# Mechanism of Error-Free Bypass of the Environmental Carcinogen N-(2′-Deoxyguanosin-8-yl)-3-aminobenzanthrone Adduct by Human DNA Polymerase η

Amritraj Patra[+],[a] Dustin A. Politica[+],[b] Arindom Chatterjee,[c] E. John Tokarsky,[d] Zucai Suo,[d] Ashis K. Basu,[c] Michael P. Stone,*[b] and Martin Egli*[a]

The environmental pollutant 3-nitrobenzanthrone produces bulky aminobenzanthrone (ABA) DNA adducts with both guanine and adenine nucleobases. A major product occurs at the C8 position of guanine (C8-dG-ABA). These adducts present a strong block to replicative polymerases but, remarkably, can be bypassed in a largely error-free manner by the human Y-family polymerase η (hPolη). Here, we report the crystal structure of a ternary Pol·DNA·dCTP complex between a C8-dG-ABA-containing template:primer duplex and hPolη. The complex was captured at the insertion stage and provides crucial insight into the mechanism of error-free bypass of this bulky lesion. Specifically, bypass involves accommodation of the ABA moiety inside a hydrophobic cleft to the side of the enzyme active site and formation of an intra-nucleotide hydrogen bond between the phosphate and ABA amino moiety, allowing the adducted guanine to form a standard Watson–Crick pair with the incoming dCTP.

3-Nitrobenzanthrone (3-nitro-7H-benz-[d,e]anthracen-7-one, 3-NBA, Scheme 1) is a product of incomplete combustion[1,2] and has been identified as a component of diesel exhaust[2–4] and as an environmental contaminant.[1,5] The International Association for Research on Cancer (IARC) has classified diesel exhaust as carcinogenic to humans,[6] and 3-NBA is a suspected culprit of this carcinogenicity.[1,6] 3-NBA has been found to be both highly mutagenic[7–10] and genotoxic,[11–15] and is itself classified as a potential human carcinogen by the IARC.[6]

Enzymatic reduction in vivo engenders a conversion of 3-NBA to aminobenzanthrone (ABA, Scheme 1).[16,17] Intermedi-
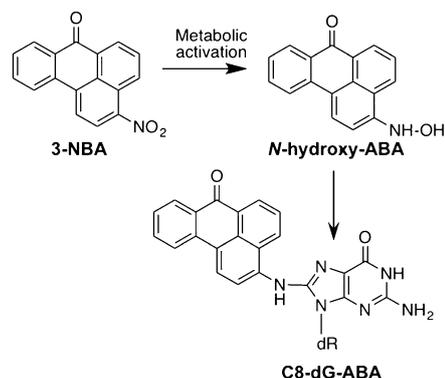
Scheme 1. Chemical structures of 3-NBA, N-OH-ABA, and one of the major in vivo ABA adducts: C8-dG-ABA.

ates in the reduction pathway can form reactive nitrenium ions that alkylate DNA, thereby resulting in ABA adducts.[18–20] Three different adducts have been identified; they occur on the C8 (Scheme 1) or N^2 positions of guanine (major adducts), or the N^6 position of adenine (minor adduct).[21,22] Due to the bulky nature of these adducts, it is believed that they present strong blocks to replicative polymerases, and bypass is instead accomplished by lower-fidelity translesion synthesis (TLS) polymerases (Pols).[8,23,24] In humans, these include the Y-family Pols η, ι, κ, and Rev1.[23–25]

However, C8-dG-ABA, one of the two major adducts formed,[8,21] has previously been implicated in significant blocking of bypass by TLS polymerases in a nucleotide excision-repair-deficient human cell line.[8] When bypass did occur, of the two dG and one dA adducts, the C8-dG-ABA adduct triggered the highest mutation frequency.[8] Nevertheless, the majority of bypass events were non-mutagenic, with nearly 70% correct incorporation of cytosine.[8] Another study specifically examined the C8-dG-ABA adduct in human embryonic kidney (HEK293T) cells, and also observed correct incorporation of cytosine opposite this lesion in 86% of bypass events.[26] A series of siRNA knockdown experiments showed that each of the Y-family polymerases contributed to the bypass of this lesion.[26] However, hPolη played the most significant role in the mutagenic insertion step, with a decrease in mutation frequency of 39% upon knockdown of hPolη.[26]

Using a physiological concentration of each nucleotide (100 μM), hPolη incorporated correct dCTP with a probability of 59% opposite the C8-dG-ABA lesion (Figure 1A) after 10 s of incubation (Figure 1B, left). The probability is defined as the

[a] Dr. A. Patra,[+] Prof. M. Egli
Department of Biochemistry, Vanderbilt University, School of Medicine
868A Robinson Research Building, Nashville, TN 37232 (USA)
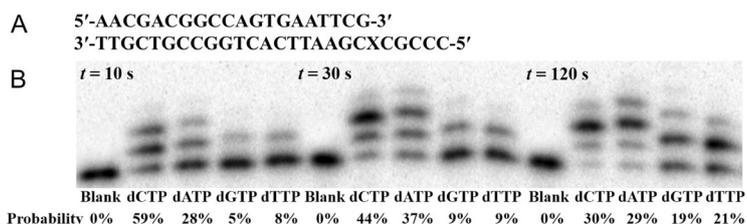E-mail: martin.egli@vanderbilt.edu

[b] Dr. D. A. Politica,[+] Prof. M. P. Stone
Department of Chemistry, Vanderbilt University, College of Arts & Science
Station B, Box 1822, Nashville, TN 37235 (USA)
E-mail: michael.p.stone@vanderbilt.edu

[c] A. Chatterjee, Prof. A. K. Basu
Department of Chemistry, University of Connecticut
55 North Eagleville Rd, Storrs, CT 06269 (USA)

[d] E. J. Tokarsky, Prof. Z. Suo
Department of Chemistry and Biochemistry, Ohio State University
484 West 12th Avenue, Columbus, OH 43210 (USA)

[+] These authors contributed equally to this work.

