**Y-Family DNA Polymerases Have Three Structural Elements
That Promote Accurate dCTP Insertion
And Minimize dATP/dGTP/dTTP Misinsertion
Opposite a N²-dG Adduct of Benzo[*a*]pyrene**

Gabriel Sholder¹

Peter Tonzi²

Sushil Chandani³

Edward L. Loechler^{1*}

¹Department of Biology, Boston University, Boston, MA 02215

²New York University School of Medicine, New York, NY 10016

³Novarus Discoveries Pvt. Ltd., Hyderabad 500073, India.

Sushil Chandani New Biological Entities, Secunderabad 500026 India Email:
sushilchandani@gmail.com

This work was supported by a grant from the NIH (R01ES03775 to ELL).

ABBREVIATIONS

²BP, benzo[*a*]pyrene; +BP-dG, [+ta]-BP-N²-dG (Figure 1); -BP-dG, [-ta]-BP-N²-dG (Figure 1); AP site, apurinic/aprimidinic site; thymine dimer, a cyclobutanepyrimidine dimer formed between adjacent thymimes by UV light; DNAP, DNA polymerase; CC-Domain, catalytic core domain that includes the thumb/palm/fingers domains, which operates as a unit and contains the active site; LF-Domain, the little finger domain, which helps grip DNA; SsDpo4 *Sulfolobus Solfataricus* Dpo4; SaDbh, *Sulfolobus acidocaldarius* Dbh; EcDNAP IV, *Escherichia coli* DNA polymerase IV; AS-NCB, active site non-covalent bridge; Distal-NCB, Distal non-covalent bridge.

ABSTRACT:

To bypass DNA damage, cells have Y-Family DNA polymerases (DNAPs). One Y-Family-class includes DNAP κ and DNAP IV, which accurately insert dCTP opposite N²-dG adducts, including from the carcinogen benzo[*a*]pyrene (BP). Another class includes DNAP η and DNAP V, which insert accurately opposite UV-damage, but inaccurately opposite BP-N²-dG. To investigate structural differences between Y-Family-classes, Dpo4 (a canonical η /V-class-member) is modified to make it more κ /IV-like, as evaluated via primer-extension studies with a BP-N²-dG-containing template. Three protein structural elements are identified that promote fidelity. (1) Watson-Crick-like [dCTP:BP-N²-dG] pairing requires the BP-moiety to be in the minor groove. Thus, as expected, dCTP insertion is facilitated by having large openings in the protein surface that can accommodate BP-bulk in the minor groove. The BP-moiety is also in the minor groove during dATP and dTTP misinsertion, though evidence suggests that each of these three minor groove BP-conformations differ. (2) Plugging an opening on the major groove side of the protein suppresses dGTP misinsertion, implying BP-N²-dG bulk is in the major groove during Hoogsteen *syn*-adduct-dG:dGTP pairing. (3) Y-Family DNAPs have non-covalent bridges (NCBs) holding their little finger-domain in contact with their catalytic core-domain; dATP/dGTP/dTTP misinsertions are suppressed by the quantity and quality of NCBs, including one near and another distal to the active site on the minor groove side. In conclusion, three protein structural elements enhance dCTP and/or suppress dATP/dGTP/dTTP insertion; four different BP-adduct conformations are responsible for the four different dNTP insertional pathways opposite BP-N²-dG; generalizations about Y-Family structure are also considered.

INTRODUCTION

Cells possess many DNA polymerases (DNAPs); e.g., humans, yeast (*S. cerevisiae*) and *E. coli* have at least seventeen, eight and five, respectively.⁽¹⁻³⁾ These DNAPs serve many functions; e.g., DNA damage caused by chemicals and radiation often block replicative DNAPs and to avoid such lethal blockage, cells possess lesion-bypass DNAPs, which conduct translesion DNA synthesis (representative recent references include: ^(1, 2, 4-9) with classic citations in reference 9). Most lesion-bypass DNAPs are in the Y-Family, which was the subject of a fine review.⁽⁷⁾ Human cells have three that are template-directed (hDNAPs η , ι and κ), while fission yeast has one (ScDNAPs η) and *E. coli* has two (EcDNAPs IV and V).

X-ray structures reveal that Y-Family DNAPs share a conserved catalytic region of ~350aa (representative references:⁽¹⁰⁻²⁸⁾). All DNA polymerases resemble a right-hand with thumb, palm and fingers domains, which in Y-Family DNAPs make up the catalytic core (CC-Domain).⁽⁷⁾ Y-Family DNAPs grip DNA with another domain,⁽¹⁰⁻¹²⁾ often called the “little finger” (LF-Domain). Y-Family DNAPs have “stubby” fingers, leaving more solvent accessible surface around the template/dNTP-binding pocket,⁽⁴⁾ which can accommodate adduct/lesion bulk. A representative X-ray structure of Dpo4 from *Sulfolobus solfataricus* (the best-studied Y-Family DNAP) is shown in Figure 1A, as viewed from the minor groove with its CC-Domain in green and LF-Domain in yellow. Representative structures are also shown for SaDbh (Figure 1B), hDNAP κ (Figure 1C) and EcDNAP IV (Figure 1D).

Steps in the mechanism of Y-Family DNAPs have been proposed for both protein structural changes based on a series of X-ray structures^(29, 30) and for chemical catalysis based on theoretical studies.⁽³¹⁾ Steps in Y-Family DNAP bypass of several lesions have been identified

based on kinetic analysis,^(32, 33) hydrogen-deuterium exchange mass spectrometry,^(34, 35) and single molecule FRET.⁽³⁶⁾

Our work has focused on benzo[*a*]pyrene (BP),⁽³⁷⁻⁴¹⁾ which is a well-studied DNA damaging agent that is a potent mutagen and carcinogen, and an example of a polycyclic aromatic hydrocarbon (PAH), a class of ubiquitous environmental substances produced by incomplete combustion.⁽⁴²⁾ PAHs in general and BP in particular induce the kinds of mutations that may be important in human cancer⁽⁴³⁾ and references therein). Figure 2 shows the major BP DNA adduct (+BP-dG) and its mirror image (–BP-dG), both of which form at N²-dG.

Based on amino acid sequence, Y-Family DNAPs are divided into six major groups.⁽⁷⁾ Based on insertional patterns opposite many lesions, classes seem to exist.⁽⁴⁴⁾ The κ /IV-class correctly inserts dCTP opposite N²-dG adducts, including from +BP-dG and –BP-dG,^(28, 37, 45-56) as well as from N²-dG adducts from endogenous sources.^(57, 58) IV/ κ -class-members also correctly insert opposite other lesions.⁽⁵⁹⁻⁶²⁾ κ /IV-class-members do not readily insert opposite UV-damage and tend to be relatively accurate in the replication of undamaged DNA.⁽⁶³⁾ κ /IV-class-members include hDNAP κ , EcDNAP IV and SaDbh. η /IV-class-members are relatively inaccurate in replicating undamaged DNA and insertion opposite N²-dG adducts, but are accurate with UV-damage, with thymine dimers being the best studies example.⁽⁴⁴⁾ η /V-class-members include hDNAP η , UmuC (polymerase subunit of EcDNAP V) and SsDpo4.

Insertional preference differences for κ /IV-class-members versus η /V-class-members must reflect an interplay between protein structural differences and adduct structural/conformational differences. Five X-ray structures exist for several BP-adducts in Dpo4,^(23, 24) and each structure is different. The adduct-dG can be either in the *anti*-orientation or *syn*-orientation, which places the BP-pyrene moiety in the minor or major groove, respectively. The BP-pyrene moiety can

also be intercalated. Furthermore, the adduct-dG can be paired with a base in the opposite strand or be extra-helical. Finally, only one structure has the adduct dG-moiety interacting with an active site dNTP, while the others have the [n+1] base pair in the active site. These Dpo4 structures would most likely relate to mutagenic pathways. X-ray structures in hDNAP κ show a likely structure for correct dCTP insertion opposite +BP-dG, which has more-or-less normal Watson-Crick pairing with the +BP-moiety in the minor groove.⁽²⁸⁾

Thus, X-ray studies reveal a diverse set of structural possibilities for BP-dG adducts, which provide tantalizing mechanistic possibilities, although this complexity makes the task of associating a particular structure/conformation with a particular dNTP insertion pathway daunting. To provide additional clues about the relationship between structure and mechanism, X-ray studies can be complemented by structure/activity studies, which can provide insights about transition state structures that control rates of reaction and, thus, mechanism.⁽⁶⁴⁾ We take a structure/activity approach herein to probe dNTP insertion mechanism opposite –BP-dG, which we used in two prior studies.^(40, 65)

Regarding misinsertion mechanism, dNTP insertion by Dpo4 has been studied opposite +BP-dG.⁽²³⁾ dATP, dGTP and dTTP misinsertion dominated in 5'-TG, 5'-CG and 5'-AG sequences respectively, which is consistent with a mutagenic insertion mechanism being dictated by the 5'-neighboring base via a slippage-type process.⁽⁴¹⁾ However, studies with +BP-dG and other BP-adducts have shown that mutagenic outcome being defined by the 5'-base is the exception rather than the rule (reviewed I, reference 41). In addition, we recently showed that—though Dpo4 does misinsert dATP opposite –BP-dG in a 5'-TG sequence context—dGTP is faster (~6.6-fold), while dTTP misinsertion, as well as correct dCTP insertion, are also significant.^(40, 65) Thus, on balance, the nature of the 5'-base does not principally define insertion.

