

Replication Bypass of the *N*-(2-Deoxy-D-erythro-pentofuranosyl)-urea DNA Lesion by Human DNA Polymerase η

Rachana Tomar, Songlin Li, Martin Egli,* and Michael P. Stone*



Cite This: <https://doi.org/10.1021/acs.biochem.3c00569>



Read Online

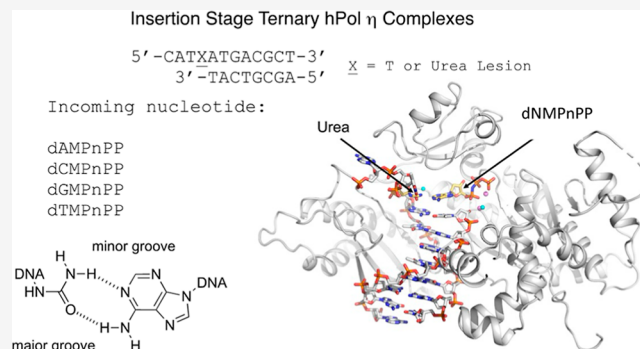
ACCESS |

Metrics & More

Article Recommendations

Supporting Information

ABSTRACT: Urea lesions in DNA arise from thymine glycol (Tg) or 8-oxo-dG; their genotoxicity is thought to arise in part due to their potential to accommodate the insertion of all four dNTPs during error-prone replication. Replication bypass with human DNA polymerase η (hPol η) confirmed that all four dNTPs were inserted opposite urea lesions but with purines exhibiting greater incorporation efficiency. X-ray crystal structures of ternary replication bypass complexes in the presence of Mg^{2+} ions with incoming dNTP analogs dAMPnPP, dCMPnPP, dGMPnPP, and dTMPnPP bound opposite urea lesions (hPol η -DNA-dNMPnPP complexes) revealed all were accommodated by hPol η . In each, the Watson–Crick face of the dNMPnPP was paired with the urea lesion, exploiting the ability of the amine and carbonyl groups of the urea to act as H-bond donors or acceptors, respectively. With incoming dAMPnPP or dGMPnPP, the distance between the imino nitrogen of urea and the N9 atoms of incoming dNMPnPP approximated the canonical distance of 9 Å in B-DNA. With incoming dCMPnPP or dTMPnPP, the corresponding distance of about 7 Å was less ideal. Improved base-stacking interactions were also observed with incoming purines vs pyrimidines. Nevertheless, in each instance, the α -phosphate of incoming dNMPnPPs was close to the 3'-hydroxyl group of the primer terminus, consistent with the catalysis of nucleotidyl transfer and the observation that all four nucleotides could be inserted opposite urea lesions. Preferential insertion of purines by hPol η may explain, in part, why the urea-directed spectrum of mutations arising from Tg vs 8-oxo-dG lesions differs.



1. INTRODUCTION

The oxidative DNA lesion thymine glycol (Tg) may undergo hydrolytic fragmentation to form *N*-(2-deoxy-D-erythro-pentofuranosyl)-urea, the urea lesion (Scheme 1).^{1–3} The urea lesion may also form from the oxidation of 7, 8-dihydro-8-oxo-deoxyguanosine (8-oxo-dG) and subsequent hydrolysis (Scheme 1).⁴ The initial synthesis of site-specific urea lesions in DNA was accomplished by Teoule and co-workers.⁵ In nucleosides^{5,6} and single-stranded DNA,⁷ urea lesions equilibrate between α - and β -anomeric configurations (Scheme 1). The urea lesion,^{1,8,9} like an abasic site (AP site),^{10–12} lacks its original coding information. It blocks high-fidelity replicative DNA synthesis and is mutagenic.^{1,2,4,13–15} Replicative pols that belong to the B- or X-families are blocked by urea lesions either immediately before the damage site (pols β and ϵ) or following preferential insertion of dAMP opposite to it (pols δ , α , and ζ).⁹

Cells utilize specialized replication machinery to synthesize DNA across chemically modified bases or lesions that block replication or compromise genome integrity.^{16,17} Human DNA polymerase η (hPol η ; UniProt Q9Y253), a translesion synthesis (TLS) DNA polymerase, plays a critical role in the error-prone bypass of oxidative damage in DNA; it bypasses UV-induced *cis*-syn pyrimidine dimers, the most common

naturally occurring DNA lesion.^{18,19} The inactivation of hPol η observed in xeroderma pigmentosum human variants^{18,20–24} is characterized by increased photosensitivity and incidence of skin cancer.^{25–27} The common DNA oxidative damage lesions Tg and 8-oxo-dG DNA and urea lesions,^{1,2} are also replicated by hPol η .^{8,9} hPol η bypasses 8-oxo-dG by inserting dCTP opposite to this lesion,^{24,28} but concomitantly has a greater potential to mis-incorporate dATP (250-fold increase relative to guanine). hPol η can incorporate all four dNTPs opposite the urea lesion and extend to full-length products beyond this lesion.⁹ As well, Dpo4, a thermostable TLS pol, elongates primers to full-length past urea lesions, preferentially inserting dATP and dGTP.⁹

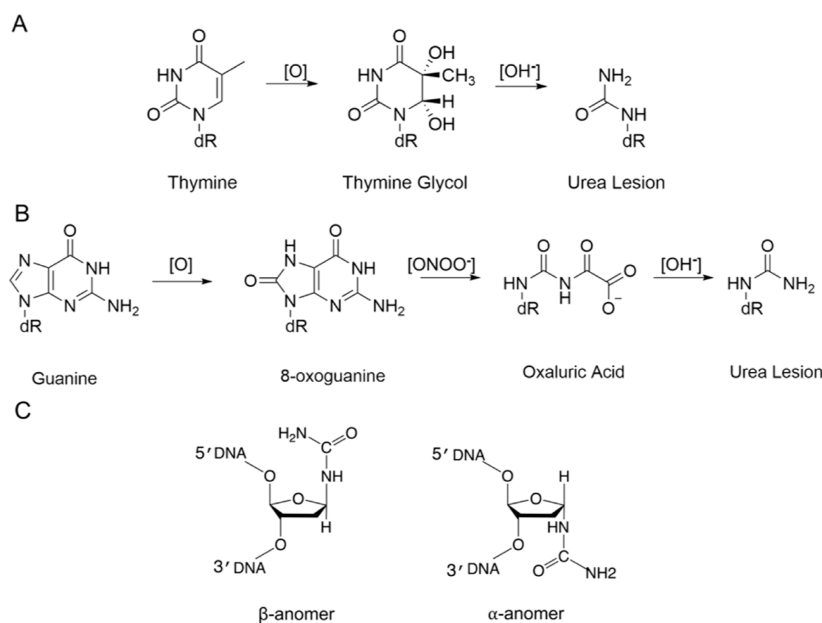
It remains unclear how hPol η incorporates all four dNTPs opposite urea lesions for successful nucleotidyl transfer reactions and error-free or error-prone replication bypass.

Received: October 17, 2023

Revised: December 26, 2023

Accepted: December 27, 2023

Scheme 1. Urea Lesion is a Secondary Oxidative DNA Damage Product; (A) Urea Lesion Arising from a Tg Lesion; (B) Urea Lesion Arising from an 8-Oxo-dG Lesion; (C) α and β Anomers of the Urea Lesion

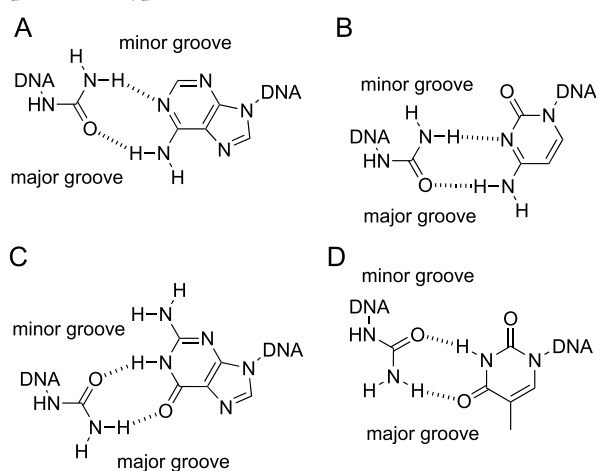


Molecular dynamics (MD) simulations predicted that during replication, urea lesions may form H-bonds and stacking interactions similar to Watson–Crick base pairing in B-DNA duplexes, with all four deoxyribonucleotide triphosphates (dNTPs) (Chart 1).²⁹ It is thought that hPol η employs conserved active site residues Arg61 and Gln38 to modulate correct vs incorrect nucleotide insertion opposite DNA lesions and improve incorporation efficiency.³⁰ The positively charged Arg61 is thought to contribute to the nucleotidyl transfer reaction in a position-dependent manner, whereas the Gln38-

mediated H-bonding interaction may stabilize the template base.³¹

Here, site-specifically urea-adducted oligodeoxynucleotide templates were prepared by base hydrolysis of corresponding site-specific thymine glycol (Tg)-containing oligodeoxynucleotide templates³ (Chart 2). We performed individual nucleotide incorporation assays with hPol η utilizing different dNTPs opposite the urea lesion; these confirm that hPol η incorporates all four nucleotides opposite the urea lesion. Moreover, our data reveal that purines exhibit greater incorporation efficiency opposite urea lesions as compared to pyrimidines. Using X-ray crystallography, we investigated ternary replication complexes of hPol η with incoming nonhydrolyzable dNTP analogs dAMPnPP, dCMPnPP, dGMPnPP, and dTMPnPP opposite the urea lesion and in the presence of Mg²⁺. These provide evidence that during error-prone bypass of the lesion, the urea moiety does form H-bonds and stack with all four possible incoming deoxyribonucleotide triphosphates (dNTPs) (Chart 1).²⁹ The active site configurations of hpol η in the complexes allow for insight into the preferential incorporation of purines. Our work affords further insight into the structural basis of error-free and mutagenic TLS by hPol η across urea lesions.

Chart 1. Urea Lesion Potentially Forms H-Bonding Interactions with All Incoming Nucleotides during Replication Bypass with hPol η ^a



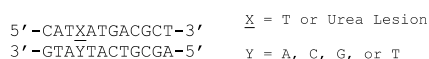
^aWith incoming (A) dATP and (B) dCTP, both with exocyclic amino groups in the major groove, the urea lesion may orient such that the urea amino nitrogen can form a H-bond with the imino nitrogen of the incoming nucleobase. With incoming (C) dGTP and (D) dTTP, both with exocyclic keto oxygens in the major groove, the urea lesion may orient such that the urea amino nitrogen can form a H-bond with the keto oxygen of the incoming nucleobase.

