

Structural and Kinetic Analysis of Nucleoside Triphosphate Incorporation Opposite an Abasic Site by Human Translesion DNA Polymerase η *

Received for publication, January 9, 2015, and in revised form, February 6, 2015. Published, JBC Papers in Press, February 9, 2015, DOI 10.1074/jbc.M115.637561

Amritaj Patra¹, Qianqian Zhang¹, Li Lei, Yan Su, Martin Egli², and F. Peter Guengerich³

From the Department of Biochemistry and Center in Molecular Toxicology, Vanderbilt University School of Medicine, Nashville, Tennessee 37232-0146

Background: Abasic sites are the most common lesion in DNA.

Results: Kinetic and mass spectrometric assays demonstrate that human polymerase (pol) η preferentially inserts A and G opposite an abasic site.

Conclusion: Crystal structures reveal H-bonding between incoming ATP and GTP and the 5'-phosphate of the abasic moiety.

Significance: Abasic site bypass by pol η follows a "purine rule" for insertion, with formation of frameshifts.

The most common lesion in DNA is an abasic site resulting from glycolytic cleavage of a base. In a number of cellular studies, abasic sites preferentially code for dATP insertion (the "A rule"). In some cases frameshifts are also common. X-ray structures with abasic sites in oligonucleotides have been reported for several microbial and human DNA polymerases (pols), e.g. Dpo4, RB69, KlenTaq, yeast pol ι , human (h) pol ι , and human pol β . We reported previously that hpol η is a major pol involved in abasic site bypass (Choi, J.-Y., Lim, S., Kim, E. J., Jo, A., and Guengerich, F. P. (2010) *J. Mol. Biol.* 404, 34–44). hpol η inserted all four dNTPs in steady-state and pre-steady-state assays, preferentially inserting A and G. In LC-MS analysis of primer-template pairs, A and G were inserted but little C or T was inserted. Frameshifts were observed when an appropriate pyrimidine was positioned 5' to the abasic site in the template. In x-ray structures of hpol η with a non-hydrolyzable analog of dATP or dGTP opposite an abasic site, H-bonding was observed between the phosphate 5' to the abasic site and water H-bonded to N1 and N6 of A and N1 and O6 of G nucleoside triphosphate analogs, offering an explanation for what appears to be a "purine rule." A structure was also obtained for an A inserted and bonded in the primer opposite the abasic site, but it did not pair with a 5' T in the template. We conclude that hpol η , a major copying enzyme with abasic sites, follows a purine rule, which can also lead to frameshifts. The phenomenon can be explained with H-bonds.

Depurination of DNA is a common event, occurring at a rate of $\sim 50,000$ abasic sites/cell/day (1). These abasic (apurinic/aprimidinic (AP)⁴) sites are by far the most frequent type of DNA damage, in that they are very blocking and, if processed by DNA polymerases, miscoding (2). In the early 1980s, three different laboratories reported a tendency of AP sites to cause incorporation of dATP, called the "A rule" (3–5). However, not all DNA polymerases stringently follow the A rule. Human (h) pol β and *Sulfolobus solfataricus* Dpo4 have been reported to prefer to produce -1 frameshift deletions to a greater extent than A insertion (6, 7). Sagher and Strauss (8) reported that hpol β preferred to insert G opposite AP sites derived from C.

Although the A rule was reported over 30 years ago, there is still no general consensus for the mechanism (9). Base stacking with the 3' base (at the end of the primer strand) has been mentioned (10), but this would not explain why different pols vary. A number of x-ray structures of pols with oligonucleotides with AP sites have been published, including *S. solfataricus* Dpo4 (11), hpol β (12), yeast Rev1 (13), hpol ι (14), and RB69 (15, 16). In some cases an amino acid of the pol is inserted into space available due to the missing base (9, 12), which is consistent with slow rates of incorporation.

One of the issues is which pol(s) is most relevant in considering the bypass of AP sites. Although hpol η has been reported to be inefficient in the bypass of such sites (17), it was found to be the most efficient single pol (at an AP site) when compared with other human translesion synthesis pols (18). Yeast pol η was reported to be important in one study (19) but not in another (20). Kokoska *et al.* (21) reported that hpol η was highly miscoding and that the majority of bypass events were insertions of A.

* This work was supported, in whole or in part, by National Institutes of Health Grants R01 ES010375 (to F. P. G. and M. E.), P01 CA160032 (to M. E.), and P30 ES000267 (to F. P. G. and M. E.).

The atomic coordinates and structure factors (codes 4RNM, 4RNN, and 4RNO) have been deposited in the Protein Data Bank (<http://www.pdb.org/>).

¹ Both authors contributed equally to this work.

² To whom correspondence may be addressed: Dept. of Biochemistry, 868A Robison Research Bldg., 2200 Pierce Ave., Nashville, TN 37232-0146. Tel.: 615-343-8070; Fax: 615-322-7122; E-mail: martin.egli@vanderbilt.edu.

³ To whom correspondence may be addressed: Dept. of Biochemistry, Vanderbilt University School of Medicine, 638B Robison Research Bldg., 2200 Pierce Ave., Nashville, TN 37232-0146. Tel.: 615-322-2261; Fax: 615-322-4349; E-mail: f.guengerich@vanderbilt.edu.

⁴ The abbreviations used are: AP, abasic site; CID, collision-induced dissociation; dNMPNPP, 2'-deoxynucleoside-5'-[(α,β) -imido]triphosphate; dAMPNPP, 2'-deoxyadenosine-5'-[(α,β) -imido]triphosphate; dGMPNPP, 2'-deoxyguanosine-5'-[(α,β) -imido]triphosphate; EIC, extracted ion chromatogram; ESI, electrospray ionization; h, human; LC, liquid chromatography; MS, mass spectrometry; pol, DNA polymerase; THF, tetrahydrofuran; UPLC, ultraperformance liquid chromatography; PCNA, proliferating cell nuclear antigen.

We analyzed hpol η in terms of its miscoding properties, in light of the evidence that it may be an important factor in mutagenesis at AP sites. We found a proclivity of hpol η to insert both purines, A and G, opposite AP sites, which is reminiscent of an earlier reported behavior of hpol β (8). A structural basis was identified using x-ray crystallography, *i.e.* water-mediated hydrogen bonding between the 5'-phosphate of the AP site and the Watson-Crick face of purines. When an appropriate pyrimidine is present in the template 5' to the AP position, extensive -1 frameshifts occur.

EXPERIMENTAL PROCEDURES

Materials—Deoxyribonucleoside triphosphates (dNTPs), T4 polynucleotide kinase, and uracil DNA glycosylase were purchased from New England Biolabs (Ipswich, MA). A mixture of four dNTPs was purchased from Invitrogen. All nonhydrolyzable dNMPNPPs were obtained from Jena Bioscience (Jena, Germany). [γ - 32 P]ATP (specific activity 3000 Ci/mmol) was purchased from PerkinElmer Life Sciences. Biospin 6 columns were purchased from Bio-Rad. All unmodified oligonucleotides (HPLC purification) were obtained from Integrated DNA Technologies (Coralville, IA). Oligonucleotides containing an AP site or an AP analog (tetrahydrofuran (THF)) were synthesized by TriLink Biotechnologies (San Diego). The catalytic core (amino acids 1–432) of hpol η was expressed in *Escherichia coli* and purified as described previously (22, 23).

Steady-state Kinetics—Steady-state kinetic assays were performed as described previously (23–25). The oligonucleotides used in this study are listed in Table 1. Prior to kinetic assays, time course experiments were run to optimize reaction conditions. All polymerase reactions were carried out in 50 mM Tris-HCl buffer (pH 7.5) containing 5 mM MgCl₂, 10 mM dithiothreitol (DTT), 100 mM KCl, 5% glycerol (v/v), and 100 μ g/ml bovine serum albumin (BSA). The 5'-6-carboxyfluorescein-labeled primer-template (18-/23-mer) duplex (5 μ M) was extended using 4–40 nM concentrations of hpol η in the presence of various concentrations of a single dNTP (0–1 mM) at 37 °C for 5–20 min. Reactions were terminated with a quench solution containing 20 mM EDTA, 95% formamide (v/v), bromophenol blue, and xylene cyanol. Products were separated on 18% (w/v) polyacrylamide gels containing 7.5 M urea. Gels were scanned by a Typhoon Scanner (GE Healthcare) and analyzed by fluorescence intensity using ImageJ software (National Institutes of Health). The values of k_{cat} and K_m were determined using GraphPad Prism (La Jolla, CA).

Pre-steady-state Kinetics—Rapid quench experiments were performed using a model RQF-3 KinTek Quench flow apparatus (KinTek, Austin, TX). The 18-mer primer was 5'-end-labeled using [γ - 32 P]ATP and T4 polynucleotide kinase and annealed to a 23-mer template. Reactions were initiated by rapidly mixing 32 P-labeled primer-template/polymerase mixtures with an equal volume of dNTP-Mg²⁺ complex at 37 °C. The final concentrations of the reactants were as follows: 25 nM hpol η , 50 nM 32 P-labeled primer-template complex, 0.5 mM dNTP. Other reaction conditions were the same as described for steady-state kinetic assays. Reactions were quenched with 0.5 M EDTA at reaction times varying from 5 ms to 5 s. Products were separated using 18% (w/v) polyacrylamide gels. Bands on the

gels were visualized using a phosphorimaging system (Bio-Rad, Molecular Imager FX) and Quantity One software as previously described (26). Pre-steady-state points were fit (GraphPad Prism) to burst Equation 1,

$$y = A(1 - e^{-k_p t}) + k_{ss} t \quad (\text{Eq. 1})$$

where A indicates the burst amplitude, k_p is the first order rate, and k_{ss} is steady-state rate of nucleotide incorporation.