Another key question is: How does a single adduct direct multiple insertional events? For example, (1) does Dpo4 incorporate dATP, dCTP, dGTP and dTTP,⁽⁴⁰⁾ because it is unable to interpret the –BP-dG adduct, such that it merely incorporates whatever dNTP happens to be in its active site, or (2) are there mechanistically distinct incorporation pathways?

We have been investigating such mechanistic questions, principally by making structural modifications in wt-Dpo4, which is in the η /V-class and has low fidelity during BP-N²-dG adduct bypass,^(23, 40, 65) to generate Dpo4-derivatives that behave more like κ /IV-class-members, with increased rates (i.e., “activity”) of dCTP insertion and decreased rates of dATP, dGTP and dTTP misinsertion opposite –BP-dG. In one study with wt-Dpo4, several Dpo4 mutants and wt-Dbh,⁽⁴⁰⁾ reactions were shown to follow simple Michaelis-Menten kinetics with respect to [dNTP]. Subsequently, the structural basis for mechanistic differences between Dpo4 and DNAP IV was investigated by studying the kinetics of various chimeric and mutant derivatives.⁽⁶⁵⁾

The study herein extends this approach by modifying structural elements in Dpo4 to make it more like Dbh, DNAP IV and DNAP κ . We find the following. (1) The –BP-moiety is in the major groove during dGTP misincorporation. In contrast the –BP-moiety is in the minor groove during dCTP, dATP and dTTP misincorporation, although evidence suggests that each of these adduct conformations is different. Thus, findings are most consistent with four different adduct –BP-dG conformations being responsible for each of the four reactions involving dCTP, dATP, dGTP and dTTP. (2) The size of an opening on the minor groove side of the active site influences the rate of correct dCTP insertion opposite a BP-dG-adduct. This is sensible because the BP-moiety must be in the minor groove to form a more-or-less normal adduct-dG:dCTP Watson-Crick-like base pair, as first predicted by molecular modeling,⁽⁶⁶⁾ and as observed in X-

ray structures.⁽²⁸⁾ (3) dATP, dGTP and dTTP misinsertions are suppressed by non-covalent bridges (NCBs) that hold the LF-Domain in contact with the CC-Domain, including one near the active site and a second that is distal to the active site. (4) Structural features are identified that κ /IV-class-members use to suppress misinsertions opposite adducts having bulk in the major groove. (5) Some of the Y-Family structural features described herein appear to be present in other Y-Family DNAPs. (6) The conclusions in this study are fundamentally consistent with those made in our previous studies.^(40, 65)

EXPERIMENTAL PROCEDURES

Vectors containing genes of interest have been described, along with methods to generate derivatives.⁽⁴⁰⁾ The one exception is N-clasp-Dpo4, which contains the first 100aa of human DNAP κ fused to Dpo4. The vector pBJ733⁽⁶⁷⁾ was PCR amplified with primers to give an appropriate N-clasp-containing fragment flanked by NdeI sites at each end. The fragment and wt-Dpo4-pET15B were each digested with NdeI, gel purified and ligated together. This method introduced an extra His residue between S100 of the N-clasp and M1 of Dpo4, which was removed via site-directed mutagenesis. Amino acid sequences of all derivatives are given in Table 1. Except as noted below, all other experimental procedures were described previously,⁽⁴⁰⁾ including the method to accurately determine rate constants, which is summarized in Supplementary Materials. Primer-extension assays were at pH 7.6 and 37°C. To obtain measurable rates for primer extensions, which typically proceeded 5-50%, a DNAP concentration (124 nM) in excess of primer/template (2 nM) had to be employed; thus, the kinetic findings can be viewed as representing virtual single turnover reactions. (We note that DNAP concentrations in excess of primer/template concentration has been used in every kinetic study with Y-Family DNAPs that we are aware of.) Pseudo-first order rate constants are reported (Table 2). Dpo4 was included in every experiment to insure that findings between experiments were comparable. K_m was previously shown to be 6.5-25-fold greater than the dNTP concentration (100 μ M) used herein (see Table 1 in reference⁽⁴⁰⁾); thus, rate constants are in a range of the Michaelis-Menten curve where V_{max}/K_m dominates. BP-adducts are considered stable, but do hydrolyze slowly to give the corresponding BP-tetraol along with the regeneration of guanine in DNA.⁽⁶⁸⁾ An experiment in Supplementary Materials describes how we established that adduct hydrolysis is not quantitatively significant.

RESULTS

Kinetics of dNTP Insertion. Using a template containing -BP-dG, ^{32}P -primer extension reactions with dNTPs were performed with purified wt-Dpo4 and wt-Dbh, along with modified mutant derivatives whose amino acid sequences are given in Table 1. Products were separated on DNA sequencing gels and subjected to phosphorimaging.⁽⁴⁰⁾ Figure 2 shows that Dpo4 inserts all dNTPs significantly opposite -BP-dG, as expected for η/V -class-members, while Dbh preferentially inserts dCTP, as expected for κ/IV -class-members. Figure 2 also shows representative phosphorimages for the other DNAP derivatives.

Phosphorimages were scanned, and Figure 3 shows a representative example. Relative amounts of extended-primer versus unextended-primer were determined from such scans using a method we established,⁽⁴⁰⁾ and as summarized in Supplemental Materials); in particular, scanning is necessary for accuracy, as other methods (e.g., drawing boxes around bands in phosphorimages) are highly inaccurate, as we showed.⁽⁴⁰⁾

A single [dNTP] concentration (100 μM) was used, both because many mutant proteins were studied and because all experiments were performed at least in triplicate. Because concentrations of both [dNTP] and [DNAP] (124nM) were in excess of [primer/template] concentration (2nM), reactions followed pseudo-first order kinetics, as we established experimentally.⁽⁴⁰⁾

For each DNAP/dNTP combination, Table 2 shows the pseudo-first order rate constant (all values $\times 10^{-3} \text{ min}^{-1}$) and fractional standard deviation, which is the actual standard deviation divided by the rate constant. The average of all fractional standard deviations is 0.20 (i.e., 20%).

Because each reaction used the same DNAP and dNTP concentration, comparisons of observed rate constants reveal how protein modifications affect transition state structure/energy.⁽⁶⁴⁾

dGTP Misinsertion Has –BP-dG Adduct Bulk in the Major Groove. Y-Family DNAPs have a loop above their active site (the “Active Site Loop” or AS-loop), which for Dpo4 is shown pictorially in Figure 4A and structurally in Figure 4B. The AS-loop for Dpo4 starts at V30 on the minor groove side of the active site (yellow amino acids in Figure 4B), makes a turn and returns to V43 on the major groove side (brown amino acids in Figure 4B). In the active site loop, A42 in Dpo4 (Figure 4B) is aligned with V40 in DNAP IV (Figure 4C), whose R-groups protrude into the major groove just above the templating base. Compared to wt-Dpo4, the mutant A42V-Dpo4 has increased bulk on its major groove side and its dGTP misinsertion ($\sim 7.4 \times 10^{-3} \text{ min}^{-1}$) was suppressed by ~ 10 -fold compared to wt-Dpo4 ($\sim 74 \times 10^{-3} \text{ min}^{-1}$), suggesting that –BP adduct bulk is in the major groove during dGTP misinsertion. The A42V-Dpo4 modification had minimal effect on dCTP, dATP and dTTP incorporations (Table 2), suggesting –BP-dG bulk is not in the major groove in these cases.