2. MATERIALS AND METHODS

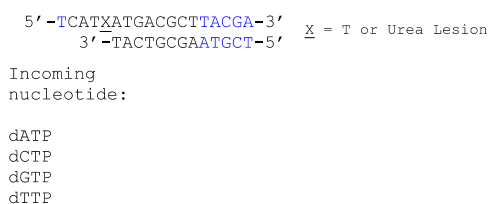
2.1. Site-specific Urea Lesion Synthesis. **2.1.1. 5'-CAT \underline{X} ATGACGCT-3'.** The template oligodeoxynucleotides 5'-CAT \underline{X} ATGACGCT-3' (\underline{X} = Tg) and 5'-AGCGTCAT-3' (Chart 2) were purchased from Integrated DNA Technologies, Inc. (Coralville, IA). Both were purified by RP-HPLC using a Gemini C₁₈ 250 mm \times 10 mm column (Phenomenex, Torrance, CA). To prepare site-specific urea adducts in this sequence for melting experiments and X-ray crystallography, 80 nmol of the Tg-containing oligodeoxynucleotides were hydrolyzed overnight in 0.2 M sodium phosphate (pH 12.0)^{3,7} at room temperature. The product 5'-CAT \underline{X} ATGACGCT-3' (\underline{X} = urea lesion) existed as an equilibrium mixture of the two species. The two species could be separated by C₁₈ HPLC at 2

Chart 2. (A) Duplex Oligodeoxynucleotides used for thermal melting (T_m) studies. (B) Urea-modified template-primer complex used for nucleotide incorporation studies in the presence of hPol η . For these studies, an additional deoxythymidine (blue) was added to the 5' end of the template, and five additional nucleotides (blue) were appended to the 3'-end of the template and the 5'-end of the primer sequences, respectively. (C) Urea-modified template primer used for X-ray crystallographic studies. The numbering of nucleotides near the active site, which corresponds to the core sequence used in the thermal melting experiments, corresponds with the nucleotide numbering system used in X-ray crystallographic studies.

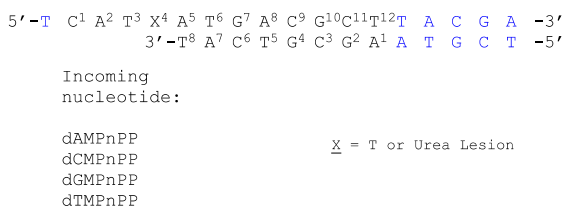
A. Thermal Melting (T_m) Experiments



B. Single Nucleotide Incorporation Assays Mediated by hPol η (insertion stage)



C. X-ray Crystallography Insertion Stage Ternary hPol η Complexes



mL/min using a 5–10% acetonitrile gradient in 0.1 M ammonium formate (pH 8.0). The identity of both fractions was confirmed by MALDI-TOF mass spectrometry; peak 1, observed mass, 3556.2; and peak 2, observed mass 3555.6; theoretical mass 3554.4.

2.1.2. 5'-TCAT $\underline{\text{X}}$ ATGACGCTTACGA-3'. The template oligodeoxynucleotide 5'-TCAT $\underline{\text{X}}$ ATGACGCTTACGA-3' ($\underline{\text{X}} = \text{Tg}$) (Chart 2) was purchased from GenScript USA Inc. (New Jersey, USA). The fluorophore FAM-labeled oligodeoxynucleotide 5'-56-FAM/TCGTAAGCGTCAT-3' was purchased from Integrated DNA Technologies, Inc. (Coralville, IA). Both were purified by RP-HPLC using a Gemini C₁₈ 250 mm \times 10 mm column (Phenomenex, Inc., Torrance, CA). To prepare site-specific urea adducts in this sequence for replication assays, 80 nmol of the Tg-containing oligodeoxynucleotide were hydrolyzed overnight in 0.2 M sodium phosphate (pH 12.0)^{3,5,7,8} at room temperature. Like before, the product 5'-TCAT $\underline{\text{X}}$ ATGACGCTTACGA-3' ($\underline{\text{X}} = \text{urea lesion}$) existed as an equilibrium mixture of two species. These could be separated by C₁₈ HPLC at 2 mL/min using a 5–15% acetonitrile gradient in 0.1 M ammonium formate (pH 8.0). The identity of the urea-adducted 18-mer oligodeoxynucleotide products was confirmed by MALDI-TOF mass spectrometry; peak 1, observed mass 5407.6; peak 2, observed mass 5406.8; theoretical mass 5405.3.

2.2. Expression and Purification of hPol η . hPol η , residues 1–432, in a pET28a plasmid construct was a gift from

W. Yang (NIDDK, National Institutes of Health, Bethesda, MD). To express hPol η , the corresponding plasmid (containing a kanamycin resistance gene) was transformed into *Escherichia coli* BL21 gold (DE3) competent cells (Agilent Technologies, Santa Clara, CA). The bacterial culture was grown at 37 °C in Luria broth medium (Research Product International, Mt Prospect, IL) containing 50 $\mu\text{g}/\text{mL}$ ampicillin (RPI) until the OD_{600 nm} reached ~ 0.6 . The incubator temperature was then changed to 18 °C, and a 0.5 mM final concentration of isopropyl β -D-1-thiogalactopyranoside (IPTG) was added to the cell culture to induce protein expression. Following a 20 h incubation at 18 °C, cells were harvested and lysed using a LM20 microfluidizer (Microfluidics, Westwood, MA) at 18,000 psi in a solution containing 500 mM NaCl, 20% glycerol, 10 mM imidazole, 5 mM β -mercaptoethanol, and a protease inhibitor cocktail (Roche, South San Francisco, CA), and 50 mM 2-Amino-2-(hydroxymethyl)propane-1,3-diol (Tris)-HCl (pH 8.0). Cell debris was removed by centrifugation at 18,000 rpm for 45 min at 4 °C. After passage through a 0.45 μm filter, the cell supernatant was loaded onto a Ni-NTA HisTrap HP 5 mL column (Cytiva Life Sciences, Marlborough, MA) using an AKTA pure 25 M (Cytiva Life Sciences, Marlborough, MA) to bind the N-terminal 6 \times His tag. The column was washed with ~ 300 mL of 500 mM NaCl, 20% glycerol, 40 mM imidazole, and 5 mM β -mercaptoethanol in 50 mM Tris-HCl (pH 8.0), and the protein was eluted with 500 mM NaCl, 20% glycerol, 300 mM imidazole, and 5 mM β -mercaptoethanol in 50 mM Tris-HCl (pH 8.0). Ni-NTA purified hPol η was concentrated using amicon centrifugal filters with membrane cutoff 30 kDa (Millipore Sigma, Burlington, MA) and then exchanged with 125 mM NaCl, 3 mM dithiothreitol (DTT) and 5% glycerol in Tris-HCl (pH 7.5). PreScission protease (APEX-BIO, Houston, TX) was added to the filtrate in a 1:100 units/ μg amount to cleave the His-tag. Cleaved protein was passed through Amicon centrifugal filters with a membrane cutoff of 10 kDa (Millipore Sigma, Burlington, MA) to be exchanged with 450 mM KCl, 5% glycerol, 1 mM Tris(2-carboxyethyl) hydroxy phosphine hydrochloride (TCEP), and 3 mM DTT in 20 mM Tris-HCl (pH 7.5). To separate hPol η from remaining impurities, the solution was passed through a size exclusion column, Superdex 200 increase (10/300 GL) (Cytiva Life Sciences, Marlborough, MA), that was pre-equilibrated with 450 mM KCl, 5% glycerol, 1 mM TCEP, and 3 mM DTT in 20 mM Tris-HCl (pH 7.5) using a Fast Performance Liquid Chromatography (FPLC) system (Cytiva Life Sciences, Marlborough, MA). Finally, protein purity was checked using 4–12% SDS PAGE. The mass of the monomeric protein was confirmed by MALDI-TOF mass spectrometry. The theoretical and observed masses for the cleaved protein were 48,556.77 and 48,580.56 Da, respectively. Protein concentration was determined by UV absorbance at 280 nm (A_{280} , 1 mg/mL ~ 1.03), and the solution was then concentrated to 2–3 mg/mL protein concentration to produce crystallization stock solutions. For long-term storage, the protein was flash frozen in liquid nitrogen and stored at -80 °C.

2.3. DNA Replication Assays. Site-specific urea lesion-containing 18-mer oligodeoxynucleotides (template strand) corresponding to each HPLC purified urea-containing strand was annealed with 5'-FAM-labeled 13-mer oligodeoxynucleotide (primer strand) (Chart 2) in a 1:1 molar ratio in 50 mM NaCl, 50 μM EDTA (sodium salt), and 20 mM Tris (pH 8.5) at room temperature for 10 min and then stored in a -20 °C

freezer. To assess different deoxynucleotide incorporation rates opposite the urea lesion, 10 nM purified hPol η was incubated with 150 nM urea-containing template:primer DNA in 50 mM NaCl, 5 mM DTT, 5% glycerol, 5 mM MgCl₂, 100 μ g/mL bovine serum albumin (BSA), and 50 mM Tris-HCl (pH 7.5), in a 50 μ L volume and pre-equilibrated at 37 °C for 10 min before the addition of 50 μ M of dATP, dCTP, dGTP, or dTTP, in separate reactions. These experiments were performed in triplicate. Reaction solutions were mixed and incubated at 37 °C for 40 min. At time points 0, 2, 5, 10, 20, and 40 min, 3.5 μ L of each reaction was mixed with 6.5 μ L of quench solution containing 95% formamide, 20 mM EDTA, and sodium salt. Thereafter, 10 μ L of 2 \times TBE-urea-bromophenol blue and xylene cyanol-containing loading buffer (Invitrogen, Waltham, MA) was added to each sample, followed by heating at 95 °C for 5 min to denature the sample. Next, samples were centrifuged, and 6 μ L of reaction mix were loaded into wells in a 15% TBE-urea gel (Invitrogen, Waltham, MA) to separate products at 0, +1, or +2 primer sites by gel electrophoresis. Gels were visualized using a Bio-Rad ChemiDoc imaging system (Bio-Rad Laboratories, Hercules, CA).

2.4. T_m Studies. Template strand oligodeoxynucleotide 12-mers with a site-specific urea lesion corresponding to each HPLC purified strand were separately annealed with equimolar amounts of the respective complementary 12-mer oligodeoxynucleotides, containing adenine (A), cytosine (C), guanine (G), or thymine (T) opposite the urea lesion (Chart 2). The solutions containing 4 μ M DNA, 100 mM NaCl, and 50 μ M EDTA (sodium salt) in 10 mM sodium phosphate (pH 8.0) were kept at room temperature for 15 min and then at 4 °C for several hours before data collection. Similarly, unmodified duplexes were prepared with all four different complementary sequences, replacing urea with thymine in the template strand. Thermal melting experiments were carried out in a 10 mm path length-quartz cuvette (1 mL volume) on a Varian Cary 4E spectrophotometer (Agilent Technologies, Santa Clara, California) in a buffer containing 1 M NaCl. The temperature was increased from 10 to 90 °C at a rate of 0.5 °C/min with a hold time of 2 min, and subsequently, the temperature was decreased to the starting temperature at the same rate. Absorbance was monitored at 260 nm. Melting points (T_m) were calculated from the maxima of absorbance versus temperature heating curves obtained from the first derivatives in each case. The melting experiments were repeated three times with separate sets of sample preparations, and then the T_m values were confirmed.