LC-MS/MS Analysis of Full-length Extended Products—A 5'-6-carboxyfluorescein-labeled primer containing a deoxyuridine (U) residue, 5'-CGG GCT CGT AAG CGT CUT-3', was used to generate shorter chains after the cleavage by uracil DNA glycosylase and piperidine treatment (26, 27). Reaction conditions were similar to those used in steady-state kinetic assays except that the final concentrations were as follows: 3 μ M hpol η and 25 μ M primer-template duplex, in a total volume of 80 μ l. The primer was extended in the presence of all four dNTPs (1 mM each) for 2–24 h at 37 °C. Reactions were terminated by removal of excess dNTP and Mg²⁺ using a spin column. The extent of reactions was monitored by electrophoresis prior to LC-MS/MS analysis. The reaction mixture was treated with 50 units of uracil DNA glycosylase and 0.25 M hot piperidine following a previous protocol (26). The cleavage solution was lyophilized and reconstituted in 60 μ l of H₂O.

LC-MS/MS analysis was performed on an Acquity ultraperformance liquid chromatography (UPLC) system (Waters Associates) coupled to a Thermo Finnigan LTQ mass spectrometer (Thermo Scientific, San Jose, CA) operating in an ESI negative mode. Samples were separated on an Acquity UPLC BEH octadecylsilane (C₁₈) column (1.7 μ m, 2.1 mm \times 100 mm) at a flow rate of 0.3 ml/min. The column temperature was maintained at 50 °C. Eluent A contained 10 mM NH₄CH₃CO₂ in 98% H₂O, 2% CH₃CN (v/v), and eluent B consisted of 10 mM NH₄CH₃CO₂ in 90% CH₃CN, 10% H₂O (v/v). A gradient program was run as follows: 0–3-min linear gradient from 0 to 3% B; 3–5-min linear gradient to 20% B; 5–6-min linear gradient to 100% B, held at 100% B for 2 min; 8–10 min from 100% B to 0% B, held at 0% B for 3 min. MS data were acquired using Xcalibur 2.1 software (Thermo). ESI settings were as follows: source voltage 4 kV; source current 100 A; capillary voltage -49 V; capillary temperature 350 °C; tube lens voltage -90 V. The most abundant species (-2 or -3 charged) were fragmented by collision-induced dissociation (CID) with a normalized collision energy of 35%. An activation Q setting of 0.25 and activation time of 30 ms were used. Oligonucleotide sequences were identified by comparing the observed CID spectra and theoretical spectra of candidate oligonucleotide sequences, which were calculated by software Mongo Oligo Calculator 2.0 from the University of Utah. The relative yields of various DNA extension products were based on their respective peak areas of the most abundant ions in extracted ion chromatograms (EICs). Some products with the same retention times and the same molecular weights were quantified by comparing the amounts of their specific CID fragments.

Crystallizations—Primer and template sequences employed in the crystallization experiments are listed in Table 1. DNA templates containing an AP site (THF-based mimic) were pur-

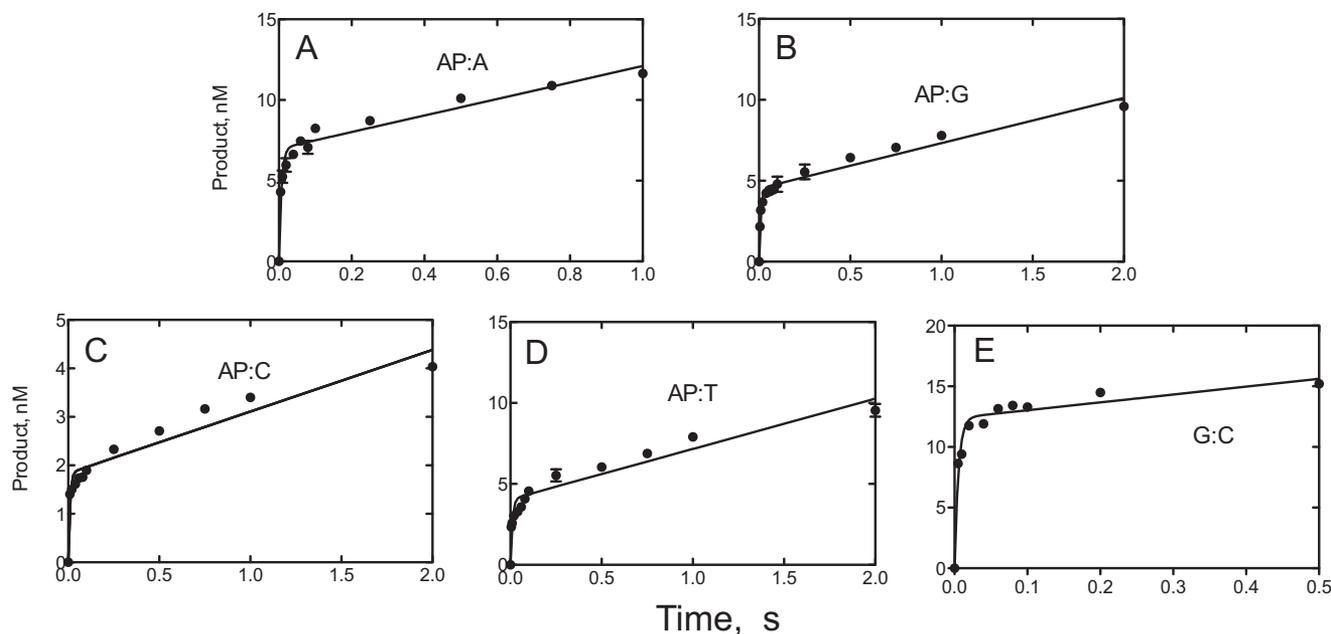


FIGURE 1. **Pre-steady-state kinetics of incorporation of individual dNTPs opposite G or an AP site.** A, dATP incorporation opposite AP site; B, dGTP incorporation opposite AP site; C, dCTP incorporation opposite AP site; D, dTTP incorporation opposite AP site; E, dCTP incorporation opposite G. hpol η (25 nM)-catalyzed incorporation of dNTP (0.5 mM) was opposite 50 nM 18/23-mer template DNA (18-mer, 5'-CGG GCT CGT AAG CGT CAT-3'; 23-mer, 3'-GCC CGA GCA TTC GCA GTA XTA CT-5', where X is either an AP site or G) at 37 °C. Reactions were quenched with 0.5 M EDTA at reaction times varying from 5 ms to 5 s. Products were analyzed by PAGE and phosphorimaging. All experiments were performed in duplicate. Data points are shown as means \pm S.D. and were fit to the equation $y = A(1 - e^{-k_p t})$, where k_p is the first order rate for the first catalytic cycle; k_{ss} is the steady-state rate, and A is the burst amplitude (GraphPad Prism).

chased from TriLink, and unmodified 8-mer primers were purchased from Integrated DNA Technologies. Template and primer strands were mixed in a 1:1 molar ratio and 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. Prior to crystallization, the DNA duplex was mixed with the protein in a 1.2:1 molar ratio in the presence of 50 mM Tris-HCl (pH 7.5) containing 450 mM KCl and 3 mM DTT. Following addition of 5 μ l of 100 mM MgCl₂ or CaCl₂, the complex was concentrated to a final concentration of \sim 2–3 mg/ml by ultrafiltration. Either non-hydrolyzable nucleoside triphosphates (dNMPNPP) or dCTP was then added to form the ternary complexes. Crystallization experiments were performed by the hanging drop vapor diffusion technique at 18 °C using a sparse matrix screen (Hampton Research, Aliso Viejo, CA) (28). One μ l of the complex solution was mixed with 1 μ l of reservoir solution and equilibrated against 500- μ l reservoir wells. Crystals appeared in droplets containing 0.1 M sodium MES (pH 5.5), containing 5 mM MgCl₂ or 5 mM CaCl₂, and 24–28% (w/v) PEG 2000 monomethyl ether within 1 day and grew to their maximum size within a week.

X-ray Diffraction Data Collection, Structural Determination, and Refinement—Crystals were mounted in nylon loops, cryo-protected in reservoir solution containing 25% glycerol (v/v), and frozen in liquid nitrogen. Diffraction data were collected on the 21-ID-D or 21-ID-F 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 (29). The structures were determined by the molecular replacement technique with the program MOLREP (30, 31) and using the hpol η structure with Protein Data Bank code 4O3N (protein

only) (23) as the search model. Structural refinement and model building were carried out with PHENIX (32) and COOT (33), respectively. Selected crystal data, data collection, and refinement parameters are listed in Table 10 (see below). Illustrations were prepared with the program UCSF Chimera (34).