The AS-loop in Dpo4 (V30-V43) has three more amino acids than in DNAP IV (I31-I41), which are accommodated in Dpo4 as a small upward bulge (Figure 4B, brown arrow) that moves protein mass away from the templating base on the major groove side (as explained in the Figure 4 legend). The three amino acids in this bulge (F37, E38 and D39) were removed to give $\Delta 3$ -Dpo4, whose dGTP misinsertion velocity ($\sim 4.8 \times 10^{-3} \text{ min}^{-1}$) was ~ 15 -fold slower than wt-Dpo4 ($\sim 74 \times 10^{-3} \text{ min}^{-1}$), while dATP, dCTP and dTTP incorporation rates changed minimally for wt-Dpo4 versus $\Delta 3$ -Dpo4. Once again, these findings suggest that –BP adduct bulk is in the major groove during dGTP misinsertion, but not during dATP, dCTP and dTTP incorporations.

The entire wt-Dpo4 AS-loop was replaced by the smaller AS-loop from DNAP IV (I31-I41), which lacks the bulge present in Dpo4, to give the mutant IVloop-Dpo4, whose dGTP misinsertion decreased ~11-fold, suggesting again that –BP-dG bulk is in the major groove. In this case, however, dCTP, dATP and dTTP incorporations were affected by addition of the DNAP IV AS-loop to Dpo4. This behavior was investigated previously⁽⁶⁵⁾ and is discussed below in a subsequent subsection that describes the active site Non-Covalent Bridge (NCB).

The entire wt-Dpo4 AS-loop was replaced by the smaller AS-loop from UmuC (V29-V37), which is the polymerase subunit of EcDNAP V, to give the derivative Vloop-Dpo4, whose dGTP misinsertion rate decreased ~9.5-fold compared to wt-Dpo4. dATP, dCTP and dTTP rates were minimally affected. The smaller UmuC AS-loop is also not expected to have a bulge like in Dpo4.

At its N-terminus, hDNAP κ has an extra 100aa, which forms the U-shaped “N-clasp” (Figure 1C, scarlet ribbon) that restricts access to the major groove side of the active site.⁽¹³⁾ hDNAP κ ’s N-clasp inhibits insertions opposite adducts with bulk on the major groove side of the active site,^(49, 69) and molecular modeling studies have shown why.^(70, 71) When the N-clasp was added to wt-Dpo4 to give N-clasp-Dpo4, dGTP misinsertion ($\sim 8.1 \times 10^{-3} \text{ min}^{-1}$) was ~9.1-fold lower than for wt-Dpo4 itself ($74 \times 10^{-3} \text{ min}^{-1}$). Though we have no direct structural evidence that the N-clasp adopted its proper structure in the N-clasp-Dpo4 derivative, this suppression of dGTP insertion is consistent with expectations if the N-clasp is indeed blocking the major groove as it does in hDNAP κ and is, thus, preventing –BP-adduct bulk from being in the major groove during dGTP misinsertion. Furthermore, N-clasp addition had little effect on dCTP insertion ($\sim 5.1 \times 10^{-3} \text{ min}^{-1}$ vs. $\sim 7.2 \times 10^{-3} \text{ min}^{-1}$), dATP misinsertion ($\sim 11.2 \times 10^{-3} \text{ min}^{-1}$ vs. $\sim 12.2 \times 10^{-3} \text{ min}^{-1}$) or dTTP misinsertion ($\sim 3.8 \times 10^{-3} \text{ min}^{-1}$ vs. $\sim 3.8 \times 10^{-3} \text{ min}^{-1}$), which is

consistent with the expectation that the N-clasp is indeed blocking the major groove, and –BP-adduct bulk is not in the major groove during dATP, dCTP or dTTP incorporation.

The five findings presented in this subsection are all consistent with the notion that –BP-adduct bulk is in the major groove during dGTP misinsertion, but not during dATP, dCTP or dTTP incorporation.

dCTP, dATP and dTTP Incorporation Rates Depend on the Size of an Opening on the Minor Groove Side of the Active Site. Canonical base pairing between –BP-N²-dG and dCTP requires the BP-moiety to be in the minor groove, since the N²-atom is in the minor groove in a Watson-Crick-like base pair. Molecular models revealed that EcDNAP IV has a large opening on the minor groove side of the active site, which helps accommodate -BP-bulk when the adduct-dG moiety is in the active site (Figure 1D)⁽⁶⁶⁾ and might facilitate dCTP insertion. Recent X-ray structures have shown this large minor groove opening in EcDNAP IV.^(17, 26) Furthermore, hDNAP κ has a large slot in the minor groove, which can accommodate BP-bulk.⁽²⁸⁾

In DNAP IV, the lack of R-group bulk on G33 and G74 results in its minor groove opening being large (Figure S1). In the equivalent positions, Dpo4 has V32 and M76, which plug the minor groove opening (Figure S1). V32 and M76 in wt-Dpo4 were converted to glycines to give the double mutant V32G/M76G-Dpo4, and dCTP insertion increased by ~2.4-fold compared to wt-Dpo4. Rates also increased for dATP (~3.5-fold) and dTTP (~3.0-fold) (Table 2). dGTP misinsertion changed minimally (~1.3-fold decrease). This suggests that –BP-bulk is in the minor groove during incorporation of dCTP, dATP and dTTP, but not dGTP.

The AS-loop in hDNAP κ is significantly smaller (G131-L136) than in Dpo4 (V30-V43), DNAP IV (I31-I41) or UmuC (V29-V37), and G131 adopts ϕ/θ -angles unique to glycines.⁽⁷²⁾ Both of these features contribute to hDNAP κ having a large minor groove opening; in fact, it is

a slot,^(13, 28) which is key for accurate bypass of +BP-dG.^(25, 28) Dpo4's AS-loop was replaced by hDNAP κ 's AS-loop to give kloop-Dpo4, whose velocities were ~4.2-fold faster than wt-Dpo4 for dCTP and ~3.4-fold faster for dTTP, while dGTP misinsertion is minimally affected. These findings also suggest that –BP-dG has bulk in the minor groove during dCTP and dTTP, but not dGTP incorporation. However, one finding is inconsistent with naïve expectations: dATP misinsertion did not increase with kloop-Dpo4 ($11.0 \times 10^{-3} \text{ min}^{-1}$ versus $11.2 \times 10^{-3} \text{ min}^{-1}$), which suggests either (i) that some feature of the kloop is able to suppress the reaction rate with whatever minor groove –BP-dG conformation is associated with dATP misinsertion, or (ii) that –BP-dG is not in the minor groove. If kloop addition to Dpo4 re-routed dATP misinsertion to a mechanism involving –BP-bulk in the major groove, then addition of N-clasp should suppress dATP misinsertion, which was not observed: the value for N-clasp/kloop-Dpo4 ($13.0 \times 10^{-3} \text{ min}^{-1}$) was similar to kloop-Dpo4 ($11.0 \times 10^{-3} \text{ min}^{-1}$). This suggests that possibility (i) is more likely, which is addressed further below.

dGTP misinsertion was similar for N-clasp-Dpo4 ($8.1 \times 10^{-3} \text{ min}^{-1}$) and the double mutant N-clasp/kloop-Dpo4 ($7.6 \times 10^{-3} \text{ min}^{-1}$). Similarly, dATP misinsertion was similar for N-clasp-Dpo4 ($8.1 \times 10^{-3} \text{ min}^{-1}$) and the N-clasp/kloop-Dpo4 ($7.6 \times 10^{-3} \text{ min}^{-1}$). dCTP insertion increased somewhat for N-clasp/kloop-Dpo4 (from $7.2 \times 10^{-3} \text{ min}^{-1}$ to $13.1 \times 10^{-3} \text{ min}^{-1}$), implying that kloop addition improved fidelity for dATP and dGTP misinsertion slightly. However, dTTP misinsertion increased ~3.1-fold (from $3.8 \times 10^{-3} \text{ min}^{-1}$ to $12.0 \times 10^{-3} \text{ min}^{-1}$), making this fidelity improvement purine-specific.