Supporting information for this article can be found under http://dx.doi.org/10.1002/cbic.201600420.

A

5′-AACGACGGCCAGTGAATTCG-3′
3′-TTGCTGCCGGTCACTTAAGCXCGCCC-5′

B

	$t = 10$ s				$t = 30$ s				$t = 120$ s			
Blank	dCTP	dATP	dGTP	dTTP	Blank	dCTP	dATP	dGTP	dTTP	Blank	dCTP	dATP	dGTP	dTTP
Probability 0%	59%	28%	5%	8%	0%	44%	37%	9%	9%	0%	30%	29%	19%	21%

Figure 1. Single-nucleotide incorporation opposite a C8-dG-ABA lesion catalyzed by hPolη at 37 °C. A) 20/26-mer C8-dG-ABA; the adducted dG is marked by an X. B) A pre-incubated solution of hPolη (130 nM) and ^{32}P-labeled 20/26-mer C8-dG-ABA DNA substrate (30 nM) was mixed with the indicated dNTP (100 μM) for 10 s, 30 s, or 120 s, then the reaction was quenched with 0.37 M EDTA. Product concentration ([P]) was quantified by calculating the intensity ratio of the P band(s) over the sum of the P band(s) and the band of remaining 20-mer, and multiplying by the total DNA concentration (30 nM). The probability for dNTP incorporation opposite C8-dG-ABA was calculated as ([P] from dNTP incorporation/Σ[P] from each dNTP incorporation)×100%. "Blank" indicates no addition of dNTPs.

percentage of the incorporation of a specific nucleotide among all four possible nucleotide incorporations opposite a template nucleotide. Opposite the lesion, there was also incorporation of incorrect nucleotides following the pattern: dATP (28%) > dTTP (9%) > dGTP (5%). At longer reaction time points (30 s and 120 s, Figure 1 B), the probability of a misincorporation event was greater because correct dCTP incorporation was almost completed whereas more misincorporation occurred.

Using pre-steady-state kinetics to measure the rates of nucleotide incorporation by hPolη onto the 20-/26-mer C8-dG-ABA DNA substrate (Figure 1 A),[27] we calculated the observed rate constants (k_{obs}) for each dNTP at 100 μM by using the hyperbolic equation ($k_{obs} = k_p + [dNTP]/(K_d + [dNTP])$), where k_p is the maximum dNTP incorporation rate constant, and K_d is the apparent equilibrium dissociation constant of the dNTP. For example, to calculate the estimated probability for correct incorporation of dCTP at 100 μM, we took the k_{obs} for dCTP and divided it by the sum of k_{obs} values obtained for all dNTPs on the same substrate ([$k_{obs,dCTP}/\Sigma k_{obs,dNTP}$]×100%). Based on our calculation, the probability for hPolη to correctly insert dCTP was 81%, higher than the 59% observed after the 10 s reaction time in the single-nucleotide incorporation assay (Figure 1 B). This suggests that 10 s was still too long to accurately estimate the correct nucleotide incorporation probability opposite the lesion. However, the kinetic probabilities of hPolη incorporating incorrect dNTPs opposite C8-dG-ABA follow the pattern: dATP (13%) > dTTP (4%) > dGTP (2%), which is similar to the one depicted in Figure 1 B.

To gain a better understanding of the mechanism of error-free bypass of the C8-dG-ABA adduct by hPolη, we determined the crystal structure of a ternary Pol·DNA·dCTP complex trapped at the insertion stage. The hPolη catalytic core encompassing residues 1–432 was expressed in *Escherichia coli* and purified following published protocols[28] (see the Supporting Information for materials and methods). The DNA duplex is composed of a 12-mer template with the C8-dG-ABA lesion (X) 5′-CATXATGACGCT-3′ and an octamer primer 5′-AGCGTCAT-3′. Crystals of the complex with dCTP were grown in the presence

of Ca^{2+} instead of Mg^{2+} to prevent incorporation of the incoming nucleoside triphosphate into the primer strand. The structure was phased with molecular replacement by using hPolη as the search model and refined to a resolution of 2.6 Å.

A summary of crystal data, data collection, and refinement parameters is given in Table 1, and an example of the electron density around the final model of the complex is depicted in Figure 2. The final model allowed visualization of ten of the template nucleotides (the 5′-terminal C and A outside the active site are disordered) and the entire primer strand. Perhaps surprisingly, the C8-dG-ABA adducted nucleotide adopts the *anti* conformation at the active site of hPolη (Figures 3 and S1). The conformation is stabilized by a hydrogen bond between the ABA

Table 1. Selected crystal data, data collection parameters, and structure refinement statistics.

Data Collection	
space group	$P6_1$
resolution [Å]	50.0–2.60 (2.64–2.60)
unit cell $a = b$, c [Å]	99.11, 81.69
completeness [%]	100 (100)
$I/\sigma(I)$	15.3 (2.0)
R merge [%]	13.3 (88.8)
redundancy	7.6 (7.5)
Refinement	
R work [%]	16.6 (21.4)
R free [%][b]	23.8 (28.1)
Avg. B [Å²]	39.0
r.m.s.d. bonds [Å]	0.009
r.m.s.d. angles [°]	1.1