RESULTS

Kinetics of dNTP Incorporation Opposite AP Sites—As a result of instability of natural AP sites, most crystal structure studies and kinetic assays have been carried out using a THF analog instead of true AP sites (Fig. 1) (18). In this study, experiments were performed using four DNA oligonucleotides containing the stable AP analog THF as well as two oligonucleotides containing a natural AP site (oligonucleotide sequences shown in Table 1). These oligonucleotides have similar sequences except for the neighboring bases at the 5' side of AP sites. Natural AP sites and the THF analog demonstrated similar steady-state kinetic properties (Table 2). In this study, undamaged G was used as the control template base because G is the most commonly lost base due to AP site formation in cells (18). Steady-state kinetic results indicated that hpol η can insert one base opposite an AP site, although the bypass efficiency is much lower than that opposite a normal base.

Many previous studies have reported that nucleotide incorporation of DNA pols follows the A rule opposite AP sites. Some DNA pols also obey a 5'-frameshift rule (11). The specificity of yeast polymerase η in the bypass AP sites has been intensively studied *in vivo* and *in vitro* (19, 35). The steady-state kinetics of AP site bypass by yeast polymerase η obeys the A rule with a dNTP selectivity of 1A: 0.53G: 0.12T: 0.051C (17). However, in this study, hpol η showed preferential incorporation of both dATP and dGTP in the different sequence contexts, with a

TABLE 1**Oligonucleotides used in this study**

X indicates AP site; Y indicates AP site analog (THF).

Oligonucleotides used for kinetic studies

Oligonucleotide name	Sequence
Primer	5'-CGG GCT CGT AAG CGT CAT-3'
Templates	
23-AP-T-mer	3'-GCC CGA GCA TTC GCA GTA XTA CT-5'
23-AP-C-mer	3'-GCC CGA GCA TTC GCA GTA XCA CT-5'
23-THF-T-mer	3'-GCC CGA GCA TTC GCA GTA YTA CT-5'
23-THF-C-mer	3'-GCC CGA GCA TTC GCA GTA YCA CT-5'
23-THF-A-mer	3'-GCC CGA GCA TTC GCA GTA YAA CT-5'
23-THF-G-mer	3'-GCC CGA GCA TTC GCA GTA YGA CT-5'
23-G-mer	3'-GCC CGA GCA TTC GCA GTA GGA CT-5'

Oligonucleotides used for crystallographic studies

Structure name	DNA sequence	Incoming nucleotide
AP:dAMPNPP	3'-TCG CAG TAY TAC-5'	dAMPNPP
	5'-AGC GTC AT-3'	
AP:dGMPNPP	3'-TCG CAG TAY TAC-5'	dGMPNPP
	5'-AGC GTC AT-3'	
AP Post-insertion complex	3'-TCG CAG TYG TAC-5'	dCTP
	5'-AGC GTC AA-3'	

TABLE 2**Steady-state kinetics of incorporation of individual dNTPs opposite G and an AP site by hpol η**

Template	dNTP	K_m μM	k_{cat} min^{-1}	k_{cat}/K_m $\mu\text{M}^{-1} \text{min}^{-1}$	dNTP selectivity ^a
23-AP-T-mer	A	44 ± 3	14.4 ± 0.2	0.33	1
	G	18 ± 1	6.0 ± 0.1	0.33	1
	C	283 ± 38	9.6 ± 0.5	0.034	0.10
	T	207 ± 17	4.5 ± 0.1	0.022	0.067
23-AP-C-mer	A	141 ± 11	21.3 ± 0.5	0.15	0.68
	G	52 ± 5	11.4 ± 0.2	0.22	1
	C	500 ± 33	6.2 ± 0.2	0.012	0.055
	T	392 ± 45	4.1 ± 0.2	0.010	0.045
23-THF-T-mer	A	108 ± 7	16 ± 0.3	0.15	1
	G	120 ± 8	10.7 ± 0.2	0.089	0.59
	T	416 ± 20	2.8 ± 0.1	0.0067	0.045
	C	817 ± 132	5.9 ± 0.5	0.0072	0.048
23-THF-C-mer	A	150 ± 6	21.1 ± 0.3	0.14	1
	G	101 ± 5	12.8 ± 0.2	0.13	0.93
	C	373 ± 23	3.2 ± 0.1	0.0086	0.061
	T	510 ± 27	4.8 ± 0.1	0.0094	0.067
23-THF-A-mer	A	234 ± 15	17 ± 0.4	0.073	1
	G	110 ± 6	7.7 ± 0.1	0.070	0.96
	C	390 ± 22	2.2 ± 0.1	0.0056	0.077
	T	576 ± 35	3.8 ± 0.1	0.0066	0.090
23-THF-G-mer	A	138 ± 18	13.1 ± 0.5	0.095	1
	G	106 ± 8	8.5 ± 0.2	0.080	0.84
	C	284 ± 23	2.2 ± 0.1	0.0077	0.081
	T	646 ± 40	3.8 ± 0.1	0.0059	0.062
23-G-mer	C ^b	1.3 ± 0.2	80 ± 3	62	

^a dNTP selectivity ratio = $(k_{\text{cat}}/K_m)_{\text{dNTP}}/(k_{\text{cat}}/K_m)_{\text{highest}}$.^b The steady-state kinetic data of dCTP incorporation opposite G are from Ref. 23.

dNTP selectivity of 0.68 to 1 for A and 0.59 to 1 for G (Table 2). With the templates containing an A or G 5' of an AP site, the insertion of dATP and dGTP was still favored over dCTP and dTTP.

TABLE 3**Pre-steady-state kinetics of incorporation of individual dNTPs opposite an AP site**See Fig. 1. The pol η concentration was 25 nM.

Template	dNTP	Amplitude	k_p	k_{ss}
		nM	s^{-1}	s^{-1}
23-AP-T-mer	A	7.0 ± 0.2	151 ± 20	0.20 ± 0.02
	G	4.5 ± 0.1	110 ± 16	0.11 ± 0.01
	C	1.8 ± 0.1	113 ± 27	0.051 ± 0.004
	T	4.1 ± 0.2	100 ± 22	0.12 ± 0.01
23-G-mer	C	12.4 ± 0.4	189 ± 30	0.26 ± 0.08

Pre-steady-state kinetics of dNTP incorporation opposite an AP site was also examined using a natural AP site-containing template with a neighboring base T (Table 3). Interestingly, the values of the burst rate k_p and the amplitude indicated the preference of dATP insertion opposite the lesion. The higher catalytic efficiency of the first turnover when inserting dATP may be a result from the stabilization of the reaction when the incoming dATP is complementary to the downstream template base (T) of an AP site. Pre-steady-state kinetic analysis showed substoichiometric bursts, 7–28%, with damaged templates and 50% burst amplitude with a normal G-contained template. The similar results were also observed in our prior work on hpol η and 8-oxoG (23). The small sub-stoichiometric bursts with AP sites in templates may be caused by the formation of nonproductive reversible complexes (37).

LC-MS/MS Analysis of Primer Extension Products—Although steady-state kinetic studies can gauge the efficiency of nucleotide incorporation catalyzed by a pol, LC-MS/MS methods we developed previously can reveal the ability of enzyme to read past and extend beyond a lesion (Fig. 2). LC-MS/MS experiments were performed following previous procedures (26, 27) using a uracil-containing primer. The sequences of extension products resulting from replication of six DNA templates with different downstream sequences and their yields were analyzed by LC-MS/MS (Tables 4–9). As shown in Tables 4 and 5, when templates contained a T 5' of either an AP site or the THF analog, the most abundant products were –1 frameshift products, with yields of 68–81%, followed by 17–30% of the products containing G opposite the lesions. Minor products (<5%) corresponded to T incorporation opposite lesions. With templates containing a C 5' of the AP (or THF) site, similar yields of –1 frameshift products were obtained (Tables 6 and 7), and A incorporation accounted for 26–32%. Only traces of products containing G or C opposite the lesion were obtained. When the T (or C) at the 5' side of THF was replaced with an A (or G), preferential incorporation of A and G catalyzed by hpol η was still observed, but there were almost no frameshift products observed (Tables 8 and 9). These results indicate that the mechanism of AP site bypass follows a “purine rule” and is partially sequence-dependent. Frameshifts were favored when the incoming dNTP was complementary to the template base downstream of an abasic site.

Crystal Structures of hpol η Tertiary Complexes at the Insertion Stage Opposite an AP Site—We determined structures of two hpol η complexes with either incoming dAMPNPP or dGMPNPP opposite the AP site (THF analog) at resolutions of 2.15 and 1.81 Å, respectively (Table 10 and Fig. 3, A and D). At the active site of both complexes, the base of the incoming