A Non-Covalent Bridge Near the Active Site Between the LF-Domain and the CC-Domain Suppresses dATP/dGTP/dTTP Misinsertions. An Active Site Non-Covalent Bridge (AS-NCB) in Dpo4 (Figure 1A) includes two Coulombic interactions: (1) The guanidinium

moiety of R36 in the AS-loop interacts with N254 in the LF-Domain, and (2) the backbone carbonyl-oxygen of R36 interacts with the ammonium moiety of K252 in the LF-Domain (Figure 4B, yellow arrows).

The significance of the AS-NCB to dNTP misinsertion was probed by eliminating both Coulombic interactions: the double mutant R36A/K252A-Dpo4 had higher activity compared to Dpo4 for all three misinsertions: dATP (~4.5-fold), dGTP (~3.7-fold), and dTTP (~2.7-fold), while dCTP changed minimally compared to wt-Dpo4. These findings show that the AS-NCB helps suppress dATP, dGTP and dTTP misinsertion, while minimally affecting correct dCTP insertion.

Considering the single mutants, no dNTP rate was significantly affected by K252A-Dpo4 compared to wt-Dpo4, while with R36A-Dpo4, dATP increased (~3.3-fold), while dCTP, dGTP and dTTP rates changed minimally. This finding is considered in greater depth in Discussion.

The AS-loop and AS-NCB is much different in Dpo4 vs. DNAP IV.⁽⁶⁵⁾ For DNAP IV, we showed that R37 interacts with E251 and R38 interacts with D252 (Figure 4C), with the major consequence being that DNAP IV's AS-NCB suppresses dATP, dGTP and dTTP misinsertions more effectively than the AS-NCB of Dpo4 by ~5-fold (see reference 65).

Distal to the Active Site a Non-Covalent Bridge (Distil-NCB) Between the LF-Domain and the CC-Domain Suppresses dATP/dGTP/dTTP Misinsertions. Dpo4 is an η /V-class-member, while Dbh is a κ /IV-class-member, based on insertion patterns opposite lesions, including a thymine dimer and an AP site.^(20, 27, 73) All four dNTPs are incorporated more slowly by Dbh than Dpo4, though misinsertions are slowed more significantly for dATP (~18-fold), dGTP (~67-fold) and dTTP (~7.8-fold) compared to dCTP (~2.6-fold). Thus, Dbh can also be classified a κ /IV-class-member based on insertion patterns opposite –BP-dG.

Studies have shown that a chimera with the LF-Domain of Dbh joined to the CC-Domain of Dpo4 (Dbh/LF-Domain-Dpo4) behaved more like Dbh than Dpo4 with a thymine dimer and an AP site.^(20, 27, 73) The same tendency is observed with –BP-dG: Dbh/LF-Domain-Dpo4 misinsertion decreased with dATP (~8-fold), dGTP (~34-fold) and dTTP (~7-fold) compared to wt-Dpo4. dCTP insertion decreased also, but less so (~2.5-fold).

To probe what in Dbh's LF-Domain might confer κ /IV-class properties on Dpo4, Dpo4 and Dbh were compared, and the greatest difference in amino acid sequence in the LF-Domain is:

Dpo4:	I ₂₃₇ R T R V R K S I ₂₄₅
Dpo4:	V ₂₃₈ E N K S K I P H ₂₄₆
DNAP IV:	N ₂₃₅ S E R L R K S V ₂₄₃

These nine amino acids from Dbh were placed into Dpo4 to give aa238-246/Dbh-Dpo4, which decreased misinsertion in comparison to wt-Dpo4 for dATP (~2.2-fold), dGTP (~7.2-fold) and dTTP misinsertion (~1.9-fold), while dCTP insertion was minimally affected. While our work was in progress, another group showed that a Dpo4 derivative containing approximately this same region behaved more like Dbh with respect to misinsertion on undamaged DNA and insertion opposite an AP site.⁽²⁰⁾

In X-ray structures^(20, 27, 73), Dbh forms a non-covalent bridge (NCB) between the LF-Domain and the CC-Domain on the minor groove side ~8Å distal to the active site (Figure 1B, blue arrow). One key contact is V101/I244—the latter amino acid being in this region. This Distal-NCB is likely to be the structural element that can suppress dATP, dGTP and dTTP misinsertions opposite –BP-dG.

Our DNAP IV molecular model was built via homology modeling using a Dpo4 starting structure.^(44, 72) Dpo4 has no Distal-NCB (Figure 1A), and yet a Distal-NCB always formed in our DNAP IV models during molecular dynamics. DNAP IV's Distal-NCB formed because P99

kinks a β -strand in the CC-Domain, forcing this β -strand $\sim 3\text{\AA}$ closer to the LF-Domain compared to Dpo4, thus bringing R240 in DNAP IV's LF-Domain close enough to interact with residues in DNAP IV's CC-Domain. At the equivalent position Dpo4 also has arginine (R242), so we investigated whether we had correctly surmised the importance of P99 in DNAP IV by studying the double mutant I101P/A102L-Dpo4, whose rates (compared to wt-Dpo4) decreased for dATP (~ 2.8 -fold) and dGTP (~ 3.5 -fold). (In DNAP IV, L100 fills a hydrophobic pocket to help anchor P99, which is why the double mutant was studied.) dCTP and dTTP insertions were minimally affected. Thus, the addition of I101P/A102L to Dpo4 also suppressed two of three dNTP misinsertions (though not as effectively as the aa238-244/Dbh change) suggesting that a DNAP IV-like Distal-NCB can also improve fidelity in Dpo4.

DISCUSSION

Four Different -BP-dG Adduct Conformations Appear to Be Responsible for the Four Different dNTP Insertions. Dpo4 incorporates all dNTPs opposite -BP-dG at significant rates (Figure 2, Table 2). One mechanism for this could be random insertion; i.e., -BP-dG binds in a single non-informational conformation, which Dpo4 cannot interpret, so it merely incorporates whatever dNTP happens to be in its active site. This mechanism is unlikely, given the following observations, which suggest that each dNTP insertion mechanism involves a unique adduct conformation.

[1] dGTP misinsertion has -BP-dG bulk in the major groove, as it was suppressed by protein structural elements that decreased the size of the major groove opening near the active site (e.g., conversion of A42 to valine), while its velocity was unaffected by changes on the minor groove side.

[2] In contrast, dCTP insertion has -BP-dG bulk in the minor groove, as it was affected by protein modifications on the minor groove side (e.g., adding V32G/M76G to Dpo4), but not the major groove side of the active site.

[3] dATP and dTTP misinsertion also involved -BP-dG bulk in the minor groove, because their rates can be affected by protein modifications on the minor groove side, but not on the major groove side. However, the -BP-dG minor groove conformations cannot be the same for dATP and dTTP compared to dCTP, because the presence of an AS-NCB and/or a Distal-NCB between the LF-Domain and the CC-Domain suppressed dATP and dTTP misinsertions, but not dCTP insertion.

[4] Protein modifications to Dpo4 tended to affect dATP and dTTP misinsertion rates similarly, suggesting that their -BP-dG conformations are more similar; however, two findings

distinguished dATP versus dTTP misinsertion. [i] The rate for dTTP misinsertion was greater for κ loop-Dpo4 than for wt-Dpo4, while dATP misinsertion was similar for κ loop-Dpo4 and wt-Dpo4. [ii] dTTP misinsertion only increased when both the R36/N254 and the K252/R36(C=O) interactions involving the AS-NCB were removed; in contrast removal of the R36/N254 interaction alone significantly suppressed dATP misinsertion. These two differences in dATP versus dTTP misinsertion pattern undoubtedly reflect some difference in –BP-dG conformational positioning in the minor groove, given that both dATP and dTTP must sit similarly in the active site.

We proposed⁽⁷⁴⁾ that different mutations arising from a single BP-dG adduct were caused by different BP-dG adduct conformations, as reviewed extensively.⁽⁷⁵⁾ The analysis in this subsection, suggesting that the four different dNTP insertion pathways arise from four different –BP-adduct conformations, provides further support for this hypothesis. The nature of these four conformations vis-à-vis the four dNTP insertions will be addressed elsewhere.

Additional Analysis of the Major Groove Opening. Adding protein bulk to the opening on the major groove side of Dpo4 decreased dGTP misinsertion, via the A42V modification, via the removal of the three amino acid bulge in Dpo4's AS-loop, or addition of the N-clasp addition. Others have noted that the N-clasp in DNAP κ inhibits bypass of adducts with bulk on the major groove side,^(49, 69) and molecular modeling studies have shown why.^(70, 71) Still others have speculated that having bulk at the position equivalent to A42 in Dpo4 might sterically exclude certain kinds of lesions; e.g., M135 in DNAP κ was proposed to help exclude the entry of thymine dimers into the active site from the major groove side.⁽¹⁶⁾ Our finding—that the small bulge in Dpo4's AS-loop moves protein bulk away from the templating base—has not been

noted previously and is a structural feature present in other η /V-class-members (discussed below).

