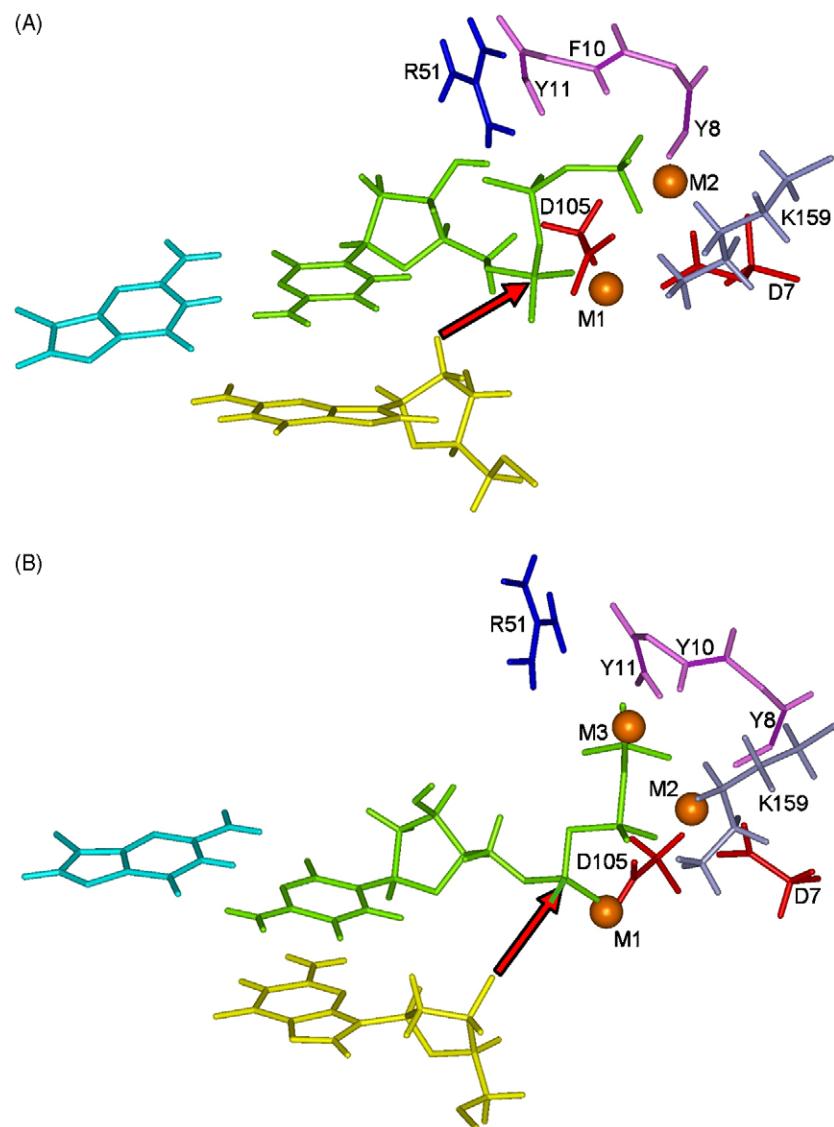


Fig. 4. Dpo4 structures of AS1 (panel A) and AS2 (panel B) with an emphasis on amino acid residues that are important in the catalytic portion of the active site (see text). The AS1 and AS2 structures are from our models with a template unadducted-dG (turquoise) paired with dCTP (green). The red arrow indicates phosphoester bond-making between  $O3'$  of the primer terminus (yellow) and  $P^*$ .

- canonical residues and they are in position to chelate divalent ions in both AS1 and AS2 (Figs. 2–4, Table S1).
- (5) In AS1,  $P\gamma$  is positioned by hydrogen bonding and chelation to a variety of amino acid residues, notably R41. On average, AS1 has 6.0 interactions with oxygens on  $P\gamma$  (Table S2), while AS2 has 7.5. In AS1, R51, K159 and the backbone hydrogen of Y10 appear important, while in AS2 R51 and the backbone hydrogens of Y10 and Y11 appear important (Fig. 4). Interestingly, one AS2 structure (1JXL) had a third divalent cation that appeared to contribute to  $P\gamma$  positioning (M3, Fig. 4B).
- (6) Watson–Crick pairing between the template base and the dNTP are reasonable in most AS1 and AS2 structures, as indicated by the distances associated with base pairing hydrogen bonds and  $C1'-C1'$  distances (Table S1).

In summary, the following can be said about AS1 versus AS2. AS2 appears somewhat more reaction-ready than AS1 based on its shorter average  $O3'-P\alpha$  distance.  $P\alpha$  is stabilized for nucleophilic attack in both AS1 and AS2, although probably AS1 is better. Concerning the leaving group,  $P\beta$  is adequately stabilized in both AS1 and AS2.  $P\gamma$  appears securely positioned in both AS1 and AS2. Base pairing appears reasonable in both AS1 and AS2. Thus, on balance, the protein context around the dNTP in AS2 seems to have structural elements that could satisfy the criteria needed to make and break the key phosphoester bonds as catalyzed by a DNA polymerase.

It is important to note, that interpreting distances in X-ray structures in terms of providing insight about a transition state at the instant of catalysis can be tricky, as illustrated by the following example. The M1-05'-dCTP distance is typically  $\sim 4.0 \text{ \AA}$  in AS1 and AS2 (Table S1), which naively might be interpreted as less than optimal, since magnesium ligand bonds can be as low as  $\sim 2.0 \text{ \AA}$ . However, the geometry around  $P\alpha$  must change from tetrahedral in reactants to trigonal–bipyramidal in the transition state and 05'-dCTP must move toward M1, which will lead to a decrease in distance. Thus, the long distance for M1-05'-dCTP in the reactants ( $\sim 4.0 \text{ \AA}$ ) is actually desirable, because it provides energetic strain that is

encouraging the reactants to proceed toward the transition state, and then on to products, in which the M1-05'-dCTP distance will be much less. Accordingly, the relatively long M1-05'-dCTP distance in reactants ( $\sim 4.0 \text{ \AA}$ ) might be viewed as “good” (i.e., favoring reaction) rather than “bad”.