2.5. Crystallization of Ternary Complexes with Non-Hydrolyzable dNTP Analogs. Site-specific urea lesion-containing 12-mer template oligodeoxynucleotides corresponding to each urea-containing strand were annealed with an equimolar amount of 8-mer oligodeoxynucleotide primers (Chart 2) in 100 mM NaCl, 50 μ M EDTA (sodium salt), and 20 mM Tris-HCl (pH 8.5) at room temperature. They were kept on ice for 30 min. To first make the binary complex of hPol η and urea-containing DNA template primer, protein in 450 mM KCl, 5% glycerol, 1 mM TCEP, 3 mM DTT, and 20 mM Tris-HCl (pH 7.5), and DNA were mixed in a 1:1.1 molar ratio, followed by incubation at room temperature for 10 min. To reduce the salt concentration to 150 mM KCl, dilution buffer containing a concentration of 5 mM MgCl₂, 5% glycerol, 1 mM TCEP, 3 mM DTT, and 20 mM Tris-HCl (pH 7.5) was added. The samples were concentrated using 10

kDa cutoff amicon centrifugal filters (Millipore Sigma, Burlington, MA) to yield 2 mg/mL protein concentration. To assemble ternary complexes, dNTP analogs 2'-deoxyadenosine-5'-[(α,β)-imido]triphosphate (sodium salt; dAMPnPP), 2'-deoxycytidine-5'-[(α,β)-imido]triphosphate (sodium salt; dCMPnPP), 2'-deoxyguanosine-5'-[(α,β)-imido]triphosphate (sodium salt; dGMPnPP), and 2'-deoxythymidine-5'-[(α,β)-imido]triphosphate (sodium salt; dTMPnPP) (Jena BioScience, Jena, Germany) were added separately to a 10 mM final concentration into a tube that contained concentrated hPol η -DNA binary complex. Solutions were kept on ice for 30 min before setting up 24-well crystallization plates. For each ternary complex crystallization, we used the hanging drop vapor diffusion method in combination with a 700 mL reservoir containing polyethylene glycol monomethyl ether 2000 (PEG MME2000) (Hampton Research, Aliso Viejo, CA) (14–24%), 5 mM MgCl₂, and 0.1 M MES hydrate (Millipore Sigma, Burlington, MA) at one of three pH values of pH 5.6, pH 6.0, and pH 6.5.^{18,30} For each drop, 0.8 μ L of ternary complex was mixed with 0.8 μ L of a reservoir solution. Plates were incubated either at 18 °C or at room temperature. Crystals were observed for samples corresponding to both urea-containing DNA strands and all four dNTPs at different pH and PEG concentrations in 1 to 2 days. Diffraction-quality crystals were obtained after incubating plates for 1 week.

2.6. X-ray Diffraction Data Collection and Structure Determination/Refinement. Diffraction data were collected on beamlines 21-ID-F and 21-ID-G at LS-CAT, Advanced Photon Source (APS, Argonne National Laboratory, Argonne, IL) at wavelengths of 0.97872 and 0.97856 Å, respectively. Initial data processing, including integration and scaling, was done in HKL2000.³² The ternary structures corresponding to each analog were determined by molecular replacement with Phaser.³³ Initial rounds of rigid body refinement were followed by restrained refinement using REFMACS.³⁴ Further refinement and model building were performed with Phenix³⁵ and Coot.³⁶ Structure figures were prepared in PyMOL (The PyMOL Molecular Graphics System, Version 2.5.2, Schrödinger, LLC).

3. RESULTS

3.1. Preparation and Characterization of Site-Specific Urea Lesions. The urea-adducted oligodeoxynucleotide obtained from base hydrolysis of thymine glycol (Tg)^{3,7} containing oligodeoxynucleotides existed as an equilibrium mixture of two species (Figure 1). These could be transiently separated by RP-HPLC on a C₁₈ column at pH 8.0, but subsequently underwent re-equilibration to produce a mixture with the original composition. Both peaks were identified as urea-adducted oligodeoxynucleotides by MALDI mass spectrometry. These mixtures of two re-equilibrating species were α - and β -anomers⁷ (Scheme 1). The experimental biochemical and spectroscopic studies below were performed on samples containing equilibrated mixtures of α - and β -anomers of the urea lesion.

3.2. Trans-Lesion Synthesis by hPol η Across Urea Adducts. To investigate hPol η -mediated bypass, site-specific urea lesions were annealed with the 13-mer 5'-FAM fluorophore-labeled primer (Chart 2). Samples were prepared from both peak 1 and 2 fractions of the HPLC-purified urea-containing 18-mer DNA template (Figure 1). Individual nucleotide incorporation assays were performed in the

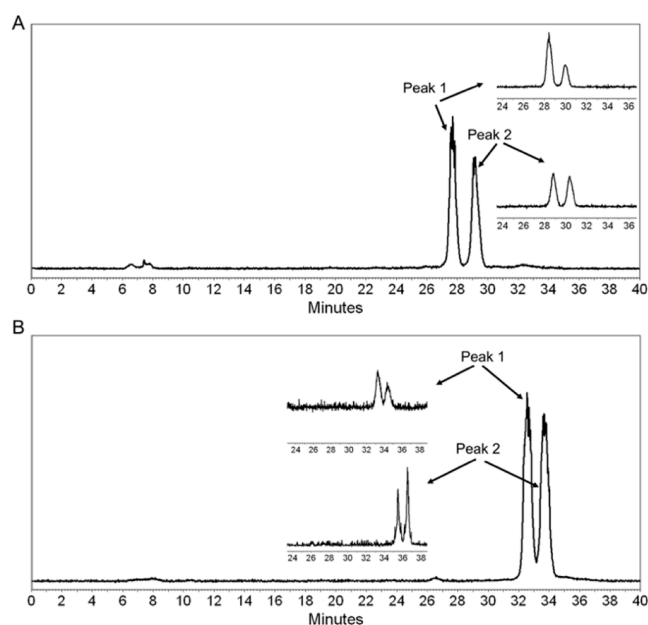


Figure 1. Reverse-phase C_{18} HPLC chromatograms monitored at 254 nm UV absorbance. (A) 12-mer oligodeoxynucleotide was used for X-ray crystallography. (B) 18-mer oligodeoxynucleotide was used for replication assays. In each panel, the two peaks correspond to the α and β anomers of the urea lesion. In each panel, the inset figures depict re-equilibration of the urea-containing oligodeoxynucleotides to mixtures of α and β anomers following reinjections of peaks 1 and 2 from the respective HPLC chromatograms.

presence of hPol η with the urea-containing template:primer duplex at 37 °C and the addition of individual dNTPs opposite the urea lesion (Figure 2). hPol η inserted the purines dATP or dGTP opposite the urea lesion with greater efficiency than it incorporated the pyrimidines dCTP or dTTP. Within 2 min of incubation with dATP or dGTP, 50–60% of the primer was utilized to form +1 dNTP addition, whereas with dCTP or dTTP, only 5–10% of the primer was utilized to form +1 dNTP addition (Figure 2). After 10 min, almost 90% of the primer was utilized using dATP or dGTP, while a significant amount of unreacted primer remained using dCTP or dTTP (Figure 2). The results in Figure 2 were obtained by using the peak 1 sample (Figure 1). Similar results were observed in dNTP incorporation using the peak 2 sample (Figure S1 in the Supporting Information).

3.3. T_m Studies. To assess the thermodynamic stability of ds oligodeoxynucleotides containing either A, C, G, or T opposite the urea lesion, site-specific urea-adducted 12-mers were separately annealed with complementary 12-mers placing either A, C, G, or T opposite urea and at equivalent concentrations. Unmodified 12-mers with T instead of urea opposite either A, C, G, or T in the complementary strand, and at equivalent concentrations, were used as controls (Chart 2). UV-absorbance-based T_m experiments were conducted at 260 nm (Figure 3). The complementary duplex having a T:A base pair showed the greatest thermal stability with a T_m value of 58 °C. Replacing the T:A base pair with an X:A base pair reduced the thermal stability of the 12-mer duplex; the X:A duplex exhibited a significantly lower T_m value of 45 °C. The duplex containing X:G exhibited a T_m value of 49 °C. In comparison, the unmodified T:G duplex also exhibited a T_m of 51 °C. The X:C and X:T duplexes exhibited T_m values of 44 and 46 °C,

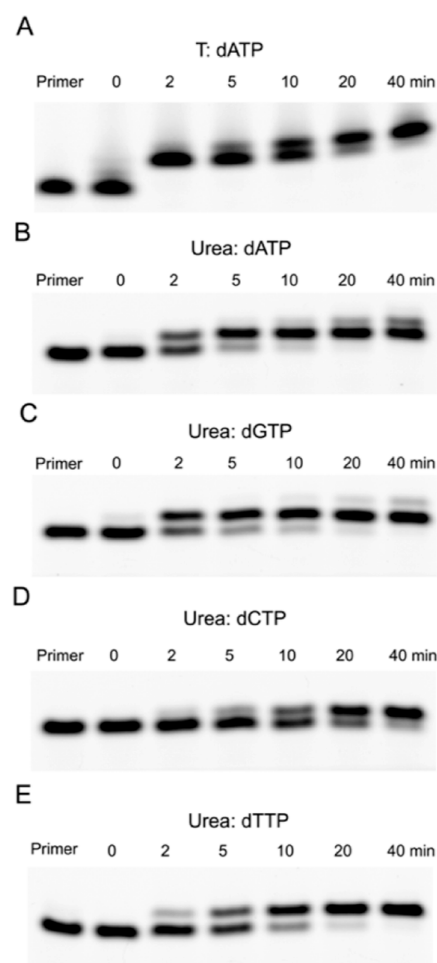


Figure 2. Individual nucleotide incorporation mediated by hPol η analyzed on 15% TBE-Urea PAGE with 7 M urea. (A) dATP incorporation opposite thymine. (B) dATP incorporation opposite the urea lesion. (C) dGTP incorporation opposite the urea lesion. (D) dCTP incorporation opposite the urea lesion. (E) dTTP incorporation opposite the urea lesion. In each assay, 10 nM hpol η was incubated with 150 nM urea-containing template:primer at 37 °C, followed by the addition of 100 μ M individual dNTPs. At the indicated time points, aliquots were removed and mixed with quencher. In each instance, the urea lesion was derived from the peak 1 HPLC eluant. In each panel, the first lane is the 5'-FAM-labeled primer band before annealing with the template.

respectively. In comparison, unmodified T:C and T:T duplexes exhibited similar T_m values of 46 and 48 °C, respectively.