Non-Covalent Bridges Orient the LF-Domain to Prevent Misinsertion Opposite

Aberrant Adduct Structures. The AS-NCB and Distal-NCB are on the minor groove side of the active site, so it is not surprising that they suppress dATP and dTTP misinsertions, which have –BP-dG bulk in the minor groove. However, the AS-NCB and Distal-NCB do not suppress dCTP insertion, even though it also has –BP-dG bulk in the minor groove. Furthermore, the AS-NCB and Distal-NCB do suppress dGTP misinsertion, which has bulk in the major groove. Thus, the formation of non-covalent bridges on the minor groove side per se cannot explain their selective ability to suppress all three misinsertions. It is more likely that these NCBs help position the LF-Domain so it can exclude aberrant DNA structures from entering the active site, independent of whether the aberrant structures have adduct bulk in the minor groove or the major groove. Others have shown that the Distal-NCB is likely to be important for Dbh functioning vis-à-vis other bypass events. ^(20, 27, 73)

Generalizations about Y-Family DNAP Structure. Dpo4 has an upward bulge on the major groove side of its active site (Figure 4C: F37/E38/D39); our findings suggest that this bulge allows adduct-bulk to be better accommodated on the major groove side of the active site based on our findings for dGTP misinsertion opposite -BP-dG. While this bulge promotes dGTP misincorporations opposite –BP-dG, which is counter-productive, it appears to be helpful for allowing entry of other lesions (e.g.) UV-damage) into the active site. We note that human and yeast DNAP η have upward bulges analogous to the one in Dpo4, ^(16, 21) suggesting that this feature might be important in many η /V-class-members.

η /V-class Y-Family DNAPs have plugged minor groove openings, as observed with Dpo4 (V32 and M76 in Figure S1). The minor groove openings are even more plugged in human- and yeast-DNAP η , which have, respectively, unique seven and thirty-four amino acid inserts that completely cap their respective minor groove openings.^(16, 21) The minor groove opening is also plugged with UmuC.⁽⁷²⁾

By simply grafting the N-clasp from hDNAP κ onto Dpo4, we generated a derivative (N-clasp-Dpo4) that suppressed dGTP misinsertion, thus improving fidelity compared to wt-Dpo4 itself (Table 2). We were not optimistic that this experiment would work as envisioned, though molecular modeling suggested that it might (data not shown); furthermore, we do not even know if these extra 100 amino acids are truly operating in N-clasp-Dpo4 analogously to how they operate in hDNAP κ . Nevertheless—by whatever mechanism—the findings show that if a Dpo4-like DNAP were simply to acquire an N-clasp-like sequence from some genetic source at some point during evolution, then fidelity of BP-adduct bypass would improve. Thereafter, one can speculate that the AS-loop might have shrunk—this supposition being based on the observation that kloop addition to N-clasp-Dpo4 (giving kloop/N-clasp-Dpo4) improved both dATP and dGTP fidelity. Next, DNAP κ 's N-clasp might have acquired its non-covalent bridge (Figure 1C, scarlet arrow), which is on the major groove side of the active site, along with the losses of the minor groove side. While this pathway for the generation of DNAP κ 's basic structure is obviously highly speculative, it seems sensible.

In summary, κ /IV-class-members have greater protein bulk on the major groove side of their active sites and more open space on the minor groove side, while η /V-class-members are the opposite. To achieve these goals, prokaryotic and Archaea Y-Family DNAPs principally vary the number of amino acids and/or R-Group bulk in the AS-Loop. Eukaryotic Y-Family DNAPs

have also accomplished these goals, though via the grafting on of significant protein inserts; e.g., the N-clasp to block the major groove in DNAP κ , and the aforementioned inserts in DNAP η (seven and thirty-four amino acids, respectively, for yeast and humans) to block the minor groove opening.

ACKNOWLEDGEMENTS: We thank Dr. Nicholas Geacintov for providing the –BP-dG- containing oligonucleotide, and Dr. Roger Woodgate for supplying several Dpo4- and Dbh- containing plasmids, along with Drs. Louise and Satya Prakash for supplying a plasmid containing the gene for human DNAP κ .

SUPPORTING INFORMATION: Supplemental information is provided.

AUTHOR INFORMATION:

Corresponding Author

*Email: loechler@bu.edu

Funding

This work was supported by a grant from the NIH (R01ES03775 to ELL).

Notes

The authors declare no competing financial interests.

REFERENCES

1. McCulloch, S. D., and Kunkel, T. A. (2008) The fidelity of DNA synthesis by eukaryotic replicative and translesion synthesis polymerases, *Cell Res* 18, 148-161.
2. Lange, S. S., Takata, K., and Wood, R. D. (2011) DNA polymerases and cancer, *Nat Rev Cancer* 11, 96-110.
3. van Loon, B., Woodgate, R., and Hubscher, U. (2015) DNA polymerases: Biology, diseases and biomedical applications, *DNA repair* 29, 1-3.
4. Yang, W. (2003) Damage repair DNA polymerases Y, *Curr Opin Struct Biol* 13, 23-30.
5. Pata, J. D. (2010) Structural diversity of the Y-family DNA polymerases, *Biochim Biophys Acta* 1804, 1124-1135.
6. Washington, M. T., Carlson, K. D., Freudenthal, B. D., and Pryor, J. M. (2010) Variations on a theme: eukaryotic Y-family DNA polymerases, *Biochim Biophys Acta* 1804, 1113-1123.
7. Yang, W. (2014) An overview of Y-Family DNA polymerases and a case study of human DNA polymerase eta, *Biochemistry* 53, 2793-2803.
8. Eoff, R. L., Choi, J. Y., and Guengerich, F. P. (2010) Mechanistic Studies with DNA Polymerases Reveal Complex Outcomes following Bypass of DNA Damage, *Journal of nucleic acids* 2010, doi: 10.4061/2010/830473.
9. Chandani, S., Jacobs, C., and Loechler, E. L. (2010) Architecture of y-family DNA polymerases relevant to translesion DNA synthesis as revealed in structural and molecular modeling studies, *Journal of nucleic acids* 2010, doi: 10.4061/2010/784081.
10. Zhou, B. L., Pata, J. D., and Steitz, T. A. (2001) Crystal structure of a DinB lesion bypass DNA polymerase catalytic fragment reveals a classic polymerase catalytic domain, *Mol Cell* 8, 427-437.
11. Ling, H., Boudsocq, F., Woodgate, R., and Yang, W. (2001) Crystal structure of a Y-family DNA polymerase in action: a mechanism for error-prone and lesion-bypass replication, *Cell* 107, 91-102.
12. Trincão, J., Johnson, R. E., Escalante, C. R., Prakash, S., Prakash, L., and Aggarwal, A. K. (2001) Structure of the catalytic core of *S. cerevisiae* DNA polymerase eta: implications for translesion DNA synthesis, *Mol Cell* 8, 417-426.
13. Lone, S., Townson, S. A., Uljon, S. N., Johnson, R. E., Brahma, A., Nair, D. T., Prakash, S., Prakash, L., and Aggarwal, A. K. (2007) Human DNA polymerase kappa encircles DNA: implications for mismatch extension and lesion bypass, *Mol Cell* 25, 601-614.
14. Ummat, A., Silverstein, T. D., Jain, R., Buku, A., Johnson, R. E., Prakash, L., Prakash, S., and Aggarwal, A. K. (2012) Human DNA polymerase eta is pre-aligned for dNTP binding and catalysis, *J Mol Biol* 415, 627-634.
15. Nair, D. T., Johnson, R. E., Prakash, L., Prakash, S., and Aggarwal, A. K. (2006) Hoogsteen base pair formation promotes synthesis opposite the 1,N⁶-ethenodeoxyadenosine lesion by human DNA polymerase iota, *Nat Struct Mol Biol* 13, 619-625.
16. Biertumpfel, C., Zhao, Y., Kondo, Y., Ramon-Maiques, S., Gregory, M., Lee, J. Y., Masutani, C., Lehmann, A. R., Hanaoka, F., and Yang, W. (2010) Structure and mechanism of human DNA polymerase eta, *Nature* 465, 1044-1048.