[a] Statistics for the highest-resolution shell are shown in parentheses. [b] Based on 5% of the reflections.
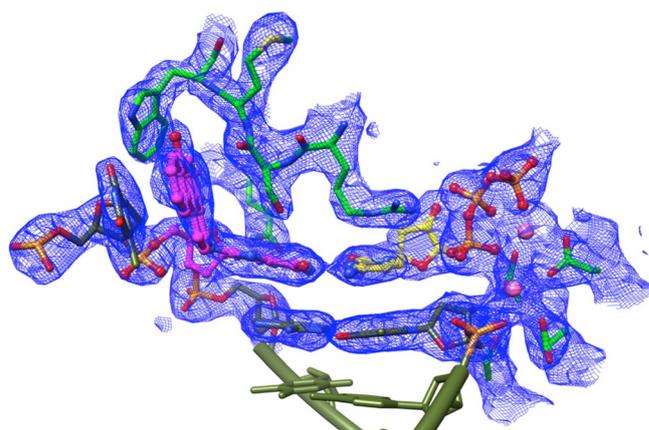
Figure 2. Quality of the final Fourier $2F_o - F_c$ sum electron density (1σ threshold) around the active site region of the ternary complex between hPolη, C8-dG-ABA adducted template:primer duplex and dCTP. Carbon atoms of hPolη, C8-dG-ABA, and dCTP are colored light green, magenta, and yellow, respectively, Ca^{2+} ions are shown as pink spheres, and the remainder of the duplex is colored in olive.
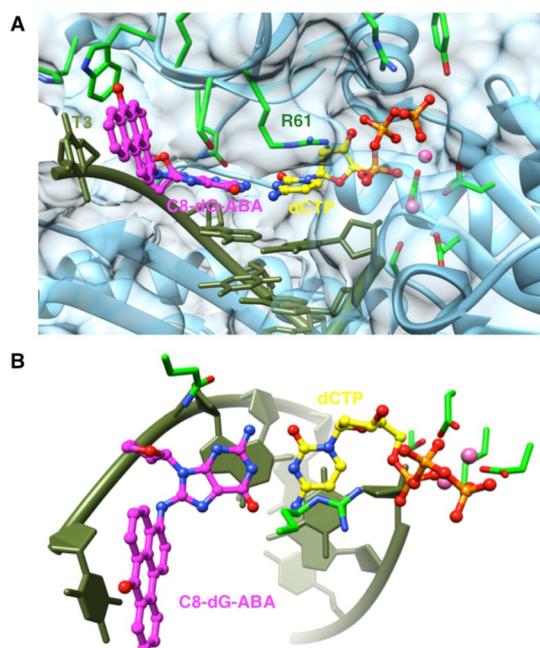
Figure 3. Active site configuration in the ternary hPolη insertion-step complex with dCTP opposite C8-dG-ABA. Views into A) the DNA major groove and B) rotated by ~90° around the horizontal axis, looking approximately along the normal to the nucleobase plane of the incoming dCTP. Carbon atoms of C8-dG-ABA, incoming dCTP, and selected hPolη side chains are colored in magenta, yellow, and light green, respectively, and Ca²⁺ ions are shown as pink spheres.

amino moiety and the phosphate group, with the C8-dG-ABA 2′-deoxyribose adopting a C3′-endo pucker.

As a consequence, the adducted G forms a standard Watson–Crick base pair with the incoming dCTP. The ABA adduct is tilted relative to the guanine plane by about 80° and rotated outwards into the major groove of the template:primer duplex.

However, rather than exhibiting the enhanced conformational flexibility that is often associated with portions of a structure that are extruded from the duplex, the structure offers evidence that the polyaromatic hydrocarbon moiety assumes a configuration that is stabilized by multiple interactions. Thus, ABA forms a sandwich with the amide group of Ser62 and the T3 template residue that is lodged outside the active site. The carbonyl oxygen of Ser62 engages in an n→π* interaction with the ABA keto group, and Cβ-H and C2′-H from Ser62 and T3, respectively, engage in C-H→π interactions with aromatic rings from opposite sides of the ABA plane (Figure 3). Trp64 seals the backside of this mostly hydrophobic pocket. Both the keto oxygens of ABA and T3 (O2) are too far removed from Trp64 N-H (distances of 3.71 and 3.83 Å, respectively; Figure S1) to engage in hydrogen bonds.

The extensive interactions observed for ABA at the hPolη active site extend to the base, sugar, and phosphate portions of the adducted dG (Figures 3 and S1). On the minor groove side, the NH₂ group of Gln38 from the finger domain forms hydrogen bonds with O4′ of the 2′-deoxyribose and N3 of the base. The Cγ2(H)₃ methyl group of Ile48 stacks onto the six-

membered ring of guanine. In the major groove, the OH group of Ser61 is positioned somewhat far away to establish hydrogen bonds with the O6 and N7 acceptor functionalities of guanine (distances of 3.80 and 3.75 Å, respectively). Finally, the OH group of Tyr39 stabilizes the template strand conformation with a hydrogen bond to the non-bridging OP1 phosphate oxygen of C8-dG-ABA. On the opposite side of the active site, the guanidino moiety of Arg61 from the finger domain stacks on top of the nucleobase of dCTP. The NH₂ moiety of the arginine forms salt bridges with the α- and β-phosphate groups and helps align the former for attack by the 3′-OH of the terminal primer nucleotide (distance O3′···Pα = 3.85 Å), together with the two divalent metal ions.