3.4. Crystallization of Ternary hPol η Complexes with Non-Hydrolyzable dNTP Analogs. Viable crystals of a ternary hPol η ternary complex were obtained with each of the nonhydrolyzable dAMPnPP, dGMPnPP, dCMPnPP, and dTNPnPP analogs. Using the peak 1 HPLC fraction (Figure 1) of the site-specific template:primer urea lesion (Chart 2), crystals with incoming dAMPnPP were obtained in 22% PEG MME 2000, 5 mM $MgCl_2$, and 0.1 M 2-(*N*-morpholino)-ethanesulfonic acid (MES) (pH 6.0). These diffracted to a resolution of 2.3 Å; the space group was $P6_1$. Using the peak 2 HPLC fraction (Figure 1) of the site-specific template:primer urea lesion (Chart 2), crystals with incoming dCMPnPP were obtained in 18% PEG MME 2000, 5 mM $MgCl_2$, and 0.1 M MES (pH 5.5). Using 25% glycerol as a cryoprotectant, these crystals diffracted to a resolution of 2.2 Å; the space group was

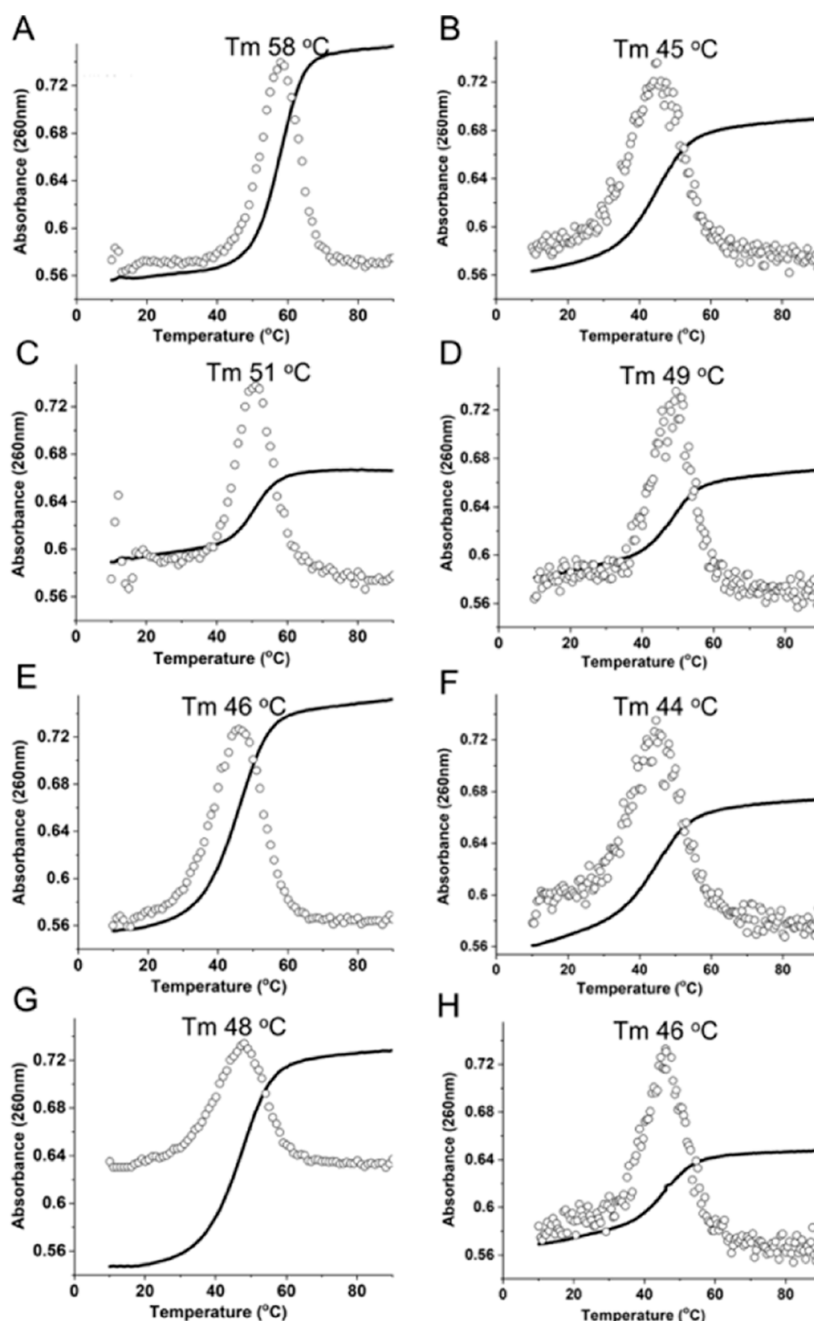


Figure 3. Thermal melting (T_m) curves were recorded by UV absorbance at 260 nm. Solid lines represent heating curves in all panels, and the corresponding first derivative curves are represented by circles. (A) Unmodified duplex 5'-d(CATTATGACGCT)-3':5'-d(AGCGTCATAATG)-3' contains a T:A base pair. (B) Modified duplex 5'-d(CATXATGACGCT)-3':5'-d(AGCGTCATAATG)-3' (X = urea lesion) containing an X:A base pair. (C) Unmodified duplex 5'-d(CATTATGACGCT)-3':5'-d(AGCGTCATCATG)-3' containing a T:C base pair. (D) Modified duplex 5'-d(CATXATGACGCT)-3':5'-d(AGCGTCATCATG)-3' (X = urea lesion) containing an X:C base pair. (E) The unmodified duplex 5'-d(CATTATGACGCT)-3':5'-d(AGCGTCATCATG)-3' containing a T:G base pair. (F) Modified duplex 5'-d(CATXATGACGCT)-3':5'-d(AGCGTCATCATG)-3' (X = urea lesion) containing an X:G base pair. (G) Unmodified duplex 5'-d(CATTATGACGCT)-3':5'-d(AGCGTCATTATG)-3' containing a T:T base pair. (H) Modified duplex 5'-d(CATXATGACGCT)-3':5'-d(AGCGTCATTATG)-3' (X = urea lesion) containing an X:T base pair.

$P6_1$. Using the peak 2 HPLC fraction (Figure 1), crystals with incoming dGMPnPP were obtained in 22% PEG MME 2000, 5 mM $MgCl_2$, and 0.1 M MES (pH 5.5). These diffracted to a resolution of 2.2 Å; the space group was $P6_1$. Using the peak 2 HPLC fraction (Figure 1), crystals with incoming dTMPnPP were obtained in 16% PEG MME 2000, 5 mM $MgCl_2$, 25% ethylene glycol, and 0.1 M MES (pH 5.6). These diffracted to a resolution of 3.02 Å; the space group was $P6_1$. Crystal data,

X-ray data collection, and refinement statistics for each of the ternary hPol η complexes are summarized in Table 1.

3.5. Structures of Ternary hPol η Complexes with Non-hydrolyzable dNTP Analogs. In DNA, urea lesions exist as mixtures of α - and β -anomers⁷ that, although chromatographically separable, re-equilibrate in solution. Because each of the ternary hpol η :DNA:dNMPnPP complexes was prepared at pH 7.5 and crystallized at acidic

Table 1. X-ray Crystallography Insertion Stage Ternary hPol η Complexes^a

complex	urea-dAMPnPP (insertion)	urea-dCMPnPP (insertion)	urea-dGMPnPP (insertion)	urea-dTMPnPP (insertion)
PDB entry	8UJT	8UJV	8UJX	8UK4
SB Grid entry	1050	1047	1048	1049
data collection				
X-ray source	APS LS-CAT	APS LS-CAT	APS LS-CAT	APS LS-CAT
wavelength [Å]	21-ID-F	21-ID-F	21-ID-G	21-ID-F
space group	P6 ₁	P6 ₁	P6 ₁	P6 ₁
unit cell				
<i>a</i> [Å]	98.99	98.95	98.93	98.80
<i>b</i> [Å]	98.99	98.95	98.93	98.80
<i>c</i> [Å]	81.82	81.5	81.6	80.3
$\alpha/\beta/\gamma$ [°]	90, 90, 120	90, 90, 120	90, 90, 120	90, 90, 120
resolution [Å]	50–2.3 ^b	50–2.23 ^b	50–2.17 ^b	50–3.02 ^b
	(2.35–2.31)	(2.27–2.23)	(2.21–2.17)	(3.07–3.02)
reflections	19,840	22,232	24,072	8880
<i>R</i> _{sym}	0.24 (1.3)	0.07 (0.4)	0.09 (0.6)	0.08 (0.67)
<i>R</i> _{pim}	0.09 (0.5)	0.04 (0.2)	0.05 (0.4)	0.04 (0.3)
<i>I</i> / σ (<i>I</i>)	29.9 (3.42)	23.6 (3.24)	19.61 (1.90)	30.48 (2.66)
completeness [%]	98.6 (100.0)	100.0 (100.0)	99.9 (100.0)	100.0 (100.0)
redundancy	8.1 (8.2)	4.1 (4.0)	4.1 (3.6)	5.7 (5.8)
CC _{1/2}	0.954 (0.65)	0.994 (0.81)	0.992 (0.68)	0.981 (0.78)
refinement				
no. of molecules per asymmetric unit	1	1	1	1
<i>R</i> _{cryst} [%]	17.8	16.6	17.1	17.4
<i>R</i> _{free} [%]	24.3	22.2	22.9	26.5
RMS deviation				
bond length [Å]	0.008	0.008	0.007	0.007
bond angles [°]	0.98	0.94	0.93	0.98
Ramachandran [%] (PROCHECK) favored	90.4	89.2	89.5	81.5
allowed	8.6	10.0	9.7	16.7
generous	1.1	0.5	0.8	1.1
outliers	0.0	0.3	0.0	0.8
<i>B</i> -factor [Å ²]	39.2	33.4	33.8	84.8
no. of atoms	3856	4018	3986	3785
no. of residues protein	431 (chain A)	435 (chain A)	435 (chain A)	431 (chain A)
DNA	19 (chains T, P)	19 (chains T, P)	19 (chains T, P)	19 (chains T, P)
water	87	231	169	13

^aCrystal data, X-ray data collection, and refinement statistics. ^bStatistics for the highest resolution shell are shown in parentheses.

pH, the observed electron density reflects an equilibrium mixture of anomeric species. At 2.2–3.0 Å resolution, the α and β anomers of urea cannot easily be distinguished. For refinements using β anomers, the electron density maps at the active site are shown in Figure 4. Structures were also refined using α anomers; the corresponding electron density maps are shown in Figure S2 in the Supporting Information. The 2Fo-Fc electron density maps computed at the 1 σ level for incoming dAMPnPP and dTMPnPP demonstrated good fits at deoxyribose C1' for both β - and α -anomers, modeled separately. The complexes with incoming dCMPnPP and dGMPnPP showed positive Fo-Fc difference electron densities (green mesh) for structures with α -anomers at deoxyribose C1', especially with incoming dCMPnPP, suggesting a better fit for the β -anomers (Figures 4 and S2 in the Supporting Information). The structures refined using the α -anomer of the urea lesion differed in the deoxyribose pucker at the urea lesion. For the α - and β -urea anomers, the deoxyribose puckers were C2'-endo and C3'-endo, respectively. A mixture of α and β urea anomers was also fitted into the electron density, which,

after refinement, suggested variable occupancies for each anomer dependent upon the identity of the incoming nucleotide. For incoming dAMPnPP, an occupancy of 64% α vs 36% β was calculated. For incoming dGMPnPP, an occupancy of 14% α vs 86% β was calculated. For incoming dCMPnPP, an occupancy of 100% β was calculated. For incoming dTMPnPP, an occupancy of 33% α vs 67% β was calculated (Figure S3 in the Supporting Information). This suggested that the peak 1 HPLC fraction used for generating crystals consisted of the α anomer and the peak 2 HPLC fraction used for generating crystals (Figure 1) consisted of the β -anomer, but that once isolated, each of these peaks re-equilibrated to an anomeric mixture.