#### 4.2. B[a]P adducts in AS1 versus AS2, and in BPmi5 versus BPmi3

NMR studies show that B[a]P-N<sup>2</sup>-dG adducts form more-or-less normal adduct-G:C base pairs in duplex DNA when the B[a]P-moiety is in the minor groove, where it points toward the base on the 5'-side (BPmi5) with [+ta]-B[a]P-N<sup>2</sup>-dG and toward base on the 3'-side with [-ta]-B[a]P-N<sup>2</sup>-dG [67–69]. The BPmi5 versus BPmi3 conformations sit very differently in the active site of a Y-family DNAP, such as Dpo4 (Fig. 5). In the BPmi3 conformation the pyrene moiety sits in the large slot between the fingers and little finger domains (Fig. 5A). In the BPmi5 conformation, the pyrene moiety sits in a notch in the fingers domain (Fig. 5B). In fact, when Dpo4 is viewed as a space-filling model, the pyrene (“smoke”) appears to be emerging from a hole (“chimney”) defined by fingers and little fingers domains. The fundamental difference in orientation had an impact on the structures that were generated in our studies (see below), as summarized in Table 2.

The no-adduct controls showed no obvious defects (Table 2, Fig. 4). The  $O3'-P\alpha$  distance in AS1 ( $4.35 \text{ \AA}$ ) was as low as the lowest found in the X-ray structures (Table S1), and AS1 had three interactions that should dampen charge on  $P\alpha$  and seven interactions for  $P\beta$ , which is reasonably similar to the average number of interactions observed in the X-ray structures (Table S2, Section 3.2). The  $O3'-P\alpha$  distance in AS2 ( $3.99 \text{ \AA}$ ) is in the range of the X-ray structures (Table S1), and AS2 had three stabilizing interactions for  $P\alpha$ , which is high compared to the average for the X-ray structures, and four stabilizing interactions for  $P\beta$ , which is low compared to average for the X-ray structures (Table S2, Section 3.2).

The BPmi5 structures seemed reasonable for both [+ta]- and [-ta]-B[a]P-N<sup>2</sup>-dG in AS2, but not in AS1 (Table 2). For example, the  $O3'-P\alpha$  distance ( $5.59 \text{ \AA}$ ) for BPmi5/[+ta]-B[a]P-

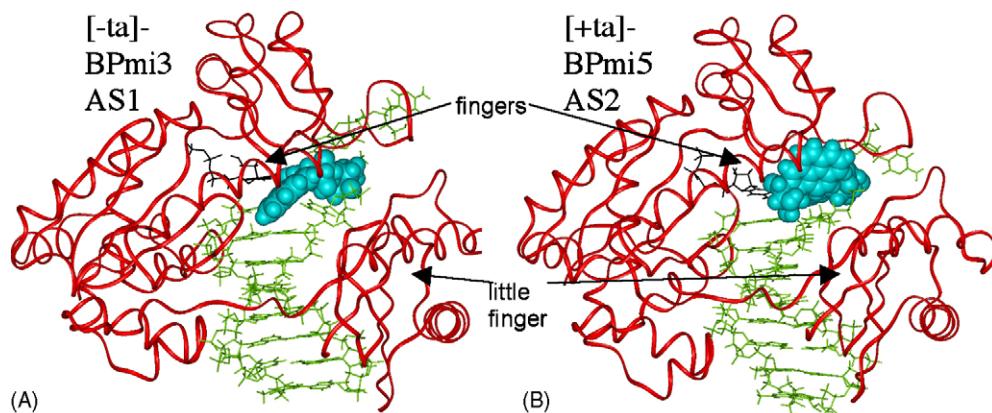


Fig. 5. B[a]P-N<sup>2</sup>-dG in DNA in Dpo4, showing that the pyrene moiety sits in very different places when in the BPmi5 orientation vs. the BPmi3 orientation. Panel A shows the pyrene (CPK, turquoise) in the BPmi3 orientation lying in the slot between the fingers domain and the little finger domain of Dpo4 (AS1). Panel B shows the pyrene (CPK, turquoise) in the BPmi5 orientation lying in a notch in the fingers domain of Dpo4 (AS2).

$\text{N}^2\text{-dG}$  in AS1 is far from reaction-ready, and the  $[n+1]$  base pair is not hydrogen bonded (Table 2).  $[-\text{ta}]\text{-B[a]P-N}^2\text{-dG}$  in the BPmi5 conformation in AS1 is much worse, as the output structure is in fact unacceptable, since the V32 side chain thrusts through the middle of the pyrene moiety in a physically impossible manner. We were unable to find a starting structure for  $[-\text{ta}]\text{-B[a]P-N}^2\text{-dG}$  that would solve this problem. The source of these defects can be understood beginning with the tight, but acceptable, fit of  $[+\text{ta}]\text{-B[a]P-N}^2\text{-dG}$  in AS2 (Fig. 6B), where the pyrene moiety (turquoise) sits in a notch in the fingers domain (shown in Fig. 5B), which is defined by amino acids K78 (gray), Y12 (brown), M76 (blue), A44 (purple) and V32 (red). To get from AS2 (Fig. 6B) to AS1 (Fig. 6A), the template  $[+\text{ta}]\text{-B[a]P-N}^2\text{-dG}$  must move helically up and to the right (arrow), which forces the pyrene moiety into this crowded region of the protein resulting in the distortions when in the AS1 active site (Table 2).

The BPmi3 structures seemed reasonable for both  $[+\text{ta}]$ - and  $[-\text{ta}]\text{-B[a]P-N}^2\text{-dG}$  in AS1, but not in AS2 (Table 2). The O3'-

P $\alpha$  distance (5.57 Å) for BPmi3/ $[+\text{ta}]\text{-B[a]P-N}^2\text{-dG}$  in AS2 is far from reaction-ready and numerous other defects are also apparent (Table 2), while BPmi3/ $[-\text{ta}]\text{-B[a]P-N}^2\text{-dG}$  is nearly as bad. The source of these defects can be understood by viewing the close fitting structure for  $[-\text{ta}]\text{-B[a]P-N}^2\text{-dG}$  in AS1 (Fig. 7A), where the pyrene moiety (turquoise) sits in the slot between the fingers domain and the little finger domain (Fig. 5A), notably contacting I104 (red), S103 (purple) and Y12 (brown). To move from AS1 (Fig. 7A) to AS2 (Fig. 7B), template  $[-\text{ta}]\text{-B[a]P-N}^2\text{-dG}$  must move helically down and to the left (arrow), which forces the pyrene moiety into another crowded region of the protein and results in distortions in the AS2 active site (Table 2).