Human pol η and Abasic Sites

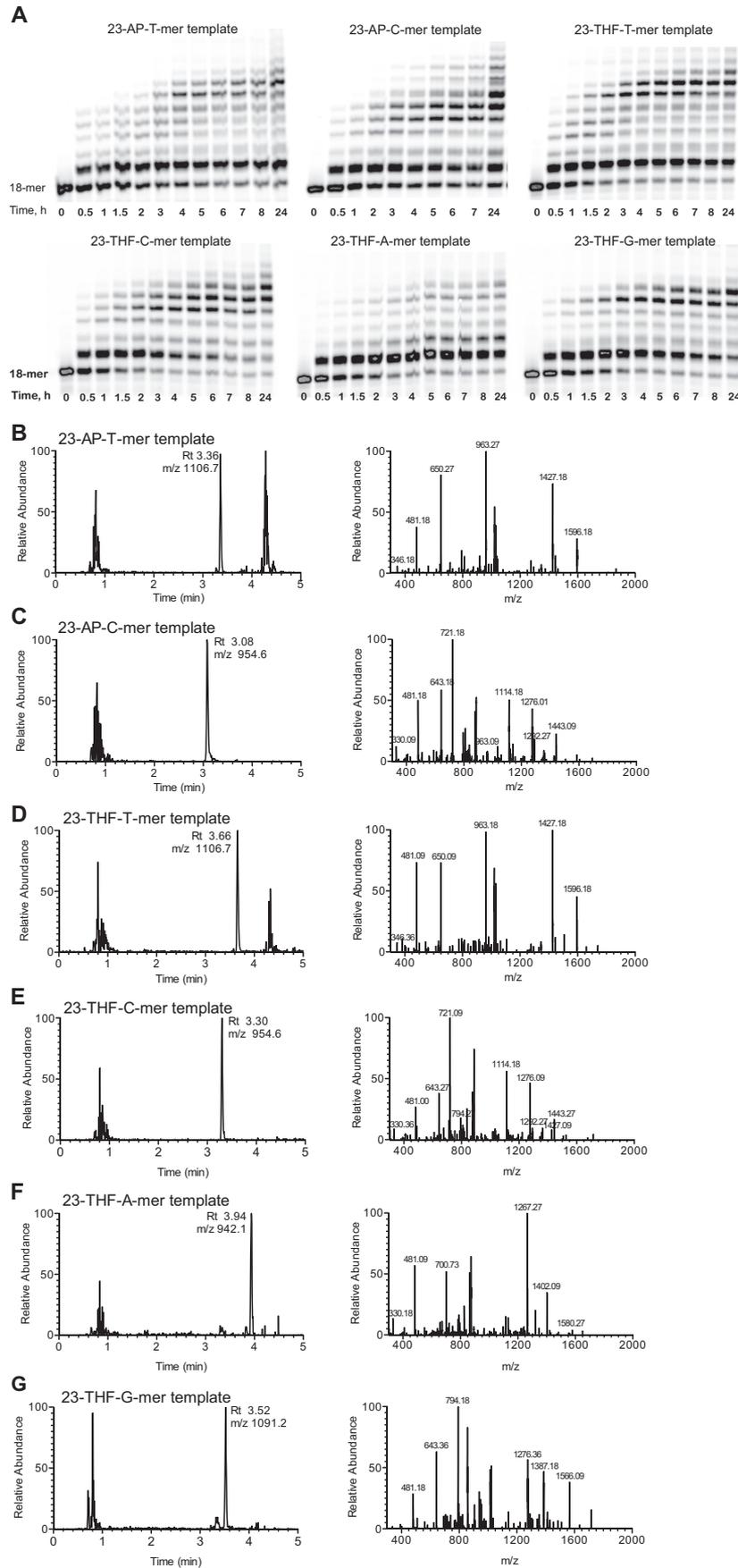


TABLE 4

LC-MS analysis of products of hpol η copying template 23-AP-T-mer (T 5' of AP site)

The _ indicates a frameshift site.

5' CGG GCT CGT AAG CGT CAT 3'
3' GCC CGA GCA TTC GCA GTA XTA CT 5' X= AP site

Fragments	<i>m/z</i> theoretical	<i>m/z</i> observed	Peak area	% Frameshift	% G	% T	
5'-pT_A-3'	634.41	634.18	3801	68			
5'-pT_ATG-3'	633.41	633.27	860				
5'-pT_ATGA-3'	790.01	789.82	2660				
5'-pT_ATGAA-3'	946.62	946.27	5522				
5'-pT_ATGAG-3'	954.62	954.36	1801				
5'-pT_ATGAT-3'	942.11	942.00	1263				
5'-pT_ATGATG-3'	1106.71	1106.27	10130				
5'-pT_ATGAGTA-3'	841.88	841.82	8114				
5'-pTGATGA-3'	954.62	954.36	1801		30		
5'-pTGATG-3'	798.01	798.27	434				
5'-pTG-3'	650.41	650.27	12778				
5'-pTT-3'	625.40	625.27	1325			2	

TABLE 5

LC-MS analysis of products of hpol η copying template 23-THF-T-mer (T 5' of THF site)

5' CGG GCT CGT AAG CGT CAT 3'
3' GCC CGA GCA TTC GCA GTA YTA CT 5' Y=THF

Products	<i>m/z</i> theoretical	<i>m/z</i> observed	Peak area	% Frameshift	% G	% T
5'-pT_ATGA-3'	790.01	789.82	1310	81		
5'-pT_ATGAA-3'	946.62	946.65	2379			
5'-pT_ATGAG-3'	954.62	954.27	1035			
5'-pT_ATGATG-3'	1106.72	1106.82	13623			
5'-pT_ATGATGA-3'	841.88	841.82	7368			
5'-pTGATGA-3'	954.62	954.27	1035			
5'-pTGATG-3'	798.01	797.82	521	17		
5'-pTG-3'	650.41	650.27	3814			
5'-pTT-3'	625.40	625.27	498			

nucleotide triphosphate is positioned opposite the AP residue (Fig. 3). In the complex with dAMPNPP, two water molecules mediate contacts between adenine (N1 and N6, Fig. 3, B and C) and the AP phosphate group in the major groove. In the minor groove, two more water molecules link N3(A) to O4' of the 2'-deoxyribose from residue A5 and to the side chain of Gln-38 (Fig. 3C). In the complex with dGMPNPP, a single water bridges N1 and O6 of G and the AP phosphate (Fig. 3, E and F). The contact in the minor groove involving N3 to Gln-38 is

TABLE 6

LC-MS analysis of products of hpol η copying template 23-AP-C-mer (C 5' of AP site)

5' CGG GCT CGT AAG CGT CAT 3'
3' GCC CGA GCA TTC GCA GTA XCA CT 5' X= AP site

Products	<i>m/z</i> theoretical	<i>m/z</i> observed	Peak area	% Frameshift	% A	% C	% G	
5'-pT_GTGA-3'	798.01	797.82	1424	66				
5'-pT_GTGAA-3'	954.62	954.27	3838					
5'-pT_GTGAG-3'	962.62	962.18	940					
5'-pT_GTGAC-3'	942.60	942.73	210					
5'-pT_GTGAT-3'	950.11	950.27	460					
5'-pT_GTGAAA-3'	1111.22	1111.18	1824					
5'-pT_GTGAAG-3'	1119.22	1118.73	627					
5'-pTAGTGA-3'	954.62	954.27	1890		32			
5'-pTAGTGA-3'	1111.22	1111.18	2321					
5'-pTAGTGA G-3'	1119.22	1118.73	350					
5'-pTCGTGA-3'	942.60	942.73	165			2		
5'-pTGGTGA-3'	1119.22	1118.73	85				<1	

TABLE 7

LC-MS analysis of products of hpol η copying template 23-THF-C-mer (C 5' of THF site)

5' CGG GCT CGT AAG CGT CAT 3'
3' GCC CGA GCA TTC GCA GTA YCA CT 5' Y=THF

Products	<i>m/z</i> theoretical	<i>m/z</i> observed	Peak area	% Frameshift	% A	% C	% G	
5'-pT_GTGA-3'	798.01	797.82	2857	72				
5'-pT_GTGAA-3'	954.62	954.36	10180					
5'-pT_GTGAG-3'	962.62	962.36	2550					
5'-pT_GTGAAA-3'	1111.22	1110.91	2179					
5'-pT_GTGAAG-3'	1119.22	1119.36	1141.4					
5'-pT_GTGAC-3'	942.60	942.27	1124					
5'-pT_GTGAT-3'	950.11	949.91	1040					
5'-pTAGTGA-3'	954.62	954.36	2085		22			
5'-pTAGTGA-3'	1111.22	1110.91	3268					
5'-pTAGTGA-3'	1119.22	1119.22	1054					
5'-pTCGTGA-3'	942.60	942.27	1111			4		
5'-pTTGTGA-3'	950.11	949.91	585				2	

direct, but the distance to O4' of A5 is too long for H-bond formation (Fig. 3F). Projections roughly along the normal to the nucleobase of the incoming nucleotide triphosphate demonstrate that the base portions of purine nucleotides cannot just establish water-mediated interactions with a portion of the AP residue but also form more optimal stacking interactions with

FIGURE 2. LC-MS analysis of products of extension of primer opposite AP sites in six template contexts by hpol η in the presence of all four dNTPs. A, gel images of primer extension past AP sites by hpol η. B, EIC and CID spectrum for *m/z* 1106.7 corresponding to sequence 5'-pTATGATG-3' extended opposite template 23-AP-T-mer. C, EIC and CID spectrum for *m/z* 954.6 corresponding to sequences 5'-pTGTGAA-3' and 5'-pTAGTGA-3' extended opposite template 23-AP-C-mer. D, EIC and CID spectrum for *m/z* 1106.7 corresponding to sequence 5'-pTATGATG-3' extended opposite template 23-THF-T-mer. E, EIC and CID spectrum for *m/z* 954.6 corresponding to sequences 5'-pTGTGAA-3' and 5'-pTAGTGA-3' extended opposite template 23-THF-C-mer. F, EIC and CID spectrum for *m/z* 942.1 corresponding to sequence 5'-pTATTGA-3' extended opposite template 23-THF-A-mer. G, EIC and CID spectrum for *m/z* 1091.2 corresponding to sequence 5'-pTACTGAA-3' extended opposite template 23-THF-G-mer. Template sequences used in this experiment are presented in Table 1. The primer sequence was 5'-CGG GCT CGT AAG CGT CTT-3'. The sequences of which EICs and CID are shown here are the most abundant species generated from different templates. DNA polymerase reactions contained 3 μM hpol η, 25 μM DNA duplex, and 1 mM dNTPs and were incubated at 37 °C for 2–24 h. LC-MS/MS analysis was performed on an Acquity UPLC system coupled to a Thermo Finnigan LTQ mass spectrometer operating in an ESI negative mode. Samples were separated on an Acquity UPLC BEH C18 column (1.7 μm, 2.1 mm × 100 mm) using a binary eluent (A, 10 mM NH₄CH₃CO₂ in 98% H₂O, 2% CH₃CN (v/v), and B 10 mM NH₄CH₃CO₂ in 90% CH₃CN, 10% H₂O (v/v)).