The tight fit between the nascent C8-dG-ABA:dCTP base pair and hPolη is further illustrated in Figure 4, which depicts a space-filling model of the active site of the polymerase. Finger residues Ser62 and Arg61 (foreground, center) protrude from the ceiling, along with Gln38, and bear down on the G:C base-pair plane at the split between the primer and template strands.
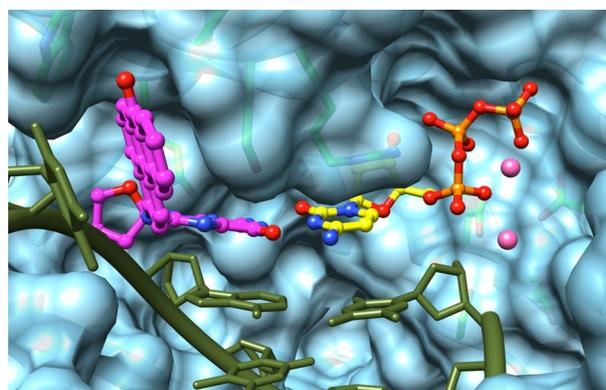
Figure 4. Space-filling model of the hPolη active site with C8-dG-ABA (magenta carbon atoms) accommodated inside a cleft to the side of the active site that allows for extensive hydrophobic interactions between the ABA moiety and surrounding residues from the polymerase finger domain. Carbon atoms of the incoming dCTP are colored yellow.

Two metal ions are accommodated in a pocket on one side of the active site by conserved Glu and Asp residues. On the other side, where the template strand enters the active site, the ABA adduct fits snugly into a cleft formed by Ser62, Met63 (main chain), and Trp64 and extended by T3 from the template strand. The structure thus offers insight into how hPolη stabilizes a configuration of the template strand that allows error-free bypass of the bulky C8-dG-ABA lesion.

We recently reported the NMR solution structure of a DNA 12-mer duplex containing a C8-dG-ABA nucleotide opposite dC.[29] In this structure, ABA was intercalated in the duplex, with the adducted G in the syn conformation and extruded into the major groove. The intercalated adduct thus takes the place of a base pair and thereby forces the opposing cytosine into the major groove (Figure S2). We superimposed this DNA duplex model with base pairs on the 5′-side of the adducted dG over the template:primer duplex of the ternary hPolη com-
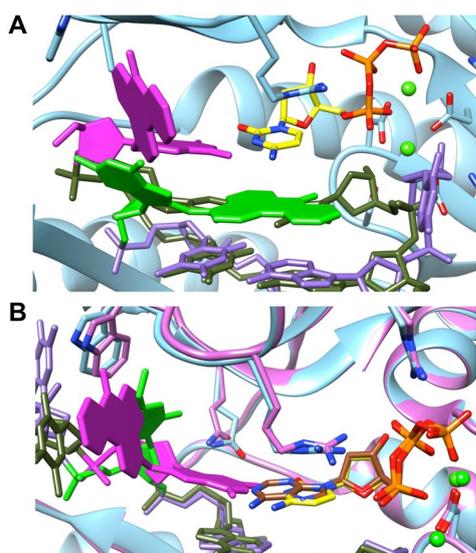
Figure 5. Model of error-prone bypass of C8-dG-ABA and comparison of hPolη complexes with the C8-dG-ABA and O^4-MedT adducts. A) Superimposition of the crystal structure of the hPolη (Pol ribbon in cyan) complex with C8-dG-ABA (magenta/olive) and dCTP (yellow), and the NMR structure of DNA (lilac) with the C8-dG-ABA (light green) lesion.[29] B) Superimposition of the hPolη complexes with C8-dG-ABA (magenta/olive)/dCTP (yellow; Pol ribbon in cyan) and O^4-MedT (light green/lilac)/dATP (brown; Pol ribbon in pink).[30]

plex, such that the intercalated ABA assumed the position of the T:A pair at the −1 position of the Pol active site (Figure 5 A). This configuration illustrates a potential basis for error-prone bypass of C8-dG-ABA in that incoming dATP or dGTP might be preferred due to favorable stacking. Insertion could then proceed in a non-templated fashion, thereby causing a mutation, or opposite a residue located 5′-adjacent to C8-dG-ABA, thus causing a frameshift.

Drastically different conformations and interactions of aromatic moieties, depending on the environment—that is, inside a DNA duplex studied alone or contained in a template–primer strand at the active site of a DNA polymerase—have previously been observed. For example, the dNaM and d5SICS components of an unnatural base pair that lack the ability to establish hydrogen bonds are stacked in the structure of an isolated DNA duplex. However, at the active site of KlenTaq DNA polymerase, dNaM in the template strand and incoming d(5SICS)TP form a side-by-side Watson–Crick-like base pair.[31]

hPolη inserts dCTP ca. 3.5-fold more efficiently than dATP opposite another C8 adduct of dG, 8-oxoG.[32] Incorporation of dATP occurs with the adducted dG in the *syn* orientation and stacked inside the template strand. However, in comparison to the C8=O8 moiety of 8-oxoG, the ABA adduct is far too bulky to be accommodated within the minor groove.

The enzyme from yeast (scPolη) is known to bypass bulky acetylaminofluorene DNA adducts, such as *N*-2′-(deoxyguanosin-8-yl)-2-acetylaminofluorene (AAF-dG). But the presence of the acetyl moiety precludes formation of a hydrogen bond between the amino group and the 5′-phosphate, and thus, the *anti* conformation. Hence, bypass of the lesion is error-prone,