17. Sharma, A., Kottur, J., Narayanan, N., and Nair, D. T. (2013) A strategically located serine residue is critical for the mutator activity of DNA polymerase IV from *Escherichia coli*, *Nucleic acids research* 41, 5104-5114.
18. Kirouac, K. N., Basu, A. K., and Ling, H. (2013) Structural mechanism of replication stalling on a bulky amino-polycyclic aromatic hydrocarbon DNA adduct by a γ family DNA polymerase, *J Mol Biol* 425, 4167-4176.
19. Kirouac, K. N., Basu, A. K., and Ling, H. (2013) Replication of a carcinogenic nitropyrene DNA lesion by human Y-family DNA polymerase, *Nucleic acids research* 41, 2060-2071.
20. Wilson, R. C., Jackson, M. A., and Pata, J. D. (2013) Y-family polymerase conformation is a major determinant of fidelity and translesion specificity, *Structure* 21, 20-31.
21. Silverstein, T. D., Johnson, R. E., Jain, R., Prakash, L., Prakash, S., and Aggarwal, A. K. (2010) Structural basis for the suppression of skin cancers by DNA polymerase eta, *Nature* 465, 1039-1043.
22. Rechkoblit, O., Kolbanovskiy, A., Malinina, L., Geacintov, N. E., Broyde, S., and Patel, D. J. (2010) Mechanism of error-free and semitargeted mutagenic bypass of an aromatic amine lesion by Y-family polymerase Dpo4, *Nat Struct Mol Biol* 17, 379-388.
23. Bauer, J., Xing, G., Yagi, H., Sayer, J. M., Jerina, D. M., and Ling, H. (2007) A structural gap in Dpo4 supports mutagenic bypass of a major benzo[a]pyrene dG adduct in DNA through template misalignment, *Proceedings of the National Academy of Sciences of the United States of America* 104, 14905-14910.
24. Ling, H., Sayer, J. M., Plosky, B. S., Yagi, H., Boudsocq, F., Woodgate, R., Jerina, D. M., and Yang, W. (2004) Crystal structure of a benzo[a]pyrene diol epoxide adduct in a ternary complex with a DNA polymerase, *Proceedings of the National Academy of Sciences of the United States of America* 101, 2265-2269.
25. Liu, Y., Yang, Y., Tang, T. S., Zhang, H., Wang, Z., Friedberg, E., Yang, W., and Guo, C. (2014) Variants of mouse DNA polymerase kappa reveal a mechanism of efficient and accurate translesion synthesis past a benzo[a]pyrene dG adduct, *Proceedings of the National Academy of Sciences of the United States of America* 111, 1789-1794.
26. Kottur, J., Sharma, A., Gore, K. R., Narayanan, N., Samanta, B., Pradeepkumar, P. I., and Nair, D. T. (2015) Unique structural features in DNA polymerase IV enable efficient bypass of the N2 adduct induced by the nitrofurazone antibiotic, *Structure* 23, 56-67.
27. Mukherjee, P., Wilson, R. C., Lahiri, I., and Pata, J. D. (2014) Three residues of the interdomain linker determine the conformation and single-base deletion fidelity of Y-family translesion polymerases, *The Journal of biological chemistry* 289, 6323-6331.
28. Jha, V., Bian, C., Xing, G., and Ling, H. (2016) Structure and mechanism of error-free replication past the major benzo[a]pyrene adduct by human DNA polymerase kappa, *Nucleic acids research* 44, 4957-4967.
29. Rechkoblit, O., Malinina, L., Cheng, Y., Kuryavyi, V., Broyde, S., Geacintov, N. E., and Patel, D. J. (2006) Stepwise translocation of Dpo4 polymerase during error-free bypass of an oxoG lesion, *PLoS Biol* 4, e11.
30. Wong, J. H., Fiala, K. A., Suo, Z., and Ling, H. (2008) Snapshots of a Y-family DNA polymerase in replication: substrate-induced conformational transitions and implications for fidelity of Dpo4, *J Mol Biol* 379, 317-330.

31. Wang, L., Yu, X., Hu, P., Broyde, S., and Zhang, Y. (2007) A water-mediated and substrate-assisted catalytic mechanism for *Sulfolobus solfataricus* DNA polymerase IV, *Journal of the American Chemical Society* 129, 4731-4737.
32. Gadkari, V. V., Tokarsky, E. J., Malik, C. K., Basu, A. K., and Suo, Z. (2014) Mechanistic investigation of the bypass of a bulky aromatic DNA adduct catalyzed by a Y-family DNA polymerase, *DNA repair* 21, 65-77.
33. Vyas, R., Efthimiopoulos, G., Tokarsky, E. J., Malik, C. K., Basu, A. K., and Suo, Z. (2015) Mechanistic Basis for the Bypass of a Bulky DNA Adduct Catalyzed by a Y-Family DNA Polymerase, *Journal of the American Chemical Society* 137, 12131-12142.
34. Nevin, P., Lu, X., Zhang, K., Engen, J. R., and Beuning, P. J. (2015) Noncognate DNA damage prevents the formation of the active conformation of the Y-family DNA polymerases DinB and DNA polymerase kappa, *The FEBS journal* 282, 2646-2660.
35. Eoff, R. L., Sanchez-Ponce, R., and Guengerich, F. P. (2009) Conformational changes during nucleotide selection by *Sulfolobus solfataricus* DNA polymerase Dpo4, *The Journal of biological chemistry* 284, 21090-21099.
36. Brenlla, A., Rueda, D., and Romano, L. J. (2015) Mechanism of aromatic amine carcinogen bypass by the Y-family polymerase, Dpo4, *Nucleic acids research* 43, 9918-9927.
37. Seo, K. Y., Nagalingam, A., Miri, S., Yin, J., Chandani, S., Kolbanovskiy, A., Shastry, A., and Loechler, E. L. (2006) Mirror image stereoisomers of the major benzo[a]pyrene N2-dG adduct are bypassed by different lesion-bypass DNA polymerases in *E. coli*, *DNA repair* 5, 515-522.
38. Yin, J., Seo, K. Y., and Loechler, E. L. (2004) A role for DNA polymerase V in G --> T mutations from the major benzo[a]pyrene N2-dG adduct when studied in a 5'-TGT sequence in *E. coli*, *DNA repair* 3, 323-334.
39. Seo, K. Y., Yin, J., Donthamsetti, P., Chandani, S., Lee, C. H., and Loechler, E. L. (2009) Amino acid architecture that influences dNTP insertion efficiency in Y-family DNA polymerase V of *E. coli*, *J Mol Biol* 392, 270-282.
40. Sholder, G., and Loechler, E. L. (2015) A method to accurately quantitate intensities of (32)P-DNA bands when multiple bands appear in a single lane of a gel is used to study dNTP insertion opposite a benzo[a]pyrene-dG adduct by *Sulfolobus* DNA polymerases Dpo4 and Dbh, *DNA repair* 25, 97-103.
41. Seo, K. Y., Nagalingam, A., Tiffany, M., and Loechler, E. L. (2005) Mutagenesis studies with four stereoisomeric N2-dG benzo[a]pyrene adducts in the identical 5'-CGC sequence used in NMR studies: G->T mutations dominate in each case, *Mutagenesis* 20, 441-448.
42. Harvey, R. G. (1997) *Polycyclic Aromatic Hydrocarbons: Chemistry and Cancer*.
43. Pfeifer, G. P., and Hainaut, P. (2003) On the origin of G --> T transversions in lung cancer, *Mutation research* 526, 39-43.
44. Lee, C. H., Chandani, S., and Loechler, E. L. (2006) Homology modeling of four Y-family, lesion-bypass DNA polymerases: the case 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* 25, 87-102.
45. Shen, X., Sayer, J. M., Kroth, H., Ponten, I., O'Donnell, M., Woodgate, R., Jerina, D. M., and Goodman, M. F. (2002) Efficiency and accuracy of SOS-induced DNA polymerases