TABLE 8

LC-MS analysis of products of hpol η copying template 23-THF-A-mer (A 5' of THF site produced with uracil DNA glycosylase)

5' CGG GCT CGT AAG CGT CAT 3'
3' GCC CGA GCA TTC GCA GTA YAA CT 5' Y=THF

Products	<i>m/z</i> theoretical	<i>m/z</i> observed	Peak area	% A	% G
5'-pTA-3'	634.41	634.27	6066	80	
5'-pTATTGA-3'	942.11	942.27	7010		
5'-pTATTGAA-3'	1098.72	1098.82	3885		
5'-pTG-3'	650.41	650.27	4151		20

TABLE 9

LC-MS analysis of products of hpol η copying template 23-THF-A-mer (G 5' of THF site)

5' CGG GCT CGT AAG CGT CAT 3'
3' GCC CGA GCA TTC GCA GTA YGA CT 5' Y=THF

Products	<i>m/z</i> theoretical	<i>m/z</i> observed	Peak area	% Frameshift	% A	% G
5'-pT_CTGA-3'	778.00	777.82	1220	5		
5'-pT_CTGAA-3'	934.60	934.82	934			
5'-pTACTG-3'	778.00	777.82	228			
5'-pTACTGA-3'	934.60	934.82	11625	67		
5'-pTACTGAA-3'	1091.21	1091.36	12617			
5'-pTACTGAG-3'	1099.21	1098.82	3739			
5'-pTACTGAC-3'	1079.20	1078.82	2284			
5'-pTGCTGA-3'	942.60	942.36	7576	28		
5'-pTGCTGAA-3'	1099.21	1098.82	5119			

the adjacent nucleotide (A) from the template strand (Fig. 3, C and F). In comparison, the overlap between incoming A or incoming G and the thymine moiety from the 3'-terminal primer residue appears to be more limited. Superimposition of the active sites from the insertion complexes with dAMPNPP and dGMPNPP indicates that guanine has shifted toward the minor groove (Fig. 3, G and H). This shift goes along with an extended conformation of the Arg-61 side chain that interacts with the Hoogsteen edge (O6 and N7) of G (Fig. 3F). In the dAMPNPP complex, Arg-61 displays two alternative, curled conformations, whereby the guanidino moiety exhibits partial stacking on the nucleobase of the incoming nucleoside triphosphate in addition to interactions with the α - and β -phosphates (Fig. 3C).

In both structures, T3 (which is located 5' to the AP residue) is lodged outside the active site (Fig. 3, B and E). The crystal structure of the dGMPNPP complex is representative of the minor product seen in the LC-MS/MS analysis. Thus, it demonstrates that G is incorporated opposite the AP residue (Tables 4 and 5 and Fig. 3I, left panel) when the template nucleotide 5' to AP is T (similar for A when the template nucleotide 5' to AP is C (Tables 6 and 7)). However, the crystal structures only represent an intermediate stage in regard to the -1 frameshift products observed by LC-MS/MS, with A pairing with T 5' to AP (or G pairing with C 5' to AP; Tables 4–7, Fig. 3I, right panel). In order for the -1 frameshift products to form, T3 (or C in the case of incoming G) has to rotate around relative to the orientation seen in the two insertion complexes (Figs. 3, B and

E, and 5) and swing into the active site to form a Watson-Crick base pair with dATP (or C with dGTP).

Crystal Structure of an hpol η Extension Stage, Tertiary Complex Following an AP Site—We determined the structure of a post-insertion complex of hpol η with dA opposite the AP site (using the THF analog), followed by template G paired with incoming dCTP at 2.85 Å resolution (Table 4 and Fig. 4A). The configuration with A positioned opposite AP and an adjacent Watson-Crick G:C pair represents the major product of the *in vitro* bypass reaction of a template sequence 5'-G(AP)-3' as established by LC-MS/MS (Table 9). Unlike the case of the insertion complexes, where the incoming purine base and AP are bridged by water molecules (Fig. 3, B, C, E, and F), the water molecule H-bonded to N1 of A in the extension complex is too far removed from the AP 5'- and 3'-phosphate groups to be able to link the moieties (Fig. 4, B and C). The looser spacing is a consequence of the orientation of the template guanine 5' to the AP site (G4, Fig. 4B) that is accommodated inside the active site, resulting in the phosphates of both G4 and AP being pushed to the periphery of the template-primer duplex (Fig. 4C). With G4 providing a stacking platform, the next residue, T3, is also directed into the active site (Fig. 4B), rather than being swung outwards as in the case of the insertion complexes, where the AP residue is unable to engage in stacking (Figs. 3, B and E, and 5). The structural data at the extension stage together with the *in vitro* bypass assays as analyzed by LC-MS/MS (Tables 8 and 9) confirm the proclivity by hpol η for insertion of A and G opposite an AP site. A is preferred over G, independent of whether the template nucleotide 5' to AP is A (Table 8) or G (Table 9). By comparison, -1 frameshifting is negligible (5%, Table 9), apparently because incoming dCTP aligning with G 5' to AP is no match for purines positioned opposite AP and stabilized by a network of water molecules (Fig. 3, B and E).

DISCUSSION

An investigation of translesion synthesis across AP sites (THF analog) catalyzed by human B-family pols α and δ and the human Y-family pols η , ι , κ , and Rev1 revealed various abilities of the individual pols to deal with this ubiquitous lesion (18). Both hpol α and δ /PCNA favored insertion of A opposite the AP site, but hpol α was severely blocked after insertion, whereas δ /PCNA were proficient at insertion and extension. Among Y-family pols, pol η was the most efficient in terms of both insertion and next-base extension. hpol κ extended from a nucleotide inserted opposite AP (C or A), but it was much less efficient in doing so than either hpol η or δ /PCNA. Rev1 preferentially inserted C opposite AP, but extension was blocked, similar to hpol ι , which was found to insert T, G, and A when encountering an AP site but then stopped (18). Overall, the results of this study were consistent with hpol η alone possessing the ability to efficiently bypass an AP site. To further characterize hpol η function and activity, we used steady-state and pre-steady-state kinetic tools in combination with LC-MS/MS assays to investigate hpol η bypass synthesis opposite AP sites. To better understand the structural basis for the observed preferences by hpol η , we determined crystal structures of ternary hpol η -DNA-dNTP (dNMPNPP analogs) complexes with AP-

TABLE 10
Crystal data, data collection parameters, and structure refinement statistics

Complex	AP:dAMPNPP (insertion)	AP:dGMPNPP (insertion)	AP post-insertion complex, G:dCTP (extension)
Data collection			
Wavelength (Å)	0.97872	0.97872	1.07810
Space group	$P6_1$	$P6_1$	$P6_1$
Resolution (Å)	50.0-2.15 (2.19-2.15) ^a	50.0-1.81 (1.84-1.81)	50.0-2.85 (2.90-2.85)
Unit cell $a = b, c$ (Å)	98.68, 81.45	98.56, 81.88	98.79, 82.01
Unique reflections	24,471 (1,213)	41,364 (2,056)	10,398 (413)
Completeness (%)	98.9 (99.9)	100 (100)	94.0 (58.6)
$I/\sigma(I)$	16.21 (1.64)	21.16 (2.94)	5.03 (0.27)
Wilson B -factor (Å ²)	32.3	21.7	71.7
R -merge	0.119 (0.944)	0.083 (0.626)	0.207 (0.989)
Redundancy	7.6 (7.4)	7.2 (6.4)	5.0 (2.0)
Refinement			
R -work	0.167 (0.226)	0.168 (0.207)	0.241 (0.391)
R -free	0.220 (0.270)	0.202 (0.244)	0.285 (0.419)
No. of atoms protein/DNA	3,409/378	3,402/383	3,329/380
dNMPNPP/Mg ²⁺	30/2	31/2	28/1 (Ca ²⁺)
Water/glycerol	343/2	412/3	63/1
Protein residues	430	430	425
B -factor (Å ²)			
Average	34.3	26.0	58.4
Protein/DNA	33.6/36.6	24.8/29.6	58.1/61.1
dNMPNPP/Mg ²⁺	25.2/18.9	19.5/15.4	51.3/47.2 (Ca ²⁺)
Water/glycerol	39.5/40.3	33.2/27.4	60.7/64.1
Root mean square deviations			
Bonds (Å)	0.008	0.008	0.003
Angles (degree)	1.0	1.1	0.7
Ramachandran			
Favored (%)	97	98	93
Allowed (%)	2.3	1.8	7
Outliers (%)	0.2	0.2	0.5
Protein Data Bank code	4RNM	4RNN	4RNO

^a Statistics for the highest resolution shell are shown in parentheses.

containing template strands (THF analogs), trapped either at the insertion or extension stages.