3.5.1. Incoming dAMPnPP Opposite Urea. This structure was solved by molecular replacement using the structure with the PDB ID code SF9L³⁷ as the search model. For the refinement using the β anomer, the electron density map at the active site is shown in Figure 4. The structure was also refined using the α anomer; the corresponding electron density map is shown in Figure S2 in the Supporting Information. Figure 5

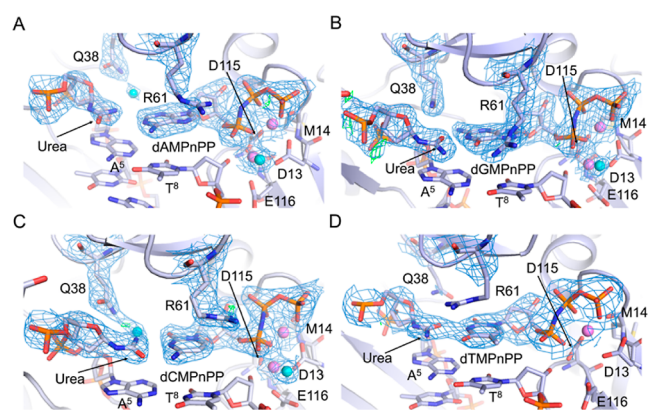


Figure 4. 2Fo-Fc Fourier sum electron density maps (light blue mesh; contoured at 1σ) at the active site region for the insertion-stage hPol η ternary complexes with urea-containing template:primer DNA duplexes and incoming nonhydrolyzable dNMPnPPs with the β -deoxyribose anomeric carbon configuration of the urea lesions. The Fo-Fc Fourier difference electron density maps (contoured at 3σ) (green mesh) are highlighted as small patches. (A) dAMPnPP, (B) dGMPnPP, (C) dCMPnPP, and (D) dTMPnPP. The color of carbon atoms is light blue throughout proteins, DNA chains, urea lesions, and incoming nucleotides. The coordination of either two Mg^{2+} ions or one Mg^{2+} ion (pink spheres) is shown. Water molecules are shown as cyan spheres.

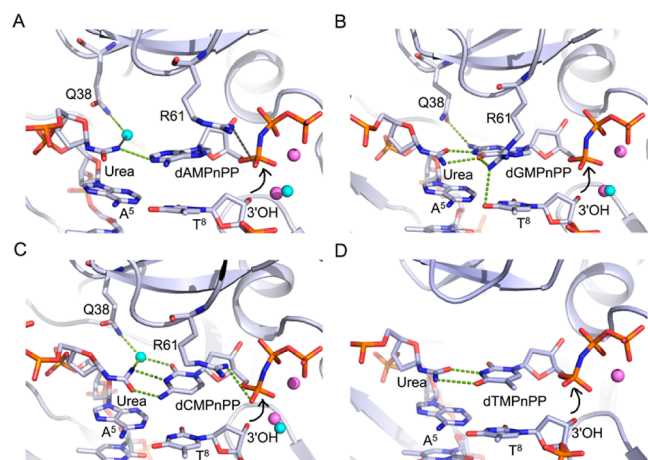


Figure 5. View of the active site interactions between urea lesions and incoming dNMPnPPs as well as the selected protein side chains and the urea-incoming dNMPnPP pairs in the insertion-stage hPol η ternary complexes. The proximity of the 3'-OH group of the primer terminal thymine T^8 to the α -phosphate P atom of the incoming dNMPnPPs for the nucleophilic attack is shown as a curved arrow symbol. (A) Urea lesion-dAMPnPP, (B) Urea lesion-dGMPnPP, (C) Urea lesion-dCMPnPP, and (D) Urea lesion-dTMPnPP. Mg^{2+} ions are pink spheres, and water molecules are cyan spheres. In each instance, the structures were refined using the β -anomer of the deoxyribose C1' carbon.

shows the orientation of the incoming dAMPnPP within the active site. At the primer 3' terminal nucleotide T^8 (Chart 2), the 3'-hydroxyl was 3.4 Å from the α -phosphate of the incoming dAMPnPP. Arg61 was ordered, as indicated by the electron density around its guanidino group (Figure 4). The Arg61 guanidino group faced the DNA major groove and was 4.0 Å from the α -phosphate oxygen O1A of dAMPnPP. The amide nitrogen of Gln38 was 3.6 Å from the deoxyribose O4' of the urea lesion in the minor groove. The Arg55 and Tyr52

side chains were 3.2 and 2.6 Å, respectively, from the γ -phosphodiester oxygens O1G and O2G of incoming dAMPnPP, and thus within the H-bonding range. Like in other ternary structures of hPol η ,^{18,30} two catalytic Mg^{2+} ions coordinated to the Asp115, Glu116, and Asp13 side chains and the Met14 backbone oxygen, water, the primer terminal T^8 3'-OH, and α -, β -, and γ -phosphate oxygens of incoming dAMPnPP. Detailed active-site interactions for all the ternary complexes are shown in Figure S4 in the Supporting Information. At this resolution, the urea amino nitrogen and carbonyl oxygen could not be distinguished in the electron density. However, the Gln38 carbonyl group engages in an H-bond with water (2.8 Å distance). This water was positioned 4.7 Å from the electron density arising from the urea moiety, which was assigned as the urea amino group. This placed the urea carbonyl group 4.6 Å from N^6H_2 of the incoming dAMPnPP. Consequently, the urea moiety was oriented such that its amino group was 2.8 Å from the dAMPnPP N1 imino nitrogen, consistent with the formation of a H-bond (Chart 1). Corresponding H-bonding interactions between urea and incoming nucleotides with α -configurations of the deoxyribose are shown in Figure S5 in the Supporting Information. The distance between the imino nitrogen of urea and dAMPnPP N9 was 8.7 Å (Figure 6).

3.5.2. Incoming dCMPnPP Opposite Urea. The structure was solved by molecular replacement using the structure with the PDB ID code 5F9N³⁷ as the search model. For the refinement using the β anomer, the electron density map at the active site is shown in Figure 4. The structure was also refined using the α anomer; the corresponding electron density map is

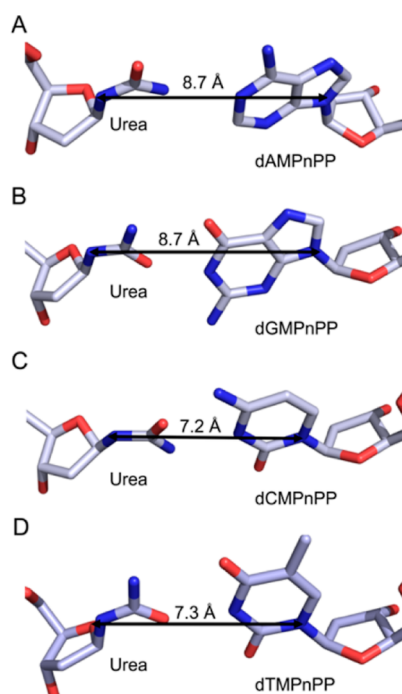


Figure 6. Expanded views of the base pair geometry at the active site in the insertion-stage hPol η ternary complexes and depicting urea imine-purine N9/pyrimidine N1 atom distances. (A) Urea lesion with incoming dAMPnPP. (B) Urea lesion with incoming dGMPnPP. (C) Urea lesion with incoming dCMPnPP. (D) Urea lesion with incoming dTMPnPP. In each instance, the structures were refined using the β -anomer of the deoxyribose C1' carbon. The carbon atoms are colored light blue.

shown in Figure S2 in the Supporting Information. At the primer-terminal nucleotide T⁸ (Chart 2), the 3'-hydroxyl was 3.5 Å from the α -phosphate of the incoming dCMPnPP. Arg61 was ordered, as indicated by the electron density surrounding its guanidino moiety (Figure 4). The NH1 moiety of the guanidino group was oriented into the major groove. It was 3.5 Å from the α -phosphodiester oxygen O1A and 3.6 Å from OS' of the incoming dCMPnPP. The side chain carbonyl of Gln38 made a 2.7 Å H-bond with water, which in turn participated in a 2.8 Å H-bond to the urea amino group. This water was also 2.9 Å from the O² of the incoming dCMPnPP. The Gln38 amide was 3.7 Å from urea O4'. Two catalytic Mg²⁺ ions were located at the active site. At this resolution of the electron density, the urea amino nitrogen and carbonyl oxygen could not be distinguished. However, the electron density occupied by the urea moiety was 3.7 Å from O² of the incoming dCMPnPP, prompting us to assign the urea density closest to the oxygen to the amino group. Consequently, the urea amino group and carbonyl oxygen were 2.9 and 2.5 Å from the N3 atom and N⁴H₂ amino group of the incoming dCMPnPP, respectively, and consistent with the formation of H-bonds (Chart 1). The distance between the imino nitrogen of urea and the N1 of dCMPnPP was 7.2 Å (Figure 6).