with insertion of all four dNTPs opposite AAF-dG.[33] Consistent with these observations at the functional level, crystal structures of the complex between scPolη and a DNA duplex that contains AAF-dG show aminofluorene stacked on the adjacent template:primer duplex in one case, with dG rotated into the major groove, and dG rotated toward the templating position and partially freeing the active site in a second case (Figure 6 A).[33] Neither structure contained an incoming dNTP. Crys-
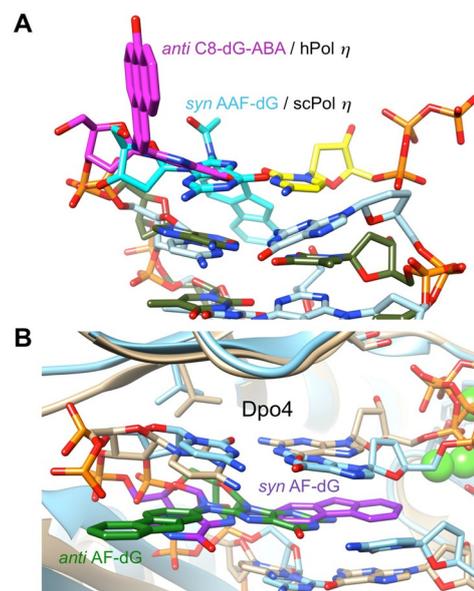
Figure 6. Conformations of dG C8-aromatic amine adducts at the active sites of TLS DNA polymerases. A) Superimposition of the template:primer duplexes with *anti*-C8-dG-ABA (hPolη, this work) and *syn*-AAF-dG in complex with scPolη (PDB ID: 2XGP, complex B).[33] The DNA conformations deviate considerably (r.m.s.d. sugar–phosphate backbone atoms > 2 Å), and the AAF moiety in the scPolη complex is lodged in the minor groove, which is impossible with the human enzyme. B) Superimposition of post-insertion complexes of Dpo4 with *syn*- and *anti*-AF-dG (PDB IDs: 3KHG and 3KHH, respectively).[34] The AF moiety in the latter is lodged in the major groove. Neither the complexes of AAF-dG with scPolη nor those of Dpo4 with AF-dG constitute good models for the bypass of C8-dG-ABA by hPolη.

tal structures of ternary complexes between the Y-family TLS DNA polymerase Dpo4 from *Sulfolobus solfataricus* and AF-dG at the −1-position (post-replicative, extension state) revealed both *syn* and *anti* orientations of the adducted nucleotide.[34] Thus, in the former structure, the AF moiety is stacked between the base pair at the −2 position and the nascent dC:dGTP pair, with the primer residue opposite AF-dG disordered in the model (Figure 6 B). In the structure with the *anti* orientation of AF-dG, the adducted nucleotide forms a Watson–Crick pair with primer dC, followed by the nascent dC:dGTP pair (Figure 6 B).[34]

Incorporation of AF-dG (but not of AAF-dG), albeit at a much reduced rate compared to that of native nucleotides, has also been observed for a replicative, high-fidelity DNA polymerase, the thermophilic *Bacillus* DNA polymerase I fragment (BF). The crystal structure of the BF:DNA complex at the pre-inserton state shows the adducted residue in a *syn* conformation, and in the *anti* conformation in the post-insertion

state and forming a Watson–Crick pair with dC, with the AF moiety protruding into the major groove.[35]

In summary, we present the first crystal structure of hPol η in complex with a bulky C8-dG aromatic amine lesion. The structure provides a model for the correct (Figure 1) but slow bypass of the C8-dG-ABA adduct.[26] Finally, the C8-dG-ABA:h-Pol η complex constitutes the second example of an insertion-stage bypass structure, with an adduct lodged in the hydrophobic cleft to the side of the active site. A somewhat similar orientation was found for O^4-MedT (Figure 5 B) that is replicated by hPol η in an error-prone fashion.[30] However, unlike in the case of C8-dG-ABA, in which only the ABA portion is inserted into the cleft, the entire O^4-MedT nucleobase can be accommodated inside it, with the lesion reaching all the way to the back of the pocket.

Experimental Section

See the Supporting Information for details.

Acknowledgements

[1] J. Kielhorn, U. Wahnschaffe, I. Mangelsdorf, *WHO Environmental Health Criteria Report*, **2003**, http://www.inchem.org/documents/ehc/ehc/ehc229.htm.

[2] T. Murahashi, *Analyst* **2003**, *128*, 42.

[3] D. Schuetzle, F. S.-C. Lee, T. J. Prater, S. B. Tejada, *Int. J. Environ. Anal. Chem.* **1981**, *9*, 93–144.

[4] M. Paputa-Peck, R. Marano, D. Schuetzle, T. Riley, C. Hampton, T. Prater, L. Skewes, T. Jensen, P. Ruehle, *Anal. Chem.* **1983**, *55*, 1946–1954.

[5] T. Murahashi, E. Iwanaga, T. Watanabe, T. Hirayama, *J. Health Sci.* **2003**, *49*, 386–390.

[6] L. Benbrahim-Tallaa, R. Baan, Y. Grosse, B. Lauby-Secretan, F. El Ghissassi, V. Bouvard, N. Guha, D. Loomis, K. Straif, *Lancet Oncol.* **2012**, *13*, 663–664.

[7] T. Enya, H. Suzuki, T. Watanabe, T. Hirayama, Y. Hisamatsu, *Environ. Sci. Technol.* **1997**, *31*, 2772–2776.

[8] M. Kawanishi, Y. Fujikawa, H. Ishii, H. Nishida, Y. Higashigaki, T. Kanno, T. Matsuda, T. Takamura-Enya, T. Yagi, *Mutat. Res.* **2013**, *753*, 93–100.

[9] V. Purohit, A. K. Basu, *Chem. Res. Toxicol.* **2000**, *13*, 673–692.

[10] S. Adachi, K. Kawanura, K. Takemoto, H. Suzuki, Y. Hisamatsu, *Relationships between Respiratory Disease and Exposure to Air Pollution*, ILSI Press, Washington, DC, **2000**, pp. 315–319.

[11] V. M. Arlt, J. Gingerich, H. H. Schmeiser, D. H. Phillips, G. R. Douglas, P. A. White, *Mutagenesis* **2008**, *23*, 483–490.

[12] N. Landvik, V. Arlt, E. Nagy, A. Solhaug, X. Tekpli, H. Schmeiser, M. Refsnes, D. Phillips, D. Lagadic-Gossmann, J. Holme, *Mutat. Res.* **2010**, *684*, 11–23.