Thus, the modeling work suggests that dCTP insertion opposite  $\text{B[a]P-N}^2\text{-dG}$  in the BPmi5 conformation might occur more readily in AS2 than in AS1 — if structure alone dictated insertion. In contrast, dCTP insertion opposite  $\text{B[a]P-N}^2\text{-dG}$  in the BPmi3 conformation might occur more readily in AS1 than AS2. There are numerous caveats to these notions, notably the

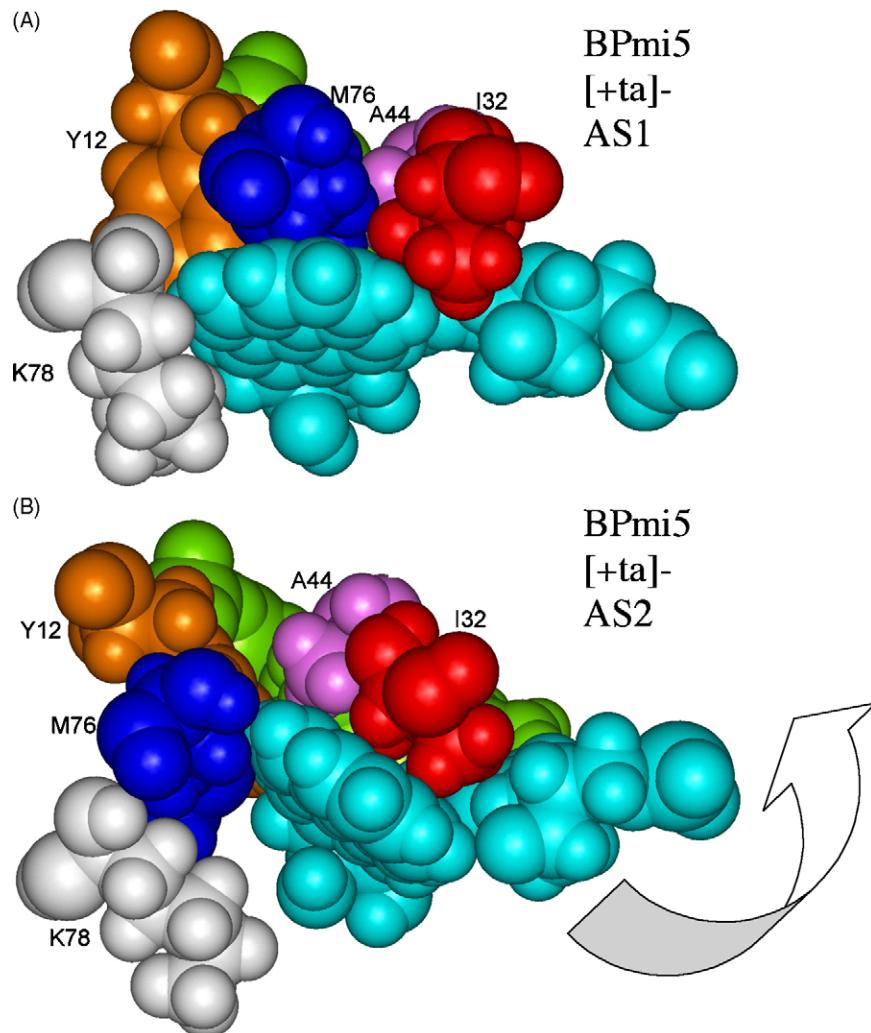


Fig. 6. The BPmi5 conformation of  $\text{B[a]P-N}^2\text{-dG}$  is better accommodated in AS2 than AS1 of Dpo4. In AS2 (panel B), the pyrene moiety of  $[+\text{ta}]\text{-B[a]P-N}^2\text{-dG}$  (turquoise) sits in a notch in the fingers domain (Fig. 5B), which is defined by amino acids: K78 (gray), Y12 (brown), M76 (blue), A44 (purple) and V32 (red). To move from AS2 (panel B) to AS1 (panel A), template  $[+\text{ta}]\text{-B[a]P-N}^2\text{-dG}$  must move helically up and to the right (arrow), which forces the pyrene moiety into a crowded region of the protein and results in severe distortions in the AS1 active site (Table 2).