3.5.3. Incoming dGMPnPP Opposite Urea. This structure was solved by molecular replacement using the structure with the PDB ID code 4O3Q³⁰ as the search model. For the refinement using the β anomer, the electron density map at the active site is shown in Figure 4. The structure was also refined using the α anomer; the corresponding electron density map is shown in Figure S2 in the Supporting Information. At the primer terminal nucleotide T⁸ (Chart 2), the 3'-hydroxyl was 3.5 Å from the α -phosphate of the incoming dGMPnPP. The side chain of Arg61 was ordered in the electron density (Figure 4) and NH1 of its guanidino group was oriented toward the major groove and engaged in a 3.0 Å H-bond with O⁶ of the incoming dGMPnPP and a 3.2 Å H-bond with the O⁴ of the primer 3' terminal thymine. It did not interact with α -phosphate oxygen O1A of the incoming dGMPnPP, as seen in the dAMPnPP- and dCMPnPP-bound structures. The Gln38 carbonyl made a 3.3 Å H-bond with the amino group of the incoming dGMPnPP. Two catalytic Mg²⁺ ions were identified. At this resolution of the electron density, the urea amino nitrogen and carbonyl oxygen could not be distinguished. However, the electron density from the urea was 3.8 Å from the N² amino group of dGMPnPP, prompting us to assign the urea density closest to the amino group to the carbonyl oxygen. The urea carbonyl oxygen and the amino group were 2.4 and 2.9 Å from the N1 imino nitrogen and the O⁶ carbonyl oxygen of dGMPnPP, respectively, consistent with the formation of H-bonds (Chart 1). The distance between the urea imino group and N9 of dGMPnPP was 8.7 Å (Figure 6).

3.5.4. Incoming dTMPnPP Opposite Urea. This structure was solved by molecular replacement using the structure with PDB ID code 5L1J³⁸ as the search model. For the refinement using the β anomer, the electron density map at the active site is shown in Figure 4. The structure was also refined using the α anomer; the corresponding electron density map is shown in Figure S2 in the Supporting Information. At the primer terminal nucleotide T⁸ (Chart 2), the 3'-hydroxyl was 3.5 Å from the α -phosphate of the incoming dTMPnPP. The electron density at the active site suggested disorder in this structure; the electron density for Arg61 and Gln38 side chains was missing (Figure 4). Residues Arg55 and Tyr52 were 3.7

and 2.6 Å, respectively, from the γ -phosphate oxygens of the incoming dTMPnPP (Figure 5). In this structure, only one catalytic Mg²⁺ ion could be identified at the electron density. Its coordination sphere included Asp115, Asp13, and Met14 side chains and the α -, β -, and γ -phosphate oxygens of the incoming dTMPnPP. At this resolution of the electron density, the urea amino nitrogen and carbonyl oxygen could not be distinguished. We assigned their identities such that the urea amino lies at 3.2 Å from O⁴ of dTMPnPP and the urea carbonyl oxygen lies at 2.6 Å distance from the N3 imino nitrogen of incoming dTMPnPP, consistent with H-bonding (Chart 1). The distance between the urea imino group and N1 of dTMPnPP was 7.3 Å (Figure 6).

4. DISCUSSION

Translesion synthesis (TLS, Y-family) DNA polymerases show reduced fidelity, catalytic efficiency, and processivity of DNA synthesis^{17,22,39} compared with high-fidelity polymerases.^{22,40} Y-family polymerases generally participate in the first step of DNA synthesis opposite lesions.^{22,41} Their solvent-exposed active sites^{16,18,26} result in a loose fit of template residues and may result in the incorporation of incorrect nucleotides during DNA synthesis.²² A second polymerase from the B-family often engages to extend the primer beyond the lesion or mismatched insertion base pair up to the next few template bases before high-fidelity polymerases resume normal DNA synthesis.^{22,42,43} The use of nonhydrolyzable nucleoside triphosphate analogues, in which an amine bridges the α - and β -phosphates, permits cocrystallization of TLS polymerase hPol η with oligodeoxynucleotide template-primer pairs in the presence of Mg²⁺ without the use of a primer with a 3'-terminal 2',3'-dideoxynucleoside and/or divalent metal ions that are not catalytically competent, enabling structural investigations of lesion bypass.

4.1. Individual Nucleotide Incorporation Opposite Urea Lesion. Nucleotide incorporation assays (Figure 2) confirm that hPol η bypasses urea lesions⁹ and inserts all four dNTPs opposite these lesions. They also reveal that the purines dATP and dGTP are preferentially inserted as compared to the pyrimidines dCTP and dTTP. This is consistent with a prior report that hPol η follows a "purine rule", i.e., preferentially incorporating dATP or dGTP at opposite AP sites.¹² Kawada et al.⁹ reported incorporation of all four dNTPs opposite urea lesions by hPol η , also with greater efficiency of dNTP incorporation with the purines dGTP and dATP vs pyrimidine dTTP, but reported higher incorporation of dCTP. They prepared site-specific urea lesions by oxidizing oligodeoxynucleotides containing 8-oxo-dG to oxaluric acid (Oxa) lesions,⁴⁴ followed by hydrolysis to urea lesions⁴ (Scheme 1).

4.2. Structures of Ternary hPol η Complexes with Non-Hydrolyzable dNTP Analogs. Our structures are consistent with the observation that hPol η incorporates each of the four dNTPs opposite urea lesions (Figure 2). In each of the four ternary complexes, the primer 3'-OH is positioned to attack the α -phosphoryl group (Figure 5). However, incoming purines are positioned differently than incoming pyrimidines. In the ternary complexes with dAMPnPP or dGMPnPP, the urea lesion is not completely coplanar but slightly above the nucleobase plane of the incoming dNMPnPP. With incoming pyrimidines, the urea lesion is coplanar with the nucleobase plane of dCMPnPP and dTMPnPP. Active site residue Arg61 is thought to play a key role in phosphoryl transfer.^{18,31,37,45} In

the absence of an incoming dNTP in the hPol η active site, Arg61 projects toward the nucleobase of the 5'-overhanging residue of the template strand. It is thought that binding of incoming dNTP and Mg^{2+} ions facilitates dNTP-specific conformational changes such that in the bound complexes, Arg61 interacts either with the α -phosphate of the incoming dNTP or the incoming nucleobase to facilitate phosphoryl transfer. In our structures, the position of Arg61 also varies. With incoming dAMPnPP, the Arg61 guanidino side chain is directed toward the α -phosphate oxygen O1A of the incoming nucleotide analog. With dCMPnPP, the Arg61 guanidino side chain NH1 is 3.5 Å from the α -phosphate oxygen O1A and 3.6 Å from OS' of the incoming nucleotide analog. In contrast, with incoming dGMPnPP, the Arg61 guanidino group interacts with O⁴ of the primer 3' terminal thymine (Figure 4) instead of the α -phosphate oxygen O1A of the incoming nucleotide analog. With incoming TMPnPP, the electron density (Figure 4) and high *B*-factor (Table 1) indicated a disordered arrangement of the Arg61 guanidino moiety, possibly due to steric interference from the thymine methyl group. Also, active site residue Gln38 is thought to help in the stabilization of the template base or incoming dNTP, with or without water-mediated interactions.^{18,31,37,38} In our structures, Gln38 interacts with either the urea lesion or the incoming nucleotide. With incoming dAMPnPP, Gln38 does not interact directly with the urea lesion but instead participates in a 2.8 Å H-bond with water, which in turn is 4.7 Å from the urea amino group. In another scenario, with incoming dGMPnPP, the Gln38 carbonyl forms a 3.3 Å H-bond with the amino group of the incoming nucleotide analog (Figure 6). With incoming dCMPnPP, the Gln38 carbonyl forms a 2.7 Å H-bond with water, which in turn forms a 2.8 Å H-bond to the urea amino group. This pulls the urea moiety toward the minor groove and thereby closer to the pyrimidine, enabling pairing between the two (Figure 5). With incoming TMPnPP, the lack of electron density for the Gln38 side chain indicates a disordered conformation at the active site (Figure 4), consistent with an average *B*-factor of 85 Å (Table 1). Only one Mg^{2+} ion is coordinated at the active site.

The favorable insertion of purines may be associated with the base pairing geometry in the active site. In a canonical B-DNA base pair, the distance between the N9 atom of a purine and the N1 atom of the complementary pyrimidine is approximately 9 Å. With incoming dAMPnPP or dGMPnPP, the 8.7 Å distance observed between the imino nitrogen of the urea and the N9 atom of the incoming purine is close to the canonical 9 Å distance (Figure 6).²⁹ In contrast, for ternary complexes featuring incoming pyrimidines, the distances between the imino group of the urea and the N1 atom of either dCMPnPP or dTMPnPP are shorter, i.e., at 7.2 and 7.3 Å, respectively (Figure 6). The more favorable incorporation of purines opposite the urea lesion may also be due, in part, to differential stacking in the ternary complexes. Despite the reduced thermal stabilities of urea-modified DNA duplexes with A, C, G, or T opposite the lesion in the complementary strand compared to the parent duplex with a T:A base pair (Figure 3) (which can be attributed to the urea moiety lacking a π -surface), Figure 7 shows that at the template:primer terminus, all four incoming dNMPnPPs stack with the 3'-flanking A⁵:T⁸ base pair (Chart 1). The most favorable stacking is seen for incoming dAMPnPP, whereas incoming dGMPnPP shifts toward the minor groove and shows reduced stacking with 3'-flanking T⁸, but still stacks with the 3'-flanking

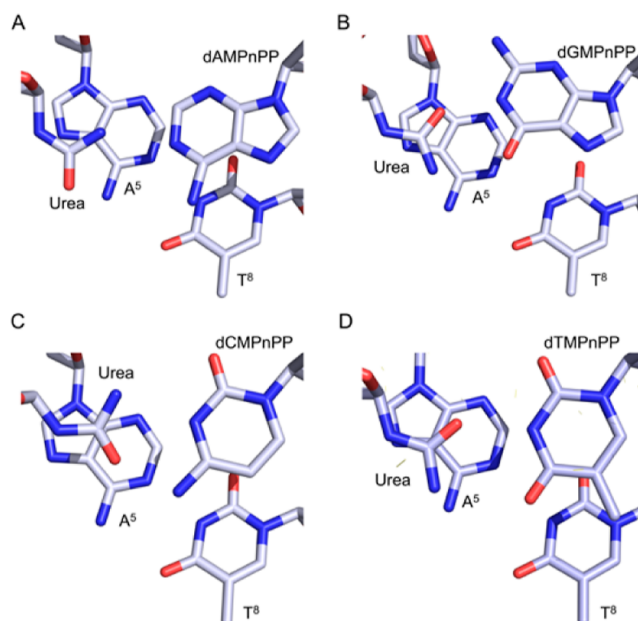


Figure 7. Stacking interactions between the urea-dNMPnPP pair and the 3'-adjacent A⁵:T⁸ pair at the active site of ternary complexes of hPol η , viewed from the top side of the urea-dNMPnPP pair. (A) Urea lesion with incoming dAMPnPP. (B) Urea lesion with incoming dGMPnPP. (C) Urea lesion with incoming dCMPnPP. (D) Urea lesion with incoming dTMPnPP. The carbon atoms are shown as light blue for urea, incoming dNMPnPPs as well as the A⁵:T⁸ pair.