[13] P. T. Phousongphouang, A. J. Grosovsky, D. A. Eastmond, M. Covarrubias, J. Arey, *Mutat. Res.* **2000**, *472*, 93–103.

[14] E. Lamy, F. Kassie, R. Gminski, H. H. Schmeiser, V. Mersch-Sundermann, *Toxicol. Lett.* **2004**, *146*, 103–109.

[15] T. Hansen, A. Seidel, J. Borlak, *Toxicol. Appl. Pharmacol.* **2007**, *221*, 222–234.

[16] J. Borlak, T. Hansen, Z.-X. Yuan, H. C. Sikka, S. Kumar, S. Schmidbauer, H. Frank, J. Jacob, A. Seidel, *Polycyclic Aromat. Compd.* **2000**, *21*, 73–86.

[17] A. Seidel, D. Dahmann, H. Krekeler, J. Jacob, *Int. J. Hyg. Environ. Health* **2002**, *204*, 333–338.

[18] V. M. Arlt, D. H. Phillips, J. Reynisson, *Org. Biomol. Chem.* **2011**, *9*, 6100–6110.

[19] Z.-Z. Yang, S.-F. Qi, D.-X. Zhao, L.-D. Gong, *J. Phys. Chem. B* **2009**, *113*, 254–259.

[20] J. Xue, L. Du, R. Zhu, J. Huang, D. L. Phillips, *J. Org. Chem.* **2014**, *79*, 3610–3614.

[21] V. M. Arlt, H. H. Schmeiser, M. R. Osborne, M. Kawanishi, T. Kanno, T. Yagi, D. H. Phillips, T. Takamura-Enya, *Int. J. Cancer* **2006**, *118*, 2139–2146.

[22] V. M. Arlt, B. L. Sorg, M. Osborne, A. Hewer, A. Seidel, H. H. Schmeiser, D. H. Phillips, *Biochem. Biophys. Res. Commun.* **2003**, *300*, 107–114.

[23] S. Prakash, R. E. Johnson, L. Prakash, *Annu. Rev. Biochem.* **2005**, *74*, 317–353.

[24] J. E. Sale, A. R. Lehmann, R. Woodgate, *Nat. Rev. Mol. Cell. Biol.* **2012**, *13*, 141–152.

[25] A. R. Lehmann, A. Niimi, T. Ogi, S. Brown, S. Sabbioneda, J. F. Wing, P. L. Kannouche, C. M. Green, *DNA Repair* **2007**, *6*, 891–899.

[26] P. Pande, C. K. Malik, A. Bose, V. P. Jasti, A. K. Basu, *Biochemistry* **2014**, *53*, 5323–5331.

[27] E. J. Tokarsky, V. V. Gadkari, W. J. Zahurancik, C. K. Malik, A. K. Basu, Z. Suo, *DNA Repair* **2016**; DOI: 10.1016/j.dnarep.2016.08.002.

[28] C. Biertümpfel, Y. Zhao, Y. Kondo, S. Ramón-Maiques, M. Gregory, J. Y. Lee, C. Masutani, A. R. Lehmann, F. Hanaoka, W. Yang, *Nature* **2010**, *465*, 1044–1048.

[29] D. A. Politica, C. K. Malik, A. K. Basu, M. P. Stone, *Chem. Res. Toxicol.* **2015**, *28*, 2253–2266.

[30] D. K. O'Flaherty, A. Patra, Y. Su, F. P. Guengerich, M. Egli, C. J. Wilds, *Chem. Sci.* **2016**, *7*, 4896–4904.

[31] K. Betz, D. A. Malyshev, T. Lavergne, W. Welte, K. Diederichs, T. J. Dwyer, P. Ordoukhanian, F. E. Romesberg, A. Marx, *Nat. Chem. Biol.* **2012**, *8*, 612–614.

[32] A. Patra, L. D. Nagy, Q. Zhang, Y. Su, L. Muller, F. P. Guengerich, M. Egli, *J. Biol. Chem.* **2014**, *289*, 16867–16882.

[33] S. Schorr, S. Schneider, K. Lammens, K.-P. Hopfner, T. Carell, *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 20720–20725.

[34] O. Rechkoblit, A. Kolbanovskiy, L. Malinina, N. E. Geacintov, S. Broyde, D. J. Patel, *Nat. Struct. Mol. Biol.* **2010**, *17*, 379–388.

[35] G. W. Hsu, J. R. Kiefer, D. Burnouf, O. J. Becherel, R. P. P. Fuchs, L. S. Beese, *J. Biol. Chem.* **2004**, *279*, 50280–50285.

CHEMBIOCHEM

Supporting Information

Mechanism of Error-Free Bypass of the Environmental Carcinogen *N*-(2′-Deoxyguanosin-8-yl)-3-aminobenzanthrone Adduct by Human DNA Polymerase η

Amritraj Patra[+],[a] Dustin A. Politica[+],[b] Arindom Chatterjee,[c] E. John Tokarsky,[d] Zucai Suo,[d] Ashis K. Basu,[c] Michael P. Stone,*[b] and Martin Egli*[a]

Contents

EXPERIMENTAL PROCEDURES

hPol η Catalytic Core Protein Expression and Purification. The hPol η plasmid (pET28a) comprising residues 1-432 was a generous gift from Dr. Wei Wang, NIDDK, NIH. The polymerase was expressed in *E. coli* and purified as previously described.[1-9] The protein solution was concentrated to 5 mg/mL.