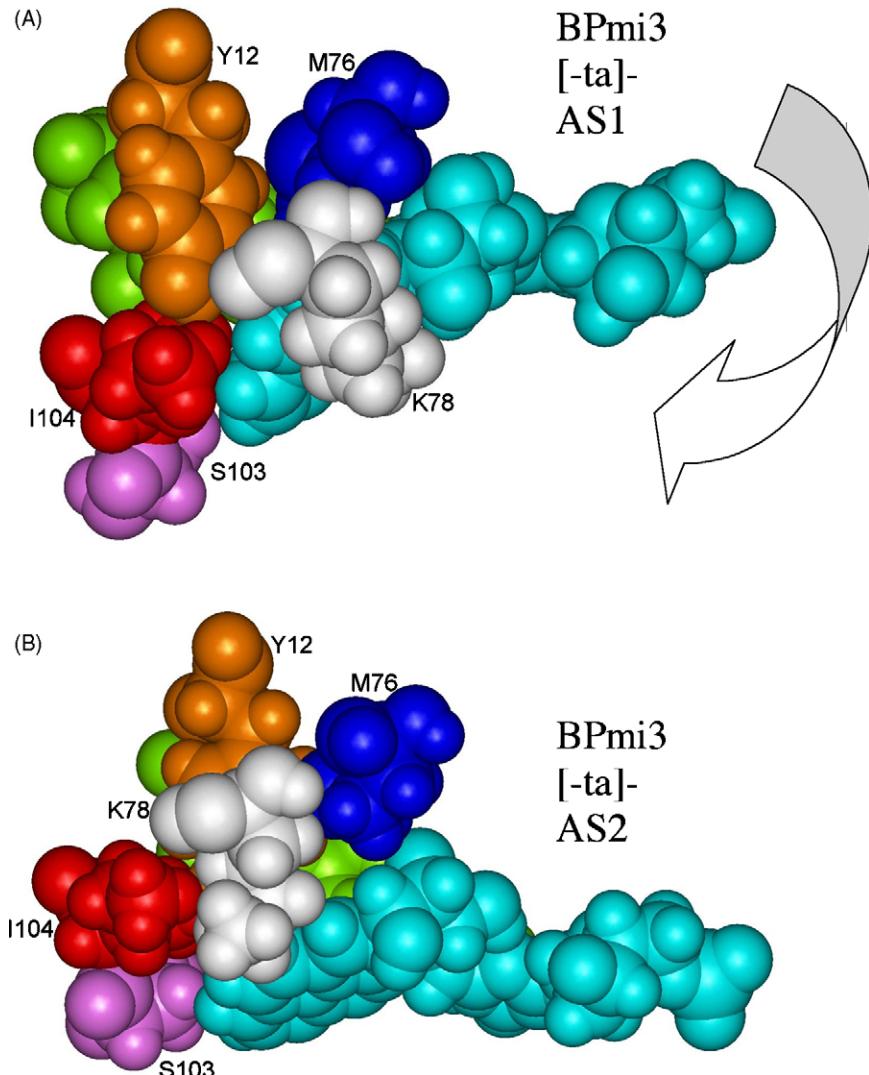


Fig. 7. The BPmi3 conformation of B[a]P-N<sup>2</sup>-dG is better accommodated in AS1 than AS2 of Dpo4. In AS1 (panel A), the pyrene moiety of [-ta]-B[a]P-N<sup>2</sup>-dG (turquoise) sits in the slot between the fingers domain and the little finger domain (Fig. 5A), most notably contacting I104 (red) and Y12 (brown). To move from AS1 (panel A) to AS2 (panel B), template [-ta]-B[a]P-N<sup>2</sup>-dG must move helically down and to the left (arrow), which forces the pyrene moiety into a crowded region of the protein and results in severe distortions in the AS2 active site (Table 2).

fact that Dpo4 was used, and it may or may not be revealing about the relevant DNAPs.

#### 4.3. AS2, the “A-rule”, and the potential usefulness of two active sites for Y-family DNAPs

The major mutagenic pathway for B[a]P-N<sup>2</sup>-dG adducts is most frequently G → T [75], which must result from dATP insertion, and we next reflect on this pathway. The preferential insertion of dATP opposite a number of DNA lesions in *E. coli* has been called the “A-rule” ([76,77] and references therein), which also rationalizes UV light mutagenesis, as C → T mutations predominate in 5'-PyC sequences, implying dATP insertion. UV mutagenesis depends on *umuD/C*, which encode DNAP V, implying that DNAP V is required for UV mutagenic dATP insertion (discussed in Ref. [25]). After reviewing the literature, we concluded that *E. coli* DNAP V appears to have

two modes of insertion opposite lesions: (i) correct dNTP insertion, and (ii) default dATP insertion, where the latter seems likely to be the basis for *E. coli*’s A-rule [66]. We speculate that UV light is a frequently encountered form of DNA damage for which a lesion-bypass DNAP might be important, and since thymine dimers (TT-CPDs) are the major UV lesion [78], a default dA insertion mode might help minimize UV mutagenesis. Based on lesion-bypass specificity and protein structure, we have argued that *E. coli* DNAP V is the ortholog of human DNAP  $\eta$  [66], which is almost certainly responsible for correct bypass of UV lesions in human cells [43–48].

Three Dpo4 X-ray structures have dATP paired with a thymine dimer (TT-CPD) in the template strand, and these structures may help illuminate why Y-family DNAPs might have two active sites. In 1RYR and 1RYS-A, dATP is positioned opposite the 5'- and 3'-T, respectively, of a TT-CPD, both in AS2 (Table 1). For UV light mutagenesis, dATP would

pair with the 3'-C in a 5'-PyC sequence. Interestingly, in 1RYS-A the adenine of dATP is in the *syn*-conformation, and perhaps *syn*-dATP is important for the “default dATP insertion” mode (previous paragraph), which might rationalize how dATP is inserted opposite the 3'-C in a 5'-PyC sequence. In this regard, it is interesting to note that the shape of the deoxyribose-phosphates in a canonical AS1-like binding site sterically prevents the adenine from adopting the *syn*-configuration in dATP (Fig. 8A). (In fact, one could imagine that the canonical binding mode for DNAPs (e.g., AS1 in Dpo4) might have evolved in part to minimize the *syn*-orientation of bases in dNTPs.) Interestingly, however, *syn*-dATP is less hindered in AS2 (Fig. 8B), as the deoxyribose-phosphates are oriented differently. We note that this notion seems to be contradicted by the fact that *syn*-dATP is also observed opposite the 3'-T of a TT-CPD in an AS1 site in 1RYS-B; however, the AS1 site in 1RYS-B is distorted in comparison to others in the AS1 collection (Fig. 3A, yellow), including the fact that the M2-O1a ligand bond is very long (4.73 Å), and a M1 divalent cation is not present in the structure. Thus, while 1RYS-B has *syn*-dATP in AS1, this active site is so distorted that it almost certainly cannot be catalytically active.

*E. coli* DNAP V and human DNAP  $\eta$  are likely to be functional orthologs, and both have isoleucine for their “roof-amino acid”, which is the residue that lies directly above the base of the dNTP in AS1, and both also have a bulky amino acid contacting the roof-aa on the major groove side (MaG1-aa, which are methionine and arginine, respectively) [66]. *E. coli* DNAP IV and human DNAP  $\kappa$  are also functional orthologs, and they have non-bulky serine as their roof-aa, as well as a

non-bulky residue for MaG1-aa (i.e., serine and alanine, respectively). Bulky amino acids (e.g., with DNAPs V and  $\eta$ ) in the roof- and MaG1-aa positions should favor dNTP binding in the lower active site AS2, while less bulky amino acids (e.g., with DNAPs IV and  $\kappa$ ) might favor dNTP binding in the upper active site AS1.