The kinetic analysis, both at the steady-state and pre-steady-state levels, attests to the preferred insertion by hpol η of purine nucleoside triphosphates opposite an AP site. This preference is maintained irrespective of whether the template strand contains a real AP site or the THF analog. A slight preference for A insertion in the pre-steady-state assays (Table 3) is likely a consequence of the sequence context in that a T 5' to the AP site may result in a -1 frameshift and pairing of the incoming dATP with template T instead of opposite the AP residue. This interpretation is supported by the LC-MS/MS data demonstrating that templates with either a 5'-TpAP-3' or a 5'-CpAP-3' step result in -1 frameshifting, formation of a T:A or C:G pair, respectively, and an orphaned AP residue. Although the proclivity by hpol η to insert A opposite an AP site is slightly pronounced relative to G, this polymerase does not obey the so-called A rule (3–5). Instead, its bypass preference opposite AP sites is perhaps more appropriately termed a purine rule. With the *in vitro* bypass experiments augmented by LC-MS/MS analysis providing a clear picture as to the outcome of hpol η synthesis past an AP site, we turned to structural investigations to gain insight into the underlying basis for the preferred insertion of A and G.

Crystal structures of hpol η ternary complexes trapped at the insertion stage reveal that dATP and dGTP are positioned opposite the AP site (THF moiety) (Fig. 3). In both complexes, O3' of the 3'-terminal T of the primer strand is optimally positioned for attack at the α -phosphate group of incoming nucleotide triphosphate (O3'...P _{α} distances of 3.32 Å and 3.21 Å in the dAMPNPP and dGMPNPP complexes, respectively, and

nearly stretched O3'...(O)₃P-O angles in both cases). The T upstream from the AP residue in the complexes is unstacked from the template strand (Figs. 3, B and E, and 5). The phosphate group linking T and the AP moiety resides in the major groove and is shifted by 4.5 Å relative to the position of the corresponding phosphate in the extension complex (5'-P of dG; Fig. 5, orange arrow). Unlike the template strand at that site in the insertion complexes, the primer-template duplex in the extension complex adopts a conformation that resembles a standard B-form geometry. Thus, the orientation of the AP 5'-phosphate in the insertion complexes is highly unusual and unlikely to be encountered in DNA duplexes devoid of mismatches or unpaired nucleotides. However, such orientations of phosphates are more commonly encountered in RNA, *e.g.* in loop regions (38) or cross-strand purine stacks (39), where the phosphate is linked to a purine base from the opposite strand by a water molecule. Indeed, the location of the 5'-phosphate of the AP residue in the major groove allows it to get into relatively close proximity of the Watson-Crick edge of the incoming nucleotide base. Thus, water molecules mediate interactions between adenine (N1 and N6) and guanine (N1 and O6) and the AP 5'-phosphate group (Figs. 3 and 5). These stabilizing contacts most likely form the underlying basis of the purine rule, because they cannot be established by the nucleobase portions of incoming pyrimidine nucleotide triphosphates. The water bridges between incoming purines and the AP phosphate group are probably more important in terms of the favorable incorporation of A and G than stacking interactions between purine moieties and the 3'-terminal primer nucleotides. However, stacking interactions between A and G of the incoming nucle-

Human pol η and Abasic Sites

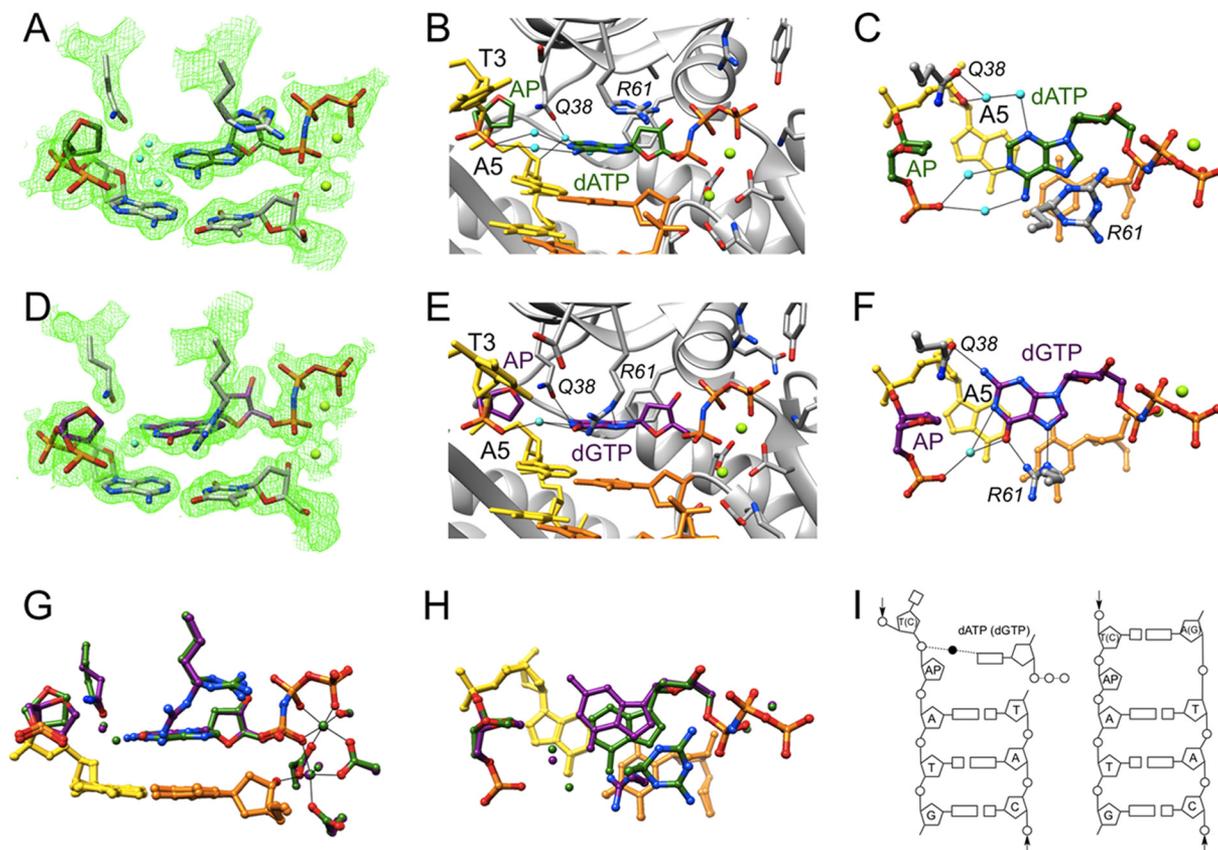


FIGURE 3. Active site configurations in the ternary hpol η insertion-step complexes with either dAMPNPP or dGMPNPP opposite the AP site. *A*, quality of the final Fourier $2F_o - F_c$ sum electron density (1σ threshold) in the active site region of the complex with dAMPNPP. *B*, view into the major groove, and *C*, rotated by $\sim 90^\circ$ around the horizontal axis and looking roughly along the normal to the nucleobase plane of the incoming dAMPNPP. The polymerase is shown as a schematic, and the DNA template (yellow)-primer (orange) duplex and selected hpol η side chains are shown in stick form (e.g. residues Gln-38 and Arg-61 from the finger domain; carbon atoms colored in gray). Carbon atoms of the AP residue and dAMPNPP are highlighted in green, and oxygen, nitrogen, and phosphorus atoms are colored in red, blue, and orange, respectively. Selected water molecules are cyan spheres; hydrogen bonds are thin solid lines, and Mg^{2+} ions are drawn as light green spheres. The side chain of Arg-61 displays two alternative conformations. *D*, quality of the final sum electron density for the complex with dGMPNPP and views of the active site from the major groove (*E*) and rotated by $\sim 90^\circ$ around the horizontal and looking roughly normal to the nucleobase plane of the incoming dGMPNPP (*F*). Carbon atoms of the AP residue and dGMPNPP are highlighted in purple, and oxygen, nitrogen, and phosphorus atoms are colored in red, blue, and orange, respectively. Overlay of the active sites in the hpol η ternary complexes with either incoming dAMPNPP or dGMPNPP, viewed into the major groove (*G*), and rotated by $\sim 90^\circ$ around the horizontal axis and looking roughly along the normal to the nucleobase plane of the incoming dNMPNPPs (*H*). Carbon atoms (G) of AP residue and dNMPNPP (the entire residues, *H*), water molecules, and Mg^{2+} ions as well as carbon atoms of amino acids Asn-38 and Arg-61 in the complexes with incoming A and G are highlighted in green and purple, respectively, and the template A and primer T of the adjacent base pair are colored in yellow and orange, respectively. The coordination of Mg^{2+} ions is indicated by thin solid lines in *G*. *I*, schematics of the bypass configuration encountered in crystal structures of complexes with incoming dATP or dGTP (left, implying insertion of the base opposite the AP site) and as established by LC-MS/MS (right, resulting in extensive frameshifting when the residue 5' to AP is T (incoming dATP) or C (incoming dGTP)). See also Tables 4–7.