Oligonucleotide Synthesis and Annealing. C8-dG-ABA modified 12mer template 3´- TCG CAG TAX TAC -5´ (X= C8-dG-ABA) and unmodified 8mer primer 5´-AGC GTC AT-3´ were used in the crystallization experiments. The synthesis of site-specifically modified oligonucleotides containing the C8-dG-ABA lesion was previously reported.[10] All modified oligodeoxynucleotides were characterized by MALDI-TOF mass spectrometry. Capillary gel electrophoresis and C-18 HPLC confirmed their purities. Unmodified DNA primer was purchased from Integrated DNA Technologies (Coralville, IA). Template and primer strands were mixed in a 1:1 molar ratio and were annealed in the presence of 10 mM sodium HEPES buffer (pH 8.0), 0.1 mM EDTA, and 50 mM NaCl by heating for 10 min. at 85°C, followed by slow cooling to room temperature.

Crystallization of the hPol η•C8-dG-ABA DNA•dCTP Insertion Complex. The DNA template-primer duplex was mixed with the protein in a 1.2:1 molar ratio in the presence of excess 50 mM Tris-HCl, pH 7.5, 450 mM KCl, and 3 mM DTT. After adding 5 μL of 100 mM $CaCl_2$ the complex was concentrated to a final concentration of ~2-3 mg/mL by ultrafiltration. dCTP was added last to form the ternary complex to a final concentration of 10 mM. Crystallization experiments were performed by the hanging drop vapor diffusion technique at 18°C using a sparse matrix screen (Hampton Research, Aliso Viejo, CA).[11] One μL of the complex solution was mixed with 1 μL of reservoir solution and equilibrated against 500 μL reservoir wells. Crystals appeared in droplets containing 0.1 M MES (pH 5.5), 5 mM $CaCl_2$, and 25% (w/v) PEG 2000 MME within one day and were harvested after a week. Crystals were mounted in nylon loops, cryo-protected in reservoir solution containing 25% glycerol (v/v), and frozen in liquid nitrogen.

X-ray Diffraction Data Collection, Structure Determination and Refinement.
Diffraction data were collected on the 21-ID-D beamline of the Life Sciences Collaborative Access Team (LS-CAT) at the Advanced Photon Source, Argonne National Laboratory (Argonne, IL). All data were integrated and scaled with the program HKL2000.[12] The structures were determined by the Molecular Replacement technique with the program MOLREP.[13,14] For the complex of the hPol η structure PDB entry 4O3N (protein only)[3] was used as the search model for molecular replacement. Structure refinement and model building were carried out with Phenix[15] and COOT,[16] respectively. Illustrations were prepared with the program UCSF Chimera.[17]

Table S1. Expanded version of Table 1 in the main paper: Crystal data, data collection parameters, and structure refinement statistics[a]

Data Collection	
Wavelength [Å]	0.97856
Space group	$P6_1$
Resolution [Å]	50.0 - 2.60 (2.64 - 2.60)[a]
Unit cell *a, b, c* [Å]	99.11, 99.11, 81.69
Unique reflections	14,197 (702)
Completeness [%]	100 (100)
I/σ(I)	15.3 (2.0)
Wilson B-factor [Å^2]	39.0
R-merge	0.133 (0.888)
Redundancy	7.6 (7.5)
Refinement	
R-work	0.166 (0.215)
R-free	0.238 (0.281)
Number of atoms	
Protein/DNA	3,377/378
dCTP/water/ Ca^{2+}	28/145/2
Protein residues	430
B-factor [Å^2]	
Average	39.0
Protein/DNA	39.3/38.8
dNTP/M^{2+} /water	29.0/37.6/39.1
R.m.s. deviations	
bonds [Å]	0.009
angles [deg.]	1.1
Ramachandran	
Favored (%)	95
Allowed (%)	4.5
Outliers (%)	0.5
PDB ID Code	5JUM