- replicating benzo[a]pyrene-7,8-diol 9,10-epoxide A and G adducts, *The Journal of biological chemistry* 277, 5265-5274.
46. Lenne-Samuel, N., Janel-Bintz, R., Kolbanovskiy, A., Geacintov, N. E., and Fuchs, R. P. (2000) 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, 299-307.
47. Napolitano, R., Janel-Bintz, R., Wagner, J., and Fuchs, R. P. (2000) All three SOS-inducible DNA polymerases (Pol II, Pol IV and Pol V) are involved in induced mutagenesis, *Embo J* 19, 6259-6265.
48. Jarosz, D. F., Godoy, V. G., Delaney, J. C., Essigmann, J. M., and Walker, G. C. (2006) A single amino acid governs enhanced activity of DinB DNA polymerases on damaged templates, *Nature* 439, 225-228.
49. Rechkoblit, O., Zhang, Y., Guo, D., Wang, Z., Amin, S., Krzeminsky, J., Louneva, N., and Geacintov, N. E. (2002) trans-Lesion synthesis past bulky benzo[a]pyrene diol epoxide N2-dG and N6-dA lesions catalyzed by DNA bypass polymerases, *The Journal of biological chemistry* 277, 30488-30494.
50. Ogi, T., Shinkai, Y., Tanaka, K., and Ohmori, H. (2002) Polkappa protects mammalian cells against the lethal and mutagenic effects of benzo[a]pyrene, *Proceedings of the National Academy of Sciences of the United States of America* 99, 15548-15553.
51. Avkin, S., Goldsmith, M., Velasco-Miguel, S., Geacintov, N., Friedberg, E. C., and Livneh, Z. (2004) Quantitative analysis of translesion DNA synthesis across a benzo[a]pyrene-guanine adduct in mammalian cells: the role of DNA polymerase kappa, *The Journal of biological chemistry* 279, 53298-53305.
52. Huang, X., Kolbanovskiy, A., Wu, X., Zhang, Y., Wang, Z., Zhuang, P., Amin, S., and Geacintov, N. E. (2003) 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, 2456-2466.
53. Suzuki, N., Ohashi, E., Kolbanovskiy, A., Geacintov, N. E., Grollman, A. P., Ohmori, H., and Shibutani, S. (2002) 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, 6100-6106.
54. Zhang, Y., Wu, X., Guo, D., Rechkoblit, O., and Wang, Z. (2002) 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* 1, 559-569.
55. Zhang, H., and Guengerich, F. P. (2010) Effect of N2-guanyl modifications on early steps in catalysis of polymerization by *Sulfolobus solfataricus* P2 DNA polymerase Dpo4 T239W, *J Mol Biol* 395, 1007-1018.
56. Zhang, H., Eoff, R. L., Kozekov, I. D., Rizzo, C. J., Egli, M., and Guengerich, F. P. (2009) Structure-function relationships in miscoding by *Sulfolobus solfataricus* DNA polymerase Dpo4: guanine N2,N2-dimethyl substitution produces inactive and miscoding polymerase complexes, *The Journal of biological chemistry* 284, 17687-17699.
57. Yuan, B., Cao, H., Jiang, Y., Hong, H., and Wang, Y. (2008) Efficient and accurate bypass of N2-(1-carboxyethyl)-2'-deoxyguanosine by DinB DNA polymerase in vitro and in vivo, *Proceedings of the National Academy of Sciences of the United States of America* 105, 8679-8684.

58. Minko, I. G., Kozekov, I. D., Harris, T. M., Rizzo, C. J., Lloyd, R. S., and Stone, M. P. (2009) Chemistry and biology of DNA containing 1,N(2)-deoxyguanosine adducts of the alpha,beta-unsaturated aldehydes acrolein, crotonaldehyde, and 4-hydroxynonenal, *Chem Res Toxicol* 22, 759-778.
59. Fukuda, H., Takamura-Enya, T., Masuda, Y., Nohmi, T., Seki, C., Kamiya, K., Sugimura, T., Masutani, C., Hanaoka, F., and Nakagama, H. (2009) Translesional DNA synthesis through a C8-guanyl adduct of 2-amino-1-methyl-6-phenylimidazo[4,5-b]pyridine (PhIP) in Vitro: REV1 inserts dC opposite the lesion, and DNA polymerase kappa potentially catalyzes extension reaction from the 3'-dC terminus, *The Journal of biological chemistry* 284, 25585-25592.
60. Fischhaber, P. L., Gerlach, V. L., Feaver, W. J., Hatahet, Z., Wallace, S. S., and Friedberg, E. C. (2002) Human DNA polymerase kappa bypasses and extends beyond thymine glycols during translesion synthesis in vitro, preferentially incorporating correct nucleotides, *The Journal of biological chemistry* 277, 37604-37611.
61. Jalszynski, P., Ohashi, E., Ohmori, H., and Nishimura, S. (2005) Error-prone and inefficient replication across 8-hydroxyguanine (8-oxoguanine) in human and mouse ras gene fragments by DNA polymerase kappa, *Genes Cells* 10, 543-550.
62. Minko, I. G., Harbut, M. B., Kozekov, I. D., Kozekova, A., Jakobs, P. M., Olson, S. B., Moses, R. E., Harris, T. M., Rizzo, C. J., and Lloyd, R. S. (2008) Role for DNA polymerase kappa in the processing of N2-N2-guanine interstrand cross-links, *The Journal of biological chemistry* 283, 17075-17082.
63. Fuchs, R. P., Fujii, S., and Wagner, J. (2004) Properties and functions of Escherichia coli: Pol IV and Pol V, *Advances in protein chemistry* 69, 229-264.
64. Wells, J. A. (1990) Additivity of mutational effects in proteins, *Biochemistry* 29, 8509-8517.
65. Sholder, G., Creech, A., and Loechler, E. L. (2015) How Y-Family DNA polymerase IV is more accurate than Dpo4 at dCTP insertion opposite an N2-dG adduct of benzo[a]pyrene, *DNA repair* 35, 144-153.
66. Chandani, S., and Loechler, E. L. (2007) Molecular modeling benzo[a]pyrene N2-dG adducts in the two overlapping active sites of the Y-family DNA polymerase Dpo4, *J Mol Graph Model* 25, 658-670.
67. Johnson, R. E., Prakash, S., and Prakash, L. (2000) The human DINB1 gene encodes the DNA polymerase Poltheta, *Proceedings of the National Academy of Sciences of the United States of America* 97, 3838-3843.
68. Drouin, E. E., Lech, J., and Loechler, E. L. (1995) The major, N2-Gua adduct of the (+)-anti-benzo[a]pyrene diol epoxide can be unstable in double-stranded DNA, *Biochemistry* 34, 2251-2259.
69. Walsh, J. M., Ippoliti, P. J., Ronayne, E. A., Rozners, E., and Beuning, P. J. (2013) Discrimination against major groove adducts by Y-family polymerases of the DinB subfamily, *DNA repair* 12, 713-722.
70. Jia, L., Geacintov, N. E., and Broyde, S. (2008) The N-clasp of human DNA polymerase kappa promotes blockage or error-free bypass of adenine- or guanine-benzo[a]pyrenyl lesions, *Nucleic acids research* 36, 6571-6584.
71. Lior-Hoffmann, L., Ding, S., Geacintov, N. E., Zhang, Y., and Broyde, S. (2014) Structural and dynamic characterization of polymerase kappa's minor groove lesion processing reveals how adduct topology impacts fidelity, *Biochemistry* 53, 5683-5691.

72. Chandani, S., and Loechler, E. L. (2009) Y-Family DNA polymerases may use two different dNTP shapes for insertion: a hypothesis and its implications, *J Mol Graph Model* 27, 759-769.
73. Boudsocq, F., Kokoska, R. J., Plosky, B. S., Vaisman, A., Ling, H., Kunkel, T. A., Yang, W., and Woodgate, R. (2004) Investigating the role of the little finger domain of Y-family DNA polymerases in low fidelity synthesis and translesion replication, *The Journal of biological chemistry* 279, 32932-32940.
74. Rodriguez, H., and Loechler, E. L. (1993) Mutagenesis by the (+)-anti-diol epoxide of benzo[a]pyrene: what controls mutagenic specificity?, *Biochemistry* 32, 1759-1769.
75. Seo, K. Y., Jelinsky, S. A., and Loechler, E. L. (2000) Factors that influence the mutagenic patterns of DNA adducts from chemical carcinogens, *Mutation research* 463, 215-246.
76. Wilson, R. C., Jackson, M. A., and Pata, J. D. Y-family polymerase conformation is a major determinant of fidelity and translesion specificity, *Structure* 21, 20-31.

