BIOCHEMISTRY
including biophysical chemistry & molecular biology
Article

Biochemistry **2009,** *48,* 7079–7088 7079
DOI: 10.1021/bi9003588

Structural and Functional Analysis of *Sulfolobus solfataricus* Y-Family DNA Polymerase Dpo4-Catalyzed Bypass of the Malondialdehyde−Deoxyguanosine Adduct[†,‡]

Robert L. Eoff,[§,⊥] Jennifer B. Stafford,[§,⊥] Jozsef Szekely,[§] Carmelo J. Rizzo,[§] Martin Egli,[∥] F. Peter Guengerich,[∥] and Lawrence J. Marnett*[,§,∥]

[§]*Department of Chemistry and* [∥]*Department of Biochemistry, A. B. Hancock