A⁵ purine (Figure 8). This is due to an H-bond between the guanine exocyclic amino group and the Gln38 side chain

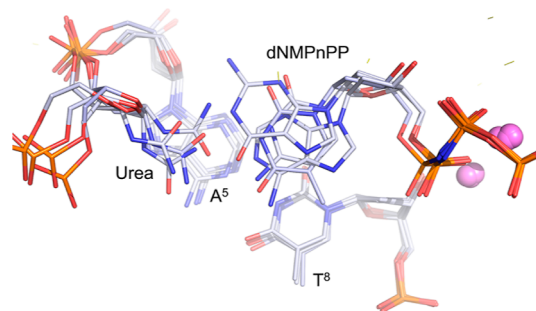


Figure 8. Comparison of urea-dNMPnPPs stacking in the hPol η ternary complexes at the active sites with reference to the template A⁵:T⁸ pair. The terminal nucleotide of the primer reveals differences in the orientation of the urea moiety with β -deoxyribose configurations and incoming nucleotides (viewed from the top side of the urea-dNMPnPP pairs). The purine ring in the urea-dGMPnPP pair is shifted into the minor groove as the N²H₂ group of guanine H-bonds with the Gln38 side chain. The urea-dCMPnPP pair is pulled into the minor groove through a water-mediated H-bond with the Gln38 side chain. The site of the nucleotidyl transfer reaction is overlaid, including Mg^{2+} (pink spheres)-mediated coordination in all the complexes. The carbon atoms are shown as light blue for urea, incoming dNMPnPPs, as well as the A⁵:T⁸ pair.

(Figure 5). The incoming pyrimidines dCMPnPP and dTMPnPP exhibit less favorable stacking with the 3'-flanking A⁵:T⁸ base pair, as revealed from superimposed ternary complex structures (Figures 7 and 8). With the incoming dCMPnPP, the urea amino group interacts with Gln38 via a water-mediated H-bond. Arg61-mediated positioning of

dCMPnPP O^2 2.9 Å away from this water pulls incoming dCMPnPP toward the minor groove, thereby reducing stacking (Figures 5 and 7). The lack of a second Mg^{2+} ion is consistent with the reduced incorporation of dTTP by hPol η compared with the other nucleotides.

4.3. Urea Lesions as Templates. The present structures reveal the ability of hPol η to accommodate each of the nonhydrolyzable dNTP analogs dAMPnPP, dCMPnPP, dGMPnPP, and dTMPnPP opposite the urea lesion. The present study confirms predictions based on MD simulations in dsDNA.²⁹ The latter predicted the formation of at least one H-bond for urea–purine pairs, whereas for urea–pyrimidine pairs, apart from a H-bond, a water-mediated interaction was proposed.²⁹ The ability of urea lesions to H-bond with each of the four incoming dNTPs stems from their ability to undergo conformational interconversion around the nitrogen–carbon bond between N1 and carbonyl carbon. This likely possesses some double bond character due to the potential for imino-amino tautomerization of the urea moiety. This conformational interconversion reorients the amino (H-bond donor) and carbonyl (H-bond acceptor) groups of urea, allowing the urea moiety to base pair with all four incoming dNTPs (Chart 2). Our data reveal that the urea lesion forms H-bonds with the Watson–Crick face of each of the incoming dNTPs (Figure 5). The ternary complex with dAMPnPP displays a single H-bond between urea and N1 of adenine, whereas with the dGMPnPP structure, there are two H-bonds with the N1H and carbonyl oxygen of guanine opposite the carbonyl and amino groups of urea, respectively. The ternary complex with dCMPnPP reveals two H-bonds linking the urea amino and carbonyl groups to N3 and N^4H_2 of dCMPnPP, respectively. Similarly, the dTMPnPP-bound structure reveals two H-bonds between the urea carbonyl oxygen and N3H of thymine and the urea amino group and O^4 of dTMPnPP (Figure 5). Of interest are the observations that B-family high-fidelity DNA polymerases δ and ϵ are blocked by urea lesions, while DNA polymerase α extends the primer with very low efficiency.⁹ One feature of urea lesions is their small size, i.e., sterically smaller than a pyrimidine base and perhaps more like abasic sites. It seems possible that steric fit requirements for incoming dNTPs by high-fidelity polymerases cannot be satisfied by urea lesions, whereas, with hpol η , the more spacious and sterically relaxed active site accommodates all four incoming nucleotides. The present results reveal the ability of urea to form base pairing interactions with all four incoming dNTPs in the case of hpol η .

4.4. Implications for Mutagenesis. The ability of urea lesions to act as templates for all four incoming dNTPs is consistent with the observation that they are mutagenic. Like AP sites, urea lesions lack the original coding information. Our data demonstrate preferential incorporation of purines dATP and dGTP opposite urea lesions (Figure 2). This is consistent with the observations of urea-derived mutagenic spectra in site-directed mutagenesis assays. Maccabbe et al.⁴⁶ reported that 62% of urea-derived mutations initially arising from thymine glycol-modified sites were T-to-C transitions, mostly distributed in sequence contexts such as TTA and TTC. While the preferential insertion of dATP opposite urea lesions arising from rearrangement of thymine glycol lesions (Scheme 1) is anticipated to result in error-free bypass, preferential incorporation of dGTP opposite deoxythymidine-derived urea lesions would result in T to C transitions. Henderson et al.⁴ showed that in bacteria, urea lesions derived from guanine

oxidation and subsequent hydrolysis (Scheme 1) generated 43% G to T and 46% G to C mutations. In this instance, preferential incorporation of either dATP or dGTP opposite the lesions would account for the observed mutations.

5. CONCLUSIONS

The Y-family TLS polymerase hPol η inserts all four dNTPs opposite the urea lesion; purines are incorporated better than pyrimidines. X-ray crystallographic data provided experimental evidence for the H-bond-mediated base-pairing interactions between the urea lesion and all four incoming dNMPnPPs. The structural analyses of ternary complexes suggested that more optimal base pairing geometries in the urea–purine pairs and improved base stacking interactions are most likely responsible for the preferential misincorporation of purines opposite the urea lesion.

■ ASSOCIATED CONTENT

Data Availability Statement

Atomic coordinates and structure factors for the reported crystal structures have been deposited with the Protein Data Bank under accession numbers dAMPnPP, PDB ID 8UJT; dCMPnPP, PDB ID 8UJV; dGMPnPP, PDB ID 8UJX; dTMPnPP, PDB ID 8UK4. Corresponding diffraction data sets have been deposited with SBGrid under accession codes-1050 (doi:10.15785/SBGRID/1050); 1047 (doi:10.15785/SBGRID/1047); 1048 (doi:10.15785/SBGRID/1048); 1049 (doi:10.15785/SBGRID/1049), respectively.

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.biochem.3c00569>.

Individual-nucleotide incorporation opposite the urea lesion using peak 2 HPLC sample, mediated by hPol η ; 2Fo-Fc Electron density maps contoured at 1σ at the active site region for hPol η ternary complexes with incoming nonhydrolyzable nucleotides with α -deoxyribose configurations; 2Fo-Fc electron density maps (1σ) showing model fit for the mixture of α - and β -anomer deoxyribose configurations of urea lesion in the hPol η ternary complexes; crystal structures of hPol η in complex with urea-containing template:primer DNA duplexes and incoming nonhydrolyzable nucleotides showing detailed interactions with β -configurations; and expanded views at the active site in the insertion-stage hPol η ternary complexes, depicting H-bonding between urea-dNMPnPP pairs with α -deoxyribose configurations (PDF)

Accession Codes

Human polymerase η . UniProt Q9Y253

■ AUTHOR INFORMATION

Corresponding Authors

Martin Egli – Department of Biochemistry, School of Medicine, Vanderbilt Ingram Cancer Center, and Vanderbilt Center for Structural Biology, Vanderbilt University, Nashville, Tennessee 37232, United States; orcid.org/0000-0003-4145-356X; Phone: +1 615-343-8070; Email: michael.p.stone@vanderbilt.edu

Michael P. Stone – Department of Chemistry, Vanderbilt Ingram Cancer Center, and Vanderbilt Center for Structural Biology, Vanderbilt University, Nashville, Tennessee 37235, United States; orcid.org/0000-0002-0922-0216;

Phone: +1 615-322-2589; Email: martin.egli@vanderbilt.edu

Authors

Rachana Tomar – Department of Chemistry, Vanderbilt Ingram Cancer Center, and Vanderbilt Center for Structural Biology, Vanderbilt University, Nashville, Tennessee 37235, United States

Songlin Li – Department of Chemistry, Vanderbilt Ingram Cancer Center, and Vanderbilt Center for Structural Biology, Vanderbilt University, Nashville, Tennessee 37235, United States

Complete contact information is available at:

<https://pubs.acs.org/10.1021/acs.biochem.3c00569>

Funding

This work was supported by NIH grants R01 ES-029357 and P01 CA-160032 (M.P.S., M.E.). The Vanderbilt-Ingram Cancer Center was funded by NIH grant P30 CA-068485. Vanderbilt University and the Vanderbilt Center for Structural Biology assisted with the purchase of in-house crystallographic instrumentation. Crystallographic data were collected on the 21-ID-F and 21-ID-G beamlines of the Life Sciences Collaborative Access Team (LS-CAT) at the Advanced Photon Source (Argonne National Laboratory, Argonne, IL). Supporting institutions may be found at <http://ls-cat.org/members.html>. The use of the Advanced Photon Source was supported by the U.S. Department of Energy, Basic Energy Sciences, Office of Science, under Contract W-31109-Eng-38. Funding for open access charge: National Institutes of Health.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank Dr. Michelle L. Reyzer for the assistance in MALDI mass spectrometry experiments. We thank Dr. Pankaj Sharma for helpful suggestions in structure data processing.

ABBREVIATIONS

AP site, apurinic/apyrimidinic site; Tg, thymine glycol; RP-HPLC, reverse-phase high-performance liquid chromatography

REFERENCES

- (1) Ide, H.; Kow, Y. W.; Wallace, S. S. Thymine glycols and urea residues in M13 DNA constitute replicative blocks in vitro. *Nucleic Acids Res.* **1985**, *13*, 8035–8052.
- (2) Evans, J.; Maccabee, M.; Hatahet, Z.; Courcelle, J.; Bockrath, R.; Ide, H.; Wallace, S. Thymine ring saturation and fragmentation products: lesion bypass, misinsertion and implications for mutagenesis. *Mutat. Res.* **1993**, *299*, 147–156.
- (3) McNulty, J. M.; Jerkovic, B.; Bolton, P. H.; Basu, A. K. Replication inhibition and miscoding properties of DNA templates containing a site-specific cis-thymine glycol or urea residue. *Chem. Res. Toxicol.* **1998**, *11*, 666–673.
- (4) Henderson, P. T.; Neeley, W. L.; Delaney, J. C.; Gu, F.; Niles, J. C.; Hah, S. S.; Tannenbaum, S. R.; Essigmann, J. M. Urea lesion formation in DNA as a consequence of 7,8-dihydro-8-oxoguanine oxidation and hydrolysis provides a potent source of point mutations. *Chem. Res. Toxicol.* **2005**, *18*, 12–18.
- (5) Guy, A.; Ahmad, S.; Téoule, R. Insertion of the fragile 2'-deoxyribosylurea residue into oligodeoxynucleotides. *Tetrahedron Lett.* **1990**, *31*, 5745–5748.
- (6) Dubey, I.; Pratiel, G.; Robert, A.; Meunier, B. Convenient method for the preparation of 2'-deoxyribosylurea by thymidine

oxidation and NMR study of both anomers. *Nucleosides, Nucleotides Nucleic Acids* **2001**, *20*, 1463–1471.