As noted in Section 1, lesion-bypass DNAPs can be involved in either the insertion or extension reaction. It is intriguing to speculate that in some circumstances, AS2 might do insertion opposite an adduct/lesion, and thereafter AS1 might do the subsequent extension step opposite the normal base on the 5'-side of the adduct/lesion. These and other ideas need further investigation.

In summary, we have noted that Y-family DNAPs may have two overlapping active sites AS1 and AS2. We have found that B[a]P-N<sup>2</sup>-dG adducts in the BPmi3 orientation (Fig. 5A) fit better in AS1, while B[a]P-N<sup>2</sup>-dG adducts in the BPmi5 orientation (Fig. 5B) fit better in AS2. Furthermore, it appears that AS2 can better accommodate *syn*-dATP, and this may be the source of default dATP insertion (thought to occur with some Y-Family DNAPs [66]) with relatively distorting adducts, which may be forced to use AS2 for TLS. In contrast, AS1 might be used for the replication of less distorting lesions and/or to synthesize undamaged DNA, as required in the extension step of TLS.

## Acknowledgement

This work was supported by United States Public Health Services Grant R01ES03775.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmgm.2006.05.003.

## References

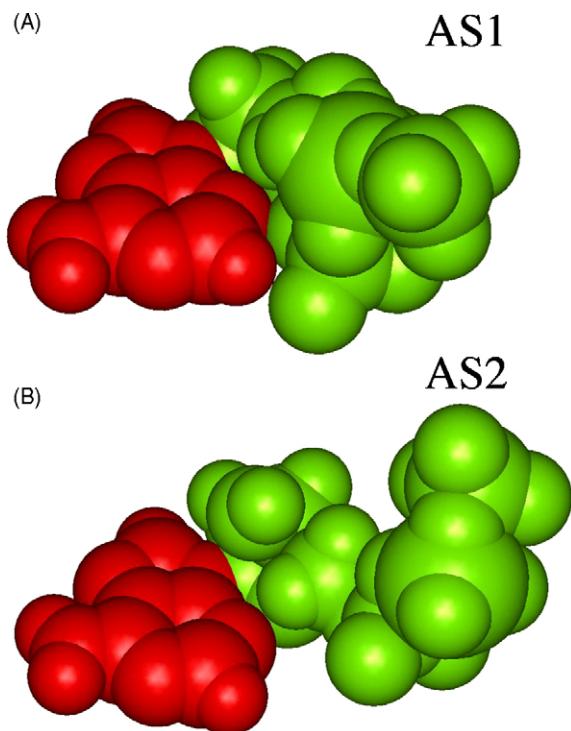


Fig. 8. Adenine can be accommodated better in the *syn*-orientation when dATP is in the shape defined by the AS2 active site (panel B) than by the AS1 active site (panel A).

- [1] R.G. Harvey, Polycyclic Aromatic Hydrocarbons: Chemistry and Cancer, Wiley–VCH Inc., New York, 1997.
- [2] D.H. Phillips, Fifty years of benzo[a]pyrene, *Nature* 303 (1983) 468–472.
- [3] B. Singer, D. Grunberger, Molecular Biology of Mutagens and Carcinogens, Plenum Press, New York, 1983.
- [4] A.H. Conney, Induction of microsomal enzymes by foreign chemicals and carcinogens by polycyclic aromatic hydrocarbons, *Cancer Res.* 42 (1982) 4875–4917.
- [5] A. Dipple, Polycyclic aromatic hydrocarbon carcinogens, in: R.G. Harvey (Ed.), Polycyclic Aromatic Hydrocarbons and Carcinogenesis, American Chemical Society Press, Washington, DC, 1985, pp. 1–17.
- [6] P.W. Jones, Polynuclear aromatic hydrocarbons, in: M.G. Bowman (Ed.), Handbook of Carcinogens and Hazardous Substances, Marcel Dekker Inc., New York, 1982, pp. 573–639.
- [7] P. Grasso, Carcinogens in food, in: G.C. Searle (Ed.), Chemical Carcinogens, 2nd ed., ACS Monograph 182, American Chemical Society Press, Washington, DC, 1984, pp. 1205–1239.
- [8] D. Chakravarti, J. Pelling, E.L. Cavalieri, E.G. Rogan, Relating aromatic hydrocarbon-induced DNA adducts and the c-H-ras mutations in mouse skin papillomas: the role of apurinic sites, *Proc. Natl. Acad. Sci. U.S.A.* 92 (1995) 10422–10426.