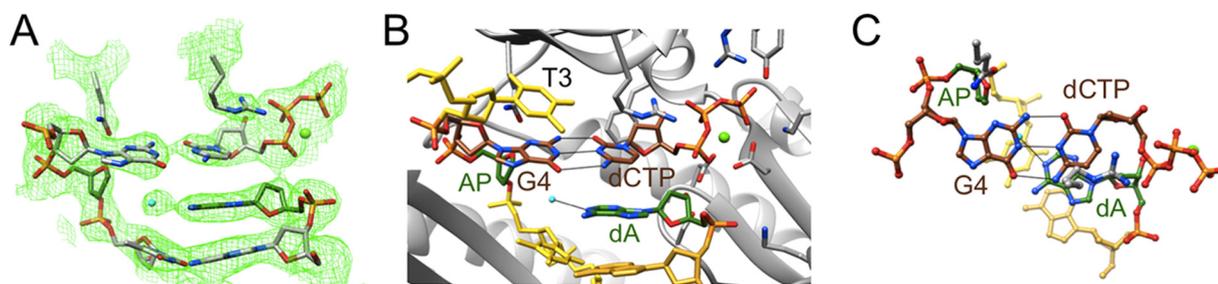


FIGURE 4. Active site configuration in the ternary hpol η extension-step complex with primer dA opposite the AP site followed by dCTP opposite template dG. *A*, quality of the final Fourier $2F_o - F_c$ sum electron density (1σ threshold) in the active site region. *B*, view into the major groove; *C*, rotated by $\sim 90^\circ$ around the horizontal axis and looking roughly along the normal to the nucleobase plane of the incoming dCTP. The polymerase is shown as a schematic, and the DNA template (yellow)-primer (orange) duplex and selected hpol η side chains are shown in stick form (e.g. residues Gln-38 and Arg-61 from the finger domain; carbon atoms colored in gray). Carbon atoms of the AP residue and dA are highlighted in green; carbon atoms of the following dG/dCTP pair are highlighted in brown, and oxygen, nitrogen, and phosphorus atoms are colored in red, blue, and orange, respectively. Selected water molecules are cyan spheres, hydrogen bonds are thin solid lines, and a Ca^{2+} ion is drawn as a light green sphere.

otides and the purine (A) 3'-adjacent to the AP site (Fig. 3, *C* and *F*) may also contribute to the preference for insertion of purines opposite an AP residue by hpol η .

Water-mediated “pairing” between A or G and the AP phosphate, as seen in the two insertion complexes, is only partially representative of the ultimate outcome of AP bypass by hpol η .

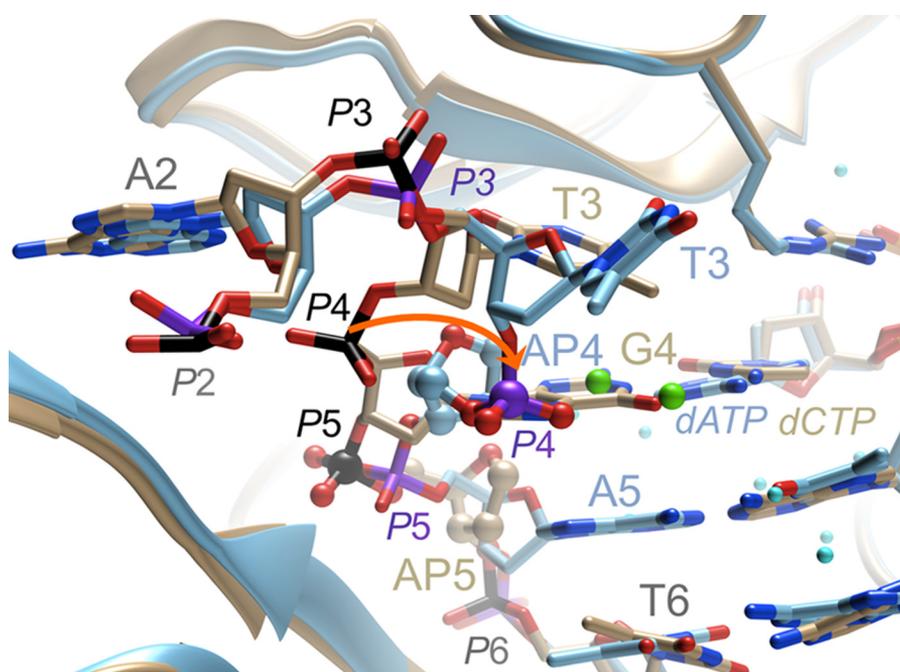


FIGURE 5. **Conformational changes in the DNA template strand containing an AP site (THF analog) between the hpol η insertion and extension stages.** Superimposition is shown of the hpol η insertion complex, with incoming dATP opposite the AP site (AP4, light blue carbon atoms and ribbon), and the hpol η extension complex, with incoming dCTP opposite G4, preceded by A opposite the AP site (AP5, beige carbon atoms and ribbon). Atoms of AP residues are shown in ball and stick mode; selected nucleotides and phosphorus atoms are labeled, and the latter are highlighted in purple in the insertion and in black in the extension complex. Water molecules are smaller spheres colored in cyan, except for two waters in green that mediate an interaction between dATP and the AP phosphate in the insertion complex. The superimposition reveals a significant movement of the T3 and AP4 residues into the major groove during insertion compared with the conformation of the template strand in the extension complex (orange arrow). Superimposition was performed using the MatchMaker option in UCSF Chimera.

Thus, the result of bypass by this pol is dependent on the identity of the template nucleotide 5'-adjacent to the AP moiety. The state trapped in the crystal structures of insertion complexes does reflect to a large extent the outcome of bypass in the case of incoming dATP, with G, A, or C (or dGTP, with G, A, or T) 5'-adjacent to the AP site. Thus, A is accommodated opposite AP, and the next step might then involve *e.g.* insertion of dCTP paired in a Watson-Crick mode with template G (as seen in the extension complex, Fig. 4). However, as the LC-MS/MS data attest (Tables 4–7), the main product of the bypass reaction catalyzed by hpol η with incoming dATP opposite an AP site with 5'-adjacent template T (Fig. 3, B and C) is a -1 frameshift, which results in formation of an A:T pair and leaves the AP site unopposed by a nucleotide from the extended primer strand (similar to the case of incoming dGTP opposite an AP site with 5'-adjacent template C). Conformational changes leading to the frameshift and starting from the geometries of the primer-template duplexes trapped in the crystal structures of the insertion complexes entail the rotation out of the major groove of the AP residue and the 5'-adjacent T toward a more standard B-form backbone geometry as seen in the extension complex (see Fig. 5 for orientation).

Comparisons between the translesion synthesis activities of hpol η and the Dpo4 pol from *S. solfataricus* reveal similar preferences in some cases, although the mechanisms of bypass may deviate. For example, both pols replicate efficiently and quite accurately past the 8-oxoG lesion (C/A ratios of 19:1 and 4:1 for Dpo4 and hpol η , respectively) but use amino acids from different domains (Arg-332, Dpo4 little finger *versus* Arg-61, hpol η finger) to mediate correct

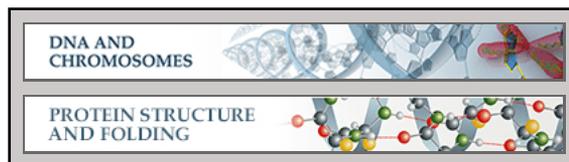
bypass (23, 36). Both pols efficiently catalyze replication of DNA past an AP lesion that results in frameshifting (-1 and $+1$ frameshifts with Dpo4 and -1 frameshifts with hpol η). However, crystal structures of Dpo4 complexes containing AP sites demonstrated that the AP residue is extrahelical, with the template nucleotide on the 5' side of the lesion directing the incoming nucleoside triphosphate and thus leading to -1 frameshifts (11) (" $5'$ rule"). Unlike the case of the hpol η insertion complexes described here, the base pairs bracketing the bulged AP moiety remain stacked seamlessly. In another configuration of the Dpo4-DNA-dNTP complex, triggering $+1$ frameshifts, the AP residue is accommodated in the solvent-exposed minor groove. But none of the Dpo4 complexes exhibit the DNA geometry seen in the insertion complexes of hpol η , with water molecules mediating contacts between A or G and the AP 5'-phosphate group that gives rise to the purine rule. Therefore, pols sculpt lesioned DNA in various ways to guarantee efficient insertion and extension reactions that result in more or less error-free replication. In the case of hpol η , the crystal structure of the extension step is consistent with efficient catalysis of the extension step in that the 3'-oxygen of the terminal primer nucleotide is positioned 3.9 Å from P_{α} of dCTP and with a near in-line orientation of O3' and the scissile P_{α} -O(P_{β}) bond. Overall, the structural and functional data presented here reveal a unique strategy by hpol η for efficiently bypassing AP sites (purine rule) that adds to the repertoire (*e.g.* A rule and $5'$ rule) used by DNA pols to cope with this common lesion.

Acknowledgments—We are grateful to K. Trisler for assistance in the preparation of this manuscript. Vanderbilt University is a member institution of the Life Sciences Collaborative Access Team at Sector 21 of the Advanced Photon Source, Argonne, IL. Use of the Advanced Photon Source at Argonne National Laboratory was supported by the United States Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract DE-AC02-06CH11357.