- (7) Tomar, R.; Minko, I. G.; Sharma, P.; Kellum, A. H.; Lei, L.; Harp, J. M.; Iverson, T. M.; Lloyd, R. S.; Egli, M.; Stone, M. P. Base excision repair of the N-(2-deoxy-d-erythro-pentofuranosyl)-urea lesion by the hNEIL1 glycosylase. *Nucleic Acids Res.* **2023**, *51*, 3754–3769.

- (8) Breimer, L.; Lindahl, T. A DNA glycosylase from *Escherichia coli* that releases free urea from a polydeoxyribonucleotide containing fragments of base residues. *Nucleic Acids Res.* **1980**, *8*, 6199–6211.

- (9) Kawada, T.; Kino, K.; Tokorodani, K.; Anabuki, R.; Morikawa, M.; Kobayashi, T.; Ohara, K.; Ohshima, T.; Miyazawa, H. Analysis of nucleotide insertion opposite urea and translesion synthesis across urea by DNA polymerases. *Genes Environ.* **2022**, *44*, 7.

- (10) Strauss, B. S. The "A" rule revisited: polymerases as determinants of mutational specificity. *DNA Repair* **2002**, *1*, 125–135.

- (11) Choi, J. Y.; Lim, S.; Kim, E. J.; Jo, A.; Guengerich, F. P. Translesion Synthesis across Abasic Lesions by Human B-Family and Y-Family DNA Polymerases α , δ , η , ι , κ , and REV1. *J. Mol. Biol.* **2010**, *404*, 34–44.

- (12) Patra, A.; Zhang, Q.; Lei, L.; Su, Y.; Egli, M.; Guengerich, F. P. Structural and Kinetic Analysis of Nucleoside Triphosphate Incorporation Opposite an Abasic Site by Human Translesion DNA Polymerase η . *J. Biol. Chem.* **2015**, *290*, 8028–8038.

- (13) Schaaper, R. M.; Glickman, B. W.; Loeb, L. A. Mutagenesis resulting from depurination is an SOS process. *Mutat. Res.* **1982**, *106*, 1–9.

- (14) Boiteux, S.; Laval, J. Coding properties of poly(deoxycytidylic acid) templates containing uracil or apyrimidinic sites: in vitro modulation of mutagenesis by deoxyribonucleic acid repair enzymes. *Biochemistry* **1982**, *21*, 6746–6751.

- (15) Efrati, E.; Tocco, G.; Eritja, R.; Wilson, S. H.; Goodman, M. F. Abasic Translesion Synthesis by DNA Polymerase β Violates the "A-rule". *J. Biol. Chem.* **1997**, *272*, 2559–2569.

- (16) Yang, W.; Woodgate, R. What a difference a decade makes: insights into translesion DNA synthesis. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 15591–15598.

- (17) Sale, J. E.; Lehmann, A. R.; Woodgate, R. Y-family DNA polymerases and their role in tolerance of cellular DNA damage. *Nat. Rev. Mol. Cell Biol.* **2012**, *13*, 141–152.

- (18) Biertumpfel, C.; Zhao, Y.; Kondo, Y.; Ramon-Maiques, S.; Gregory, M.; Lee, J. Y.; Masutani, C.; Lehmann, A. R.; Hanaoka, F.; Yang, W. Structure and mechanism of human DNA polymerase η . *Nature* **2010**, *465*, 1044–1048.

- (19) Suarez, S. C.; Beardslee, R. A.; Toffton, S. M.; McCulloch, S. D. Biochemical analysis of active site mutations of human polymerase η . *Mutat. Res.* **2013**, *745–746*, 46–54.

- (20) Patra, A.; Banerjee, S.; Johnson Salyard, T. L.; Malik, C. K.; Christov, P. P.; Rizzo, C. J.; Stone, M. P.; Egli, M. Structural Basis for Error-Free Bypass of the 5-N-Methylformamidopyrimidine-dG Lesion by Human DNA Polymerase η and *Sulfolobus solfataricus* P2 Polymerase IV. *J. Am. Chem. Soc.* **2015**, *137*, 7011–7014.

- (21) Su, Y.; Egli, M.; Guengerich, F. P. Mechanism of Ribonucleotide Incorporation by Human DNA Polymerase η . *J. Biol. Chem.* **2016**, *291*, 3747–3756.

- (22) Yang, W. An Overview of Y-Family DNA Polymerases and a Case Study of Human DNA Polymerase η . *Biochemistry* **2014**, *53*, 2793–2803.

- (23) Masutani, C.; Kusumoto, R.; Iwai, S.; Hanaoka, F. Mechanisms of accurate translesion synthesis by human DNA polymerase η . *EMBO J.* **2000**, *19*, 3100–3109.

- (24) Haracska, L.; Yu, S. L.; Johnson, R. E.; Prakash, L.; Prakash, S. Efficient and accurate replication in the presence of 7,8-dihydro-8-oxoguanine by DNA polymerase η . *Nat. Genet.* **2000**, *25*, 458–461.

- (25) Masutani, C.; Kusumoto, R.; Yamada, A.; Dohmae, N.; Yokoi, M.; Yuasa, M.; Araki, M.; Iwai, S.; Takio, K.; Hanaoka, F. The XPV (xeroderma pigmentosum variant) gene encodes human DNA polymerase η . *Nature* **1999**, *399*, 700–704.

- (26) Silverstein, T. D.; Johnson, R. E.; Jain, R.; Prakash, L.; Prakash, S.; Aggarwal, A. K. Structural basis for the suppression of skin cancers by DNA polymerase η . *Nature* **2010**, *465*, 1039–1043.
- (27) Johnson, R. E.; Kondratick, C. M.; Prakash, S.; Prakash, L. hRAD30 mutations in the variant form of xeroderma pigmentosum. *Science* **1999**, *285*, 263–265.
- (28) Carlson, K. D.; Washington, M. T. Mechanism of Efficient and Accurate Nucleotide Incorporation Opposite 7,8-Dihydro-8-Oxoguanine by *Saccharomyces cerevisiae* DNA Polymerase η . *Mol. Cell. Biol.* **2005**, *25*, 2169–2176.
- (29) Suresh, G.; Padhi, S.; Patil, I.; Priyakumar, U. D. Urea Mimics Nucleobases by Preserving the Helical Integrity of B-DNA Duplexes via Hydrogen Bonding and Stacking Interactions. *Biochemistry* **2016**, *55*, 5653–5664.
- (30) Patra, A.; Nagy, L. D.; Zhang, Q.; Su, Y.; Muller, L.; Guengerich, F. P.; Egli, M. Kinetics, Structure, and Mechanism of 8-Oxo-7,8-dihydro-2'-deoxyguanosine Bypass by Human DNA Polymerase η . *J. Biol. Chem.* **2014**, *289*, 16867–16882.
- (31) Su, Y.; Patra, A.; Harp, J. M.; Egli, M.; Guengerich, F. P. Roles of Residues Arg-61 and Gln-38 of Human DNA Polymerase η in Bypass of Deoxyguanosine and 7,8-Dihydro-8-oxo-2'-deoxyguanosine. *J. Biol. Chem.* **2015**, *290*, 15921–15933.
- (32) Otwinowski, Z.; Minor, W. [20] Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **1997**, *276*, 307–326.
- (33) McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J. Phaser crystallographic software. *J. Appl. Crystallogr.* **2007**, *40*, 658–674.
- (34) Murshudov, G. N.; Vagin, A. A.; Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1997**, *53*, 240–255.
- (35) Adams, P. D.; Afonine, P. V.; Bunkoczi, 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.; Richardson, J. S.; Terwilliger, T. C.; Zwart, P. H. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 213–221.
- (36) Emsley, P.; Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004**, *60*, 2126–2132.
- (37) Ghodke, P. P.; Mali, J. R.; Patra, A.; Rizzo, C. J.; Guengerich, F. P.; Egli, M. Enzymatic bypass and the structural basis of miscoding opposite the DNA adduct 1,N2-ethenodeoxyguanosine by human DNA translesion polymerase η . *J. Biol. Chem.* **2021**, *296*, 100642.
- (38) Patra, A.; Zhang, Q.; Guengerich, F. P.; Egli, M. Mechanisms of Insertion of dCTP and dTTP Opposite the DNA Lesion O6-Methyl-2'-deoxyguanosine by Human DNA Polymerase η . *J. Biol. Chem.* **2016**, *291*, 24304–24313.
- (39) Vaisman, A.; Woodgate, R. Translesion DNA polymerases in eukaryotes: what makes them tick? *Crit. Rev. Biochem. Mol. Biol.* **2017**, *52*, 274–303.
- (40) Reha-Krantz, L. J. DNA polymerase proofreading: Multiple roles maintain genome stability. *Biochim. Biophys. Acta* **2010**, *1804*, 1049–1063.
- (41) Livneh, Z.; Z, O.; Shachar, S. Multiple two-polymerase mechanisms in mammalian translesion DNA synthesis. *Cell Cycle* **2010**, *9*, 729–735.
- (42) Wang, F.; Yang, W. Structural insight into translesion synthesis by DNA Pol II. *Cell* **2009**, *139*, 1279–1289.
- (43) Lee, Y. S.; Gregory, M. T.; Yang, W. Human Pol zeta purified with accessory subunits is active in translesion DNA synthesis and complements Pol η in cisplatin bypass. *Proc. Natl. Acad. Sci. U.S.A.* **2014**, *111*, 2954–2959.
- (44) Kino, K.; Morikawa, M.; Kobayashi, T.; Kobayashi, T.; Komori, R.; Sei, Y.; Miyazawa, H. The oxidation of 8-oxo-7,8-dihydroguanine by iodine. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 3818–3820.
- (45) Ummat, A.; Silverstein, T. D.; Jain, R.; Buku, A.; Johnson, R. E.; Prakash, L.; Prakash, S.; Aggarwal, A. K. Human DNA polymerase η is pre-aligned for dNTP binding and catalysis. *J. Mol. Biol.* **2012**, *415*, 627–634.
- (46) Maccabee, M.; Evans, J. S.; Glackin, M. P.; Hatahet, Z.; Wallace, S. S. Pyrimidine Ring Fragmentation Products. *J. Mol. Biol.* **1994**, *236*, 514–530.