- [9] L. Chen, R. Devanesan, S. Higginbotham, F. Ariese, R. Jankowiak, G.J. Small, E.G. Rogan, E. Cavalieri, Expanded analysis of benzo[a]pyrene-DNA adducts formed *in vitro* and in mouse skin: their significance in tumor initiation, *Chem. Res. Toxicol.* 9 (1996) 897–903.
- [10] A. Balmain, K. Brown, Oncogene activation in chemical carcinogenesis, *Adv. Cancer Res.* 51 (1988) 147–182.
- [11] B.A. Ruggeri, B. Bauer, S.-Y. Zhang, A.J.P. Klein-Szanto, Murine squamous cell carcinoma cell lines produced by a complete carcinogenesis protocol with benzo[a]pyrene exhibit characteristic p53 mutations and the absence of H-ras and cyclin D1 abnormalities, *Carcinogenesis* 15 (1994) 1613–1619.
- [12] M. Hall, P.L. Grover, Chemical carcinogenesis and mutagenesis, in: G.S. Cooper, P.L. Grover (Eds.), *Handbook of Experimental Pharmacology*, vol. 94/1, Springer-Verlag, Heidelberg, 1990, pp. 327–372.
- [13] G.P. Pfeifer, P. Hainaut, On the origin of G-to-T transversions in lung cancer, *Mutat. Res.* 526 (2003) 39–43.
- [14] H. Rodriguez, E.L. Loehler, Mutational spectra of the (+)-anti-diol epoxide of benzo[a]pyrene in a *supF* gene of an *Escherichia coli* plasmid: DNA sequence context influences hotspots, mutational specificity and the extent of SOS enhancement of mutagenesis, *Carcinogenesis* 14 (1993) 373–383.
- [15] H. Rodriguez, E.L. Loehler, Mutagenesis by the (+)-anti-diol epoxide of benzo[a]pyrene: what controls mutagenic specificity? *Biochemistry* 32 (1993) 373–383.
- [16] S.J. Wei, R.L. Chang, C.Q. Wong, X.X. Cui, N. Dandamudi, Y.P. Lu, K.A. Merkler, J.M. Sayer, A.H. Conney, D.M. Jerina, The ratio of deoxyadenosine to deoxyguanosine adducts formed by (+)-(7R,8S,9S,10R)-7,8-dihydroxy-9,10-epoxy-7,8,9,10-tetrahydrobenzo[a]pyrene in purified calf thymus DNA and DNA in V-79 cells is independent of dose, *Int. J. Oncol.* 14 (1999) 509–513.
- [17] M. Schiltz, X.X. Cui, Y.P. Lu, H. Yagi, D.M. Jerina, M.Z. Zdzienicka, R.L. Chang, A.H. Conney, S.J. Wei, Characterization of the mutational profile of (+)-7R,8S-dihydroxy-9S, 10R-epoxy-7,8,9,10-tetrahydrobenzo[a]pyrene at the hypoxanthine (guanine) phosphoribosyltransferase gene in repair-deficient Chinese hamster V-H1 cells, *Carcinogenesis* 20 (1999) 2279–2286.
- [18] K.-Y. Seo, A. Nagalingam, Shadi Miri, Jun Yin, Alexander Kolbanovskiy, Anant Shastry, E.L. Loehler, Mirror image stereoisomers of the major benzo[a]pyrene N<sup>2</sup>-dG adduct are bypassed by different lesion-bypass DNA polymerases in *E. coli* DNA repair 5 (2006) 515–527.
- [19] R. Napolitano, R. Janel-Bintz, J. Wagner, R.P. Fuchs, All three SOS-inducible DNA polymerases (Pol II, Pol IV and Pol V) are involved in induced mutagenesis, *EMBO J.* 19 (2000) 6259–6265.
- [20] N. Lenne-Samuel, R. Janel-Bintz, A. Kolbanovskiy, N.E. Geacintov, R.P. Fuchs, The processing of a benzo(a)pyrene adduct into a frameshift or a base substitution mutation requires a different set of genes in *Escherichia coli*, *Mol. Microbiol.* 38 (2000) 299–307.
- [21] J. Wagner, H. Etienne, R. Janel-Bintz, R.P. Fuchs, Genetics of mutagenesis in *E. coli*: various combinations of translesion polymerases (Pol II, IV and V) deal with lesion/sequence context diversity, *DNA Repair* 1 (2002) 159–167.
- [22] D.F. Jarosz, V.G. Godoy, J.C. Delaney, J.M. Essigmann, G.C. Walker, A single amino acid governs enhanced activity of DinB DNA polymerases on damaged templates, *Nature* 439 (2006) 225–228.
- [23] J. Yin, K.-Y. Seo, E.L. Loehler, A role for DNA polymerase V in G → T mutagenesis from the major benzo[a]pyrene N<sup>2</sup>-dG adduct when studied in a 5'-TGT sequence in *Escherichia coli*, *DNA Repair* 3 (2004) 323–334.
- [24] K. Bebenek, T.A. Kunkel, Functions of DNA polymerases, *Adv. Protein Chem.* 69 (2004) 137–165.
- [25] M.F. Goodman, Error-prone repair DNA polymerases in prokaryotes and eukaryotes, *Annu. Rev. Biochem.* 71 (2002) 17–50.
- [26] U. Hubscher, G. Maga, S. Spadari, Eukaryotic DNA polymerases, *Annu. Rev. Biochem.* 71 (2002) 133–163.
- [27] A.J. Rattray, J.N. Strathern, Error-prone DNA polymerases: when making a mistake is the only way to get ahead, *Annu. Rev. Genet.* 37 (2003) 31–66.