Table 1: Sequences of Dpo4 Derivatives Used in This Work.¹

	10	20	30	40	50	60	70	80	90	100	110	120	130	140	150	160	170	180	190	200	210	220	230	240	250	
wt-Dpo4	MIVLFVDFDYFYAQVEEVLNPSLKGKPVVVCVFSGRFEDSGAVATANYEA																									
NC	X																									
A42V																										
D3																										
Vloop																										
IVLoop																										
R37A/IVloop/M76																										
R38A/IVloop/M76																										
V32G/M76G																										
Kloop																										
NC/Kloop	X																									
R36A																										
K252A																										
R36A/K252A																										
aa238-246/Dbh																										
I101P/A102L																										
LF/Dbh-Dpo4																										
QNKYSEFVENKSKIPHGRYLTLPYN->																										
X (N-clasp)	MDSTKEKCDSYKDDLLRLMGLNDNKAGMEGLDKKINKIIMEATKGSRYGNEIKKEKQVNR	IENMMQKAQITTSQQLRKAQLQVDRFAMELEQSRNLS																								

¹ The sequences of relevant regions of wt-Dpo4 are given at the top of the table, with amino acid changes for each Dpo4 derivative shown below. The “^” symbol represents absent or deleted amino acids. The derivative LF/Dbh-Dpo4 is a chimera with the Dpo4 sequence to A229, at which point it switches to Dbh sequences for the LF-Domain as indicated. “X” represents the N-clasp, which is the first 100 amino acids of hDNAP κ , whose sequence is given at the bottom of the table. The N-clasp is present in the derivatives “NC-Dpo4” and “NC/ κ loop-Dpo4.”

Table 2: Pseudo First-Order Rate Constants for dNTP Insertion Opposite –BP-dG for the DNA Polymerases wt-Dpo4 and wt-Dbh, as well as Derivatives of Dpo4.^{1,2}

Wild Type	dATP		dCTP		dGTP		dTTP	
wt-Dpo4	11.2	0.25	5.1	0.25	73.8	0.20	3.8	0.15
wt-Dbh	0.64	0.029	2.0	0.10	1.1	0.29	0.49	0.29
Dpo4 Derivatives	Major Groove Considerations							
A42V	12.0	0.10	4.9	0.081	7.4	0.21	3.4	0.24
Δ3	9.7	0.27	5.5	0.23	4.8	0.41	3.5	0.29
IVloop	4.9	0.19	9.6	0.21	6.7	0.21	2.7	0.31
Vloop	9.8	0.16	5.2	0.075	7.8	0.12	3.0	0.23
N-Clasp	12.2	0.21	7.2	0.24	8.1	0.12	3.8	0.29
	Minor Groove Opening Size Considerations							
V32G/M76G	39.4	0.20	12.3	0.13	55.1	0.12	11.3	0.12
κloop	11.0	0.20	21.8	0.19	70.9	0.26	13.0	0.11
N-Clasp/κloop	13.5	0.15	13.1	0.071	7.6	0.11	12.0	0.036
	AS-NCB Considerations							
R36A	36.6	0.19	6.7	0.11	71.0	0.25	4.2	0.17
K252A	9.8	0.049	5.0	0.16	67.7	0.12	4.3	0.21
R36A/K252A	50.3	0.38	6.5	0.21	276.	0.28	10.3	0.57
	Distal-NCB Considerations							
LF/Dbh	1.4	0.15	2.1	0.12	2.2	0.063	0.54	0.35
aa238-246/Dbh	5.2	0.24	5.2	0.29	10.3	0.38	2.0	0.21
I101P/A102L	4.0	0.36	4.8	0.94	21.2	0.29	3.1	0.80

¹ Pseudo first-order rate constants are reported for dNTP insertion opposite a –BP-dG adduct in a primer/template (left-most column for each dNTP) using standard conditions (dNTPs/100μM and DNAP/124nM, at pH 7.6 and 37°C); e.g., 11.2 x 10⁻³min⁻¹ is the value for Dpo4 with dATP. All values are the average of at least three determinations and are reported [x 10⁻³min⁻¹]. Purified proteins included wt-Dpo4, wt-Dbh and Dpo4 mutant derivatives (Table 1). Methodological details are summarized in Supplemental Materials and described in reference ⁽⁴⁰⁾.

² Standard deviations are reported in the right-most columns for each dNTP and are expressed as a fraction of the rate constants; e.g., for wt-Dpo4 with dATP, the true standard deviation (2.8 x 10⁻³min⁻¹) is divided by its pseudo first-order rate constant (11.2 x 10⁻³min⁻¹) to give its fractional standard deviation (i.e., 0.25).

FIGURE 1: Views from the minor groove side of the active site for the Y-Family DNAPs:

SsDpo4 (Panel A), SaDbh (Panel B), hDNAP κ (Panel C) and EcENAP IV (Panel D). The CC-Domains are shown in green and the LF-Domains are shown in yellow. Amino acids involved in the formation of noncovalent bridges (NCBs) between the CC-Domain and the LF-Domain are shown in brown, which can be of three types: near the active site (AS-NCB, yellow arrow) and distal to the active site on the minor groove side (Distal-NCB, blue arrow), or on the major groove side near the active site (N-clasp-NCB, scarlet arrow), as is present uniquely in hDNAP κ , because of its N-clasp (scarlet structure). X-ray sources of structures include SsDpo4 ⁽²⁴⁾ and hDNAP κ , ⁽¹³⁾ while modeled structures are used for SaDbh and DNAP IV ⁽⁷²⁾ for the following reasons. The X-ray structures for SaDbh ^(27, 73, 76) are with undamaged DNA containing a one-nucleotide bulge, which prevents AS-NCB formation; the structure shown is modeled with the DNA bulge removed, which allowed the AS-NCB to reform during molecular dynamics. Though X-ray structures exist for DNAP IV ^(17, 26), they do not have a BP-adduct, so we include our model with –BP-dG (red), ⁽⁷²⁾ which can be seen through the opening on the minor groove side of the active site. DNA is gray.

FIGURE 2: Structures of +BP-dG and –BP-dG, along with a portion of the primer/template sequence and a phosphorimage showing Y-Family DNAPs SsDpo4, SaDbh and derivatives incorporating dATP, dCTP, dGTP and dTTP incorporation into a ³²P-containing primer. Primer alone is shown in lane 1. Images for the SsDpo4 set, the SaDbh set and the primer were in the same gel (same images as in reference 37), but were repositioned in the figure. In addition, typical phosphorimages for each of the DNAPs listed in Table 2 are shown. The images come

from many different experiments. Four lanes are shown for each DNAP arranged (left-to-right) dATP, dCTP, dGTP and dTTP, with incubations for 60min, 60min, 20min and 60min, respectively, except R36A/K252A-Dpo4, whose dGTP reaction was 5min.

FIGURE 3: Scan of the dCTP primer extension reaction by SsDpo4 shown in Figure 2. Primer alone is shown in red. The method for computing the fraction of extended primer versus unextended primer is described in detail in reference ⁽⁴⁰⁾.

FIGURE 4: The active site loop (AS-loop) of Dpo4 (V30-V43) is shown pictorially in Panel A and structurally in Panel B, while DNAP IV AS-loop (I31-I40) is shown structurally in Panel C. The view is from the major groove side of the active site. AS-loops in Y-Family DNAPs begin on the minor groove side of the active site, which is represented by amino acids in yellow, and then turns back to the major groove side of the active site, which is represented by amino acids in brown. On the major groove side above the templating base, Dpo4 has A42, while DNAP IV has V40. Dpo4's AS-NCB between the CC-Domain and LF-Domain-Doman is anchored by two Coulombic interactions: R36/N254 and K252/R36(C=O) (yellow arrows). The AS-loop in DNAP IV has three fewer amino acids than in Dpo4, which are accommodated as a bulge (F37/E38/D39, brown arrow) that moves bulk away from the active site. R36 in Dpo4 (purple) is aligned with R37 in DNAP IV (purple), and the bulge in Dpo4 insures that the next amino acid (F37, red) is above R36 and away from the active site, whereas the next amino acid in DNAP IV (R38, red) is below R37 and closer to the active site. DNAP IV's AS-NCB between the CC-Domain and LF-Domain-Doman is anchored by two Coulombic interactions: R37/E251 and R38/D252 (yellow arrows).

Sholder & Loechler Figure 1

Shen et al. 2019 Figure 1

Sholder & Loechler Figure 3

Sholder & Loechler Figure 4