[a] Statistics for the highest-resolution shell are shown in parentheses.
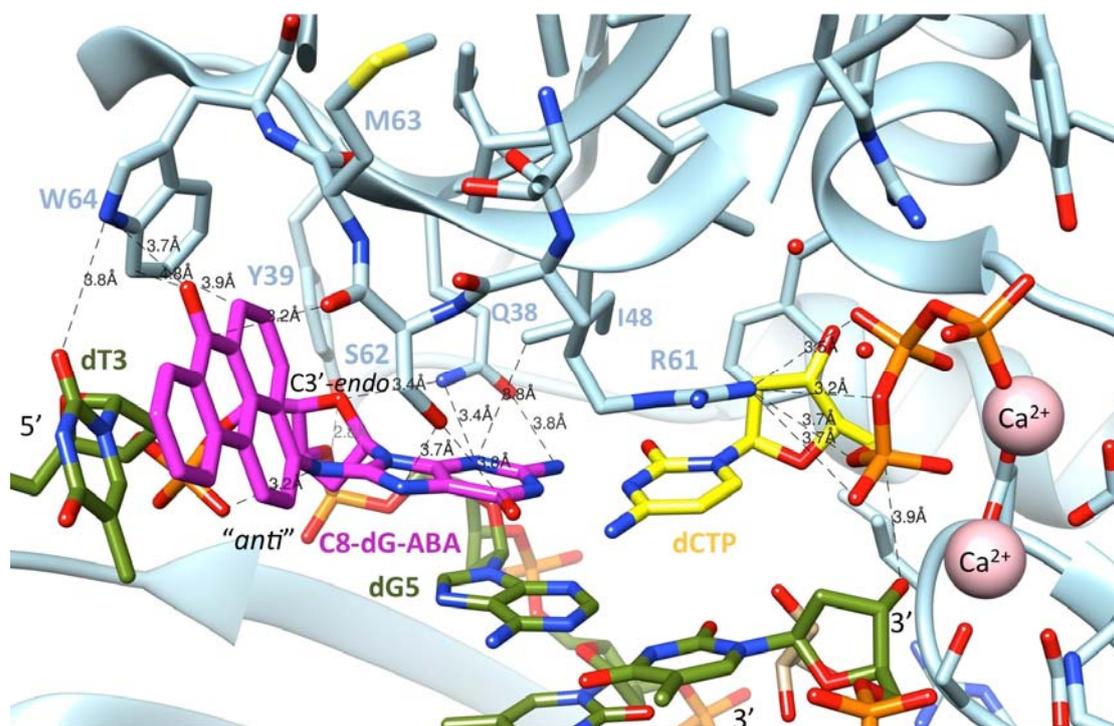
Figure S1: Close-up view of the active site in the crystal structure of the hPol η•C8-dG-ABA DNA•dCTP insertion-step complex. The color code matches that in Figure 2 (main paper) and selected distances in Å are indicated with dashed lines.
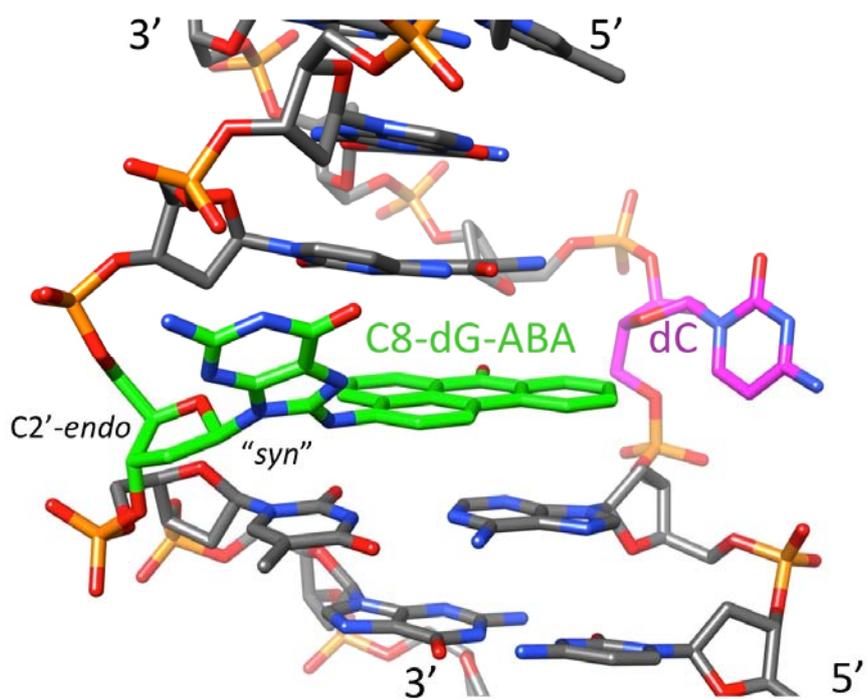
Figure S2: NMR solution structure of C8-dG-ABA inside B-form DNA.[18]

ACKNOWLEDGEMENTS

REFERENCES

[1] C. Biertümpfel, Y. Zhao, Y. Kondo, S. Ramon-Maiques, M. Gregory, J. Y. Lee, C. Masutani, A. R. Lehmann, F. Hanaoka, W. Yang, *Nature* **2010**, *465*, 1044-1048.

[2] Y. Zhao, C. Biertümpfel, M. T. Gregory, Y.-J. Hua, F. Hanaoka, W. Yang, *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, 7269-7274.

[3] A. Patra, L. D. Nagy, Q. Zhang, Y. Su, L. Müller, F. P. Guengerich, M. Egli, *J. Biol. Chem.* **2014**, *289*, 16867-16882.

[4] A. Patra, Q. Zhang, L. Lei, Y. Su, M. Egli, F. Peter Guengerich, *J. Biol. Chem.* **2015**, *290*, 8028-8038.

[5] Y. Su, A. Patra, J. Harp, M. Egli, F. Peter Guengerich, *J. Biol. Chem.* **2015**, *290*, 15921-15933.

[6] A. Patra, S. Banerjee, T. Johnson Salyard, C. Malik, P. Christov, C. J. Rizzo, M. P. Stone, M. Egli, *J. Am. Chem. Soc.* **2015**, *137,* 7011-7014.

[7] Y. Su, M. Egli, F. P. Guengerich, *J. Biol. Chem.* **2016**, *291*, 3747-3756.

[8] D. K. O'Flaherty, A. Patra, Y. Su, F. P. Guengerich, M. Egli, C. J. Wilds, *Chemical Science* **2016**, *7*, 4896-4904.

[9] A. Patra, Y. Su, Q. Zhang, K. M. Johnson, F. P. Guengerich, M. Egli, *J. Biol. Chem.* **2016**, *291*, 14134-14145.

[10] V. V. Gadkari, E. J. Tokarsky, C. K. Malik, A. K. Basu, Z. C. Suo, *DNA Repair* **2014**, *21*, 65-77.

[11] J. Jancarik, S.-H. Kim, *J. Appl. Cryst.* **1991**, *24*, 409-411.

[12] Z. Otwinowski, W. Minor, *Methods Enzymol.* **1997**, *276*, 307-326.

[13] A. Vagin, A. Teplyakov, *J. Appl. Cryst.* **1997**, *30*, 1022-1025.

[14] The CCP4 suite: programs for protein crystallography. *Acta Cryst. Section D, Biol. Cryst.* **1994**, *50*, 760-763.

[15] P. D. Adams, P. V. Afonine, G. Bunkoczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L.W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, J. S. Richardson, T. C. Terwilliger, P. H. Zwart, *Acta Cryst. D* **2010**, *66*, 213-221.

[16] P. Emsley, K. Cowtan, *Acta Cryst. Section D, Biol. Cryst.* **2004**, *60*, 2126-2132.

[17] E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin, *J. Comput. Chem.* **2004**, *25*, 1605-1612.

[18] D. A. Politica, C. K. Malik, A. K. Basu, M. P. Stone, *Chem. Res. Toxicol.* **2015**, *28*, 2253–2266.