- [28] H. Ohmori, E.C. Friedberg, R.P. Fuchs, M.F. Goodman, F. Hanaoka, D. Hinkle, T.A. Kunkel, C.W. Lawrence, Z. Livneh, T. Nohmi, L. Prakash, S. Prakash, T. Todo, G.C. Walker, Z. Wang, R. Woodgate, The Y-family of DNA polymerases, *Mol. Cell* 8 (2001) 7–8.
- [29] F. Boudsocq, H. Ling, W. Yang, R. Woodgate, Structure-based interpretation of missense mutations in Y-family DNA polymerases and their implications for polymerase function and lesion bypass, *DNA Repair (Amst.)* 1 (2002) 343–358.
- [30] E.C. Friedberg, R. Wagner, M. Radman, Specialized DNA polymerases, cellular survival, and the genesis of mutations, *Science* 296 (2002) 1627–1630.
- [31] E.C. Friedberg, Why do cells have multiple error-prone DNA polymerases? *Environ. Mol. Mutagen.* 38 (2001) 105–110.
- [32] Z. Wang, Translesion synthesis by the UmuC family of DNA polymerases, *Mutat. Res.* 486 (2001) 59–70.
- [33] R.E. Johnson, M.T. Washington, S. Prakash, L. Prakash, Bridging the gap: a family of novel DNA polymerases that replicate faulty DNA, *Proc. Natl. Acad. Sci. U.S.A.* 96 (1999) 12224–12226.
- [34] V. Pages, R.P. Fuchs, How DNA lesions are turned into mutations within cells? *Oncogene* 21 (2002) 8957–8966.
- [35] W. Yang, Damage repair DNA polymerases Y, *Curr. Opin. Struct. Biol.* 13 (2003) 23–30.
- [36] P.M. Burgers, E.V. Koonin, E. Bruford, L. Blanco, K.C. Burtis, M.F. Christman, W.C. Copeland, E.C. Friedberg, F. Hanaoka, D.C. Hinkle, C.W. Lawrence, M. Nakanishi, H. Ohmori, L. Prakash, S. Prakash, C.A. Reynaud, A. Sugino, T. Todo, Z. Wang, J.C. Weill, R. Woodgate, Eukaryotic DNA polymerases: proposal for a revised nomenclature, *J. Biol. Chem.* 276 (2001) 43487–43490.
- [37] H. Ohmori, E.C. Friedberg, R.P. Fuchs, M.F. Goodman, F. Hanaoka, D. Hinkle, T.A. Kunkel, C.W. Lawrence, Z. Livneh, T. Nohmi, L. Prakash, S. Prakash, T. Todo, G.C. Walker, Z. Wang, R. Woodgate, The Y-family of DNA polymerases, *Mol. Cell* 8 (2001) 7–8.
- [38] B.L. Zhou, J.D. Pata, T.A. Steitz, Crystal structure of a DinB lesion bypass DNA polymerase catalytic fragment reveals a classic polymerase catalytic domain, *Mol. Cell* 8 (2001) 427–437.
- [39] H. Ling, F. Boudsocq, R. Woodgate, W. Yang, Crystal structure of a Y-family DNA polymerase in action: a mechanism for error-prone and lesion-bypass replication, *Cell* 107 (2001) 91–102.
- [40] H. Ling, F. Boudsocq, B.S. Plosky, R. Woodgate, W. Yang, Replication of a *cis-syn* thymine dimer at atomic resolution, *Nature* 424 (2003) 1083–1087.
- [41] H. Ling, J.M. Sayer, B.S. Plosky, H. Yagi, F. Boudsocq, R. Woodgate, D.M. Jerina, W. Yang, 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 (2004) 2265–2269.
- [42] H. Ling, F. Boudsocq, R. Woodgate, W. Yang, Snapshots of replication through an abasic lesion; structural basis for base substitutions and frameshifts, *Mol. Cell* 13 (2004) 751–762.
- [43] J. Trincao, R.E. Johnson, W.T. Wolfe, C.R. Escalante, S. Prakash, L. Prakash, A.K. Aggarwal, Dpo4 is hindered in extending a G-T mismatch by a reverse wobble, *Nat. Struct. Mol. Biol.* 11 (2004) 457–462.
- [44] A. Vaisman, H. Ling, R. Woodgate, W. Yang, Fidelity of Dpo4: effect of metal ions, nucleotide selection and pyrophosphorylation, *EMBO J.* 25 (2005) 2957–2967.
- [45] H. Zang, A.K. Goodenough, J.Y. Choi, A. Irimia, L.V. Loukachevitch, I.D. Kozekov, K.C. Angel, C.J. Rizzo, M. Egli, F.P. Guengerich, 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,N<sub>2</sub>-ethenoguanine, *J. Biol. Chem.* 280 (2005) 29750–29764.
- [46] H. Zang, A. Irimia, J.-Y. Choi, K.C. Angel, L.V. Loukachevitch, M. Egli, F.P. Guengerich, Efficient and high fidelity incorporation of dCTP opposite 7,8-dihydro-8-oxodeoxyguanosine by *Sulfolobus solfataricus* DNA polymerase Dpo4, *J. Biol. Chem.* 281 (2006) 2358–2372.
- [47] O. Rechkoblit, L. Malinina, Y. Cheng, V. Kuryavyi, S. Broyde, N.E. Geacintov, D.J. Patel, Stepwise translocation of Dpo4 polymerase during error-free bypass of an oxoG lesion, *Plos. Biologia* 4 (2006) 1–18.
- [48] J. Trincao, R.E. Johnson, C.R. Escalante, S. Prakash, L. Prakash, A.K. Aggarwal, Structure of the catalytic core of *S. cerevisiae* DNA polymerase eta: implications for translesion DNA synthesis, *Mol. Cell* 8 (2001) 417–426.