REFERENCES

- Friedberg, E. C., Walker, G. C., Siede, W., Wood, R. D., Schultz, R. A., and Ellenberger, T. (2006) *DNA Repair and Mutagenesis*, 2nd Ed., American Society for Microbiology, Washington, D. C.
- Strauss, B. S. (2002) The “A” rule revisited: polymerases as determinants of mutational specificity. *DNA Repair* **1**, 125–135
- Boiteux, S., and Laval, J. (1982) Coding properties of poly(deoxycytidylic acid) templates containing uracil or apyrimidinic sites: *in vitro* modulation of mutagenesis by deoxyribonucleic acid repair enzymes. *Biochemistry* **21**, 6746–6751
- Sagher, D., and Strauss, B. (1983) Insertion of nucleotides opposite apurinic/apyrimidinic sites in deoxyribonucleic acid during *in vitro* synthesis: uniqueness of adenine nucleotides. *Biochemistry* **22**, 4518–4526
- Schaaper, R. M., Glickman, B. W., and Loeb, L. A. (1982) Mutagenesis resulting from depurination is an SOS process. *Mutat. Res.* **106**, 1–9
- Efrati, E., Tocco, G., Eritja, R., Wilson, S. H., and Goodman, M. F. (1997) Abasic translesion synthesis by DNA polymerase β violates the “A-rule.” Novel types of nucleotide incorporation by human DNA polymerase β at an abasic lesion in different sequence contexts. *J. Biol. Chem.* **272**, 2559–2569
- Fiala, K. A., Hypes, C. D., and Suo, Z. (2007) Mechanism of abasic lesion bypass catalyzed by a Y-family DNA polymerase. *J. Biol. Chem.* **282**, 8188–8198
- Sagher, D., and Strauss, B. (1985) Abasic sites from cytosine as termination signals for DNA synthesis. *Nucleic Acids Res.* **13**, 4285–4298
- Obeid, S., Blatter, N., Kranaster, R., Schnur, A., Diederichs, K., Welte, W., and Marx, A. (2010) Replication through an abasic DNA lesion: structural basis for adenine selectivity. *EMBO J.* **29**, 1738–1747
- Zahn, K. E., Belrhali, H., Wallace, S. S., and Doublé, S. (2007) Caught bending the A-rule: crystal structures of translesion DNA synthesis with a non-natural nucleotide. *Biochemistry* **46**, 10551–10561
- Ling, H., Boudsocq, F., Woodgate, R., and Yang, W. (2004) Snapshots of replication through an abasic lesion; structural basis for base substitutions and frameshifts. *Mol. Cell* **13**, 751–762
- Beard, W. A., Shock, D. D., Batra, V. K., Pedersen, L. C., and Wilson, S. H. (2009) DNA polymerase β substrate specificity: side chain modulation of the “A-rule”. *J. Biol. Chem.* **284**, 31680–31689
- Nair, D. T., Johnson, R. E., Prakash, L., Prakash, S., and Aggarwal, A. K. (2011) DNA synthesis across an abasic lesion by yeast REV1 DNA polymerase. *J. Mol. Biol.* **406**, 18–28
- Nair, D. T., Johnson, R. E., Prakash, L., Prakash, S., and Aggarwal, A. K. (2009) DNA synthesis across an abasic lesion by human DNA polymerase ι . *Structure* **17**, 530–537
- Freisinger, E., Grollman, A. P., Miller, H., and Kisker, C. (2004) Lesion (in)tolerance reveals insights into DNA replication fidelity. *EMBO J.* **23**, 1494–1505
- Hogg, M., Wallace, S. S., and Doublé, S. (2004) Crystallographic snapshots of a replicative DNA polymerase encountering an abasic site. *EMBO J.* **23**, 1483–1493
- Haracska, L., Washington, M. T., Prakash, S., and Prakash, L. (2001) Inefficient bypass of an abasic site by DNA polymerase η . Inefficient bypass of an abasic site by DNA polymerase η . *J. Biol. Chem.* **276**, 6861–6866
- Choi, J.-Y., Lim, S., Kim, E. J., Jo, A., and Guengerich, F. P. (2010) Translesion synthesis across abasic lesions by human B-family and Y-family DNA polymerases α , δ , η , ι , κ , and REV1. *J. Mol. Biol.* **404**, 34–44
- Zhao, B., Xie, Z., Shen, H., and Wang, Z. (2004) Role of DNA polymerase η in the bypass of abasic sites in yeast cells. *Nucleic Acids Res.* **32**, 3984–3994
- Gibbs, P. E., McDonald, J., Woodgate, R., and Lawrence, C. W. (2005) The relative roles *in vivo* of *Saccharomyces cerevisiae* pol η , pol ζ , Rev1 protein and pol32 in the bypass and mutation induction of an abasic site, T-T (6-4) photoadduct and T-T *cis-syn* cyclobutane dimer. *Genetics* **169**, 575–582
- Kokoska, R. J., McCulloch, S. D., and Kunkel, T. A. (2003) The efficiency and specificity of apurinic/apyrimidinic site bypass by human DNA polymerase η and *Sulfolobus solfataricus* Dpo4. *J. Biol. Chem.* **278**, 50537–50545
- Biertümpfel, C., Zhao, Y., Kondo, Y., Ramón-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 η . *Nature* **465**, 1044–1048
- Patra, A., Nagy, L. D., Zhang, Q., Su, Y., Müller, L., Guengerich, F. P., and Egli, M. (2014) Kinetics, structure, and mechanism of 8-oxo-7,8-dihydro-2'-deoxyguanosine bypass by human DNA polymerase η . *J. Biol. Chem.* **289**, 16867–16882
- Boosalis, M. S., Petruska, J., and Goodman, M. F. (1987) DNA polymerase insertion fidelity. Gel assay for site-specific kinetics. *J. Biol. Chem.* **262**, 14689–14696
- O'Flaherty, D. K., and Guengerich, F. P. (2014) Steady-state kinetic analysis of DNA polymerase single-nucleotide incorporation products. *Curr. Protoc. Nucleic Acid Chem.* **59**, 7.21.1–7.21.13
- Zang, H., Goodenough, A. K., Choi, J. Y., Irimia, A., Loukachevitch, L. V., Kozekov, I. D., Angel, K. C., Rizzo, C. J., Egli, M., and Guengerich, F. P. (2005) DNA adduct bypass polymerization by *Sulfolobus solfataricus* DNA polymerase Dpo4: analysis and crystal structures of multiple base pair substitution and frameshift products with the adduct 1,N²-ethenoguanine. *J. Biol. Chem.* **280**, 29750–29764
- Christov, P. P., Angel, K. C., Guengerich, F. P., and Rizzo, C. J. (2009) Replication past the N⁵-methyl-formamidopyrimidine lesion of deoxyguanosine by DNA polymerases and an improved procedure for sequence analysis of *in vitro* bypass products by mass spectrometry. *Chem. Res. Toxicol.* **22**, 1086–1095
- Jancarik, J., and Kim, S.-H. (1991) Sparse matrix sampling: a screening method for crystallization of proteins. *J. Appl. Crystallogr.* **24**, 409–411
- Otwinowski, Z., and Minor, W. (1997) Processing of x-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326
- Vagin, A., and Teplyakov, A. (1997) MOLREP: an automated program for molecular replacement. *J. Appl. Crystallogr.* **30**, 1022–1025
- Collaborative Computational Project No. 4 (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **50**, 760–763
- Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., et al. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221
- Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132
- Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004) UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612
- Yuan, F., Zhang, Y., Rajpal, D. K., Wu, X., Guo, D., Wang, M., Taylor, J. S., and Wang, Z. (2000) Specificity of DNA lesion bypass by the yeast DNA polymerase η . *J. Biol. Chem.* **275**, 8233–8239
- Zang, H., Irimia, A., Choi, J.-Y., Angel, K. C., Loukachevitch, L. V., Egli, M., and Guengerich, F. P. (2006) Efficient and high fidelity incorporation of dCTP opposite 7,8-dihydro-8-oxodeoxyguanosine by *Sulfolobus solfataricus* DNA polymerase Dpo4. *J. Biol. Chem.* **281**, 2358–2372
- Furge, L. L., and Guengerich, F. P. (1999) Explanation of pre-steady-state kinetics and decreased burst amplitude of HIV-1 reverse transcriptase at sites of modified DNA bases with an additional, nonproductive enzyme-DNA-nucleotide complex. *Biochemistry* **38**, 4818–4825
- Jucker, F. M., and Pardi, A. (1995) GNRA tetraloops make a U-turn. *RNA* **1**, 219–222
- Correll, C. C., Freeborn, B., Moore, P. B., and Steitz, T. A. (1997) Metals, motifs, and recognition in the crystal structure of a 5S rRNA domain. *Cell* **91**, 705–712

**DNA and Chromosomes:
Structural and Kinetic Analysis of
Nucleoside Triphosphate Incorporation
Opposite an Abasic Site by Human
Translesion DNA Polymerase η**



Amritaj Patra, Qianqian Zhang, Li Lei, Yan
Su, Martin Egli and F. Peter Guengerich
J. Biol. Chem. 2015, 290:8028-8038.

doi: 10.1074/jbc.M115.637561 originally published online February 9, 2015

Access the most updated version of this article at doi: [10.1074/jbc.M115.637561](https://doi.org/10.1074/jbc.M115.637561)

Find articles, minireviews, Reflections and Classics on similar topics on the [JBC Affinity Sites](#).

Alerts:

- [When this article is cited](#)
- [When a correction for this article is posted](#)

[Click here](#) to choose from all of JBC's e-mail alerts

This article cites 38 references, 16 of which can be accessed free at
<http://www.jbc.org/content/290/13/8028.full.html#ref-list-1>