- [49] D.T. Nair, R.E. Johnson, S. Prakash, L. Prakash, A.K. Aggarwal, Replication by human DNA polymerase-*iota* occurs by Hoogsteen base-pairing, *Nature* 430 (2004) 377–380.
- [50] S.N. Uljon, R.E. Johnson, T.A. Edwards, S. Prakash, L. Prakash, A.K. Aggarwal, Crystal structure of the catalytic core of human DNA polymerase kappa, *Structure (Camb.)* 12 (2004) 1395–1404.
- [51] C. Masutani, R. Kusumoto, A. Yamada, N. Dohmae, M. Yokoi, M. Yuasa, M. Araki, S. Iwai, K. Takio, F. Hanaoka, The XPV (xeroderma pigmentosum variant) gene encodes human DNA polymerase eta, *Nature* 399 (1999) 700–704.
- [52] R.E. Johnson, S. Prakash, L. Prakash, Efficient bypass of a thymine-thymine dimer by yeast DNA polymerase, Pol eta, *Science* 283 (1999) 1001–1004.
- [53] R.E. Johnson, C.M. Kondratick, S. Prakash, L. Prakash, hRAD30 mutations in the variant form of xeroderma pigmentosum, *Science* 285 (1999) 263–265.
- [54] M.T. Washington, R.E. Johnson, S. Prakash, L. Prakash, Fidelity and processivity of *Saccharomyces cerevisiae* DNA polymerase eta, *J. Biol. Chem.* 274 (1999) 36835–36838.
- [55] M.T. Washington, R.E. Johnson, S. Prakash, L. Prakash, Accuracy of thymine-thymine dimer bypass by *Saccharomyces cerevisiae* DNA polymerase eta, *Proc. Natl. Acad. Sci. U.S.A.* 97 (2000) 3094–3099.
- [56] R.E. Johnson, M.T. Washington, S. Prakash, L. Prakash, Fidelity of human DNA polymerase eta, *J. Biol. Chem.* 275 (2000) 7447–7450.
- [57] O. Rechkoblit, Y. Zhang, D. Guo, Z. Wang, S. Amin, J. Krzeminsky, N. Louneva, N.E. Geacintov, *trans*-Lesion synthesis past bulky benzo[a]pyrene diol epoxide and N<sup>2</sup>-dG and N<sup>6</sup>-dA lesions catalyzed by DNA bypass polymerases, *J. Biol. Chem.* 277 (2002) 30488–30494.
- [58] D. Chiappерino, H. Kroth, I.H. Kramarczuk, J.M. Sayer, C. Masutani, F. Hanaoka, D.M. Jerina, A.M. Cheh, Preferential misincorporation of purine nucleotides by human DNA polymerase eta opposite benzo[a]pyrene 7,8-diol, 9,10-epoxide deoxyguanosine adducts, *J. Biol. Chem.* 277 (2002) 11765–11771.
- [59] Y. Zhang, X. Wu, D. Guo, O. Rechkoblit, N.E. Geacintov, Z. Wang, Two-step error-prone bypass of the (+)- and (−)-trans-anti-BPDE-N(2)-dG adducts by human DNA polymerases eta and kappa, *Mutat. Res.* 510 (2002) 23–35.
- [60] Y. Zhang, F. Yuan, X. Wu, O. Rechkoblit, J.S. Taylor, N.E. Geacintov, Z. Wang, Error-prone lesion bypass by human DNA polymerase eta, *Nucleic Acids Res.* 28 (2000) 4717–4724.
- [61] N. Suzuki, E. Ohashi, A. Kolbanovskiy, N.E. Geacintov, A.P. Grollman, H. Ohmori, S. Shibusaki, Translesion synthesis by human DNA polymerase kappa on a DNA template containing a single stereoisomer of dG-(+)- or dG-(−)-anti-N(2)-BPDE (7,8-dihydroxy-anti-9,10-epoxy-7,8,9,10-tetrahydrobenzo[a]pyrene), *Biochemistry* 41 (2002) 6100–6106.
- [62] X. Huang, A. Kolbanovskiy, X. Wu, Y. Zhang, Z. Wang, P. Zhuang, S. Amin, N.E. Geacintov, Effects of base sequence context on translesion synthesis past a bulky (+)-trans-anti-B[a]P-N2-dG lesion catalyzed by the Y-family polymerase pol kappa, *Biochemistry* 42 (2003) 2456–2466.
- [63] Y. Zhang, X. Wu, D. Guo, O. Rechkoblit, Z. Wang, Activities of human DNA polymerase kappa in response to the major benzo[a]pyrene DNA adduct: error-free lesion bypass and extension synthesis from opposite the lesion, *DNA Repair (Amst.)* 1 (2002) 559–569.
- [64] T. Ogi, Y. Shinkai, K. Tanaka, H. Ohmori, Pol kappa protects mammalian cells against the lethal and mutagenic effects of benzo[a]pyrene, *Proc. Natl. Acad. Sci. U.S.A.* 99 (2002) 15548–15553.
- [65] S. Avkin, M. Goldsmith, S. Velasco-Miguel, N. Geacintov, E.C. Friedberg, Z. Livneh, Quantitative analysis of translesion DNA synthesis across a benzo[a]pyrene-guanine adduct in mammalian cells: the role of DNA polymerase kappa, *J. Biol. Chem.* 279 (2004) 53298–53305.
- [66] C.H. Lee, S. Chandani, E.L. Loechler, Homology modeling of four lesion-bypass DNA polymerases: structure and lesion bypass findings suggest that *E. coli* pol IV and human Pol kappa are orthologs, and *E. coli* pol V and human Pol eta are orthologs, *J. Mol. Graph. Model.*, in press.
- [67] M. Cosman, C. de los Santos, R. Fiala, B.E. Hingerty, V. Ibanez, L.A. Margulis, D. Live, N.E. Geacintov, S. Broyde, D.J. Patel, Solution conformation of the major adduct between the carcinogen (+)-anti-benzo[a]pyrene diol epoxide and DNA, *Proc. Natl. Acad. Sci. U.S.A.* 89 (1992) 1914–1918.
- [68] M.A. Fountain, T.R. Krugh, Structural characterization of a (+)-trans-anti-benzo[a]pyrene-DNA adduct using NMR, restrained energy minimization and molecular dynamics, *Biochemistry* 34 (1995) 3152–3161.
- [69] C. de los Santos, M. Cosman, B. Hingerty, V. Ibanez, L.A. Margulis, N.E. Geacintov, S. Broyde, D.J. Patel, Influence of benzo[a]pyrene diol epoxide chirality on solution conformations of DNA covalent adducts: the (−)-trans-anti-[BP]G-C adduct structure and comparison with the (+)-trans-anti-[BP]G-C enantiomer, *Biochemistry* 31 (1992) 5245–5252.
- [70] Insight II, version 98.0, Accelrys Inc., San Diego, 1998.
- [71] L. Nilsson, M. Karplus, Empirical energy functions for energy minimization and dynamics of nucleic acids, *J. Comp. Chem.* 7 (1986) 591–616.
- [72] A.D. MacKerell Jr., N. Banavali, N. Foloppe, Development and current status of the CHARMM force field for nucleic acids, *Biopolymers* 56 (2001) 257–265.
- [73] S. Chandani, C.H. Lee, E.L. Loechler, Free energy perturbation methods to study structure and energetics of DNA adducts: results for the major N<sup>2</sup>-dG adduct of benzo[a]pyrene in two conformations and different sequence contexts, *Chem. Res. Toxicol.* 18 (2005) 1108–1123.
- [74] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne, The protein data bank, *Nucleic Acids Res.* 28 (2000) 235–242.
- [75] K.-Y. Seo, A. Nagalingam, E.L. Loechler, Mutagenesis studies on four stereoisomeric N<sup>2</sup>-dG benzo[a]pyrene adducts in the identical 5'-CGC sequence used in NMR studies: although adduct conformation differs, mutagenesis outcome does not as G → T mutations dominate in each case, *Mutagenesis* 20 (2005) 441–448.
- [76] B.S. Strauss, The “A” rule revisited: polymerases as determinants of mutational specificity, *DNA Repair* 1 (2002) 125–135.
- [77] E.L. Loechler, Mechanism by which aflatoxins and other bulky carcinogens induce mutations, in: Eaton, Groopman (Eds.), *The Toxicology of Aflatoxins: Human Health, Veterinary, and Agricultural Significance*, Chapter 8, Academic Press, Orlando, 1994, pp. 149–178.
- [78] D.E. Brash, W.A. Haseltine, UV-induced mutation hotspots occur at DNA damage hotspots, *Nature* 298 (1982) 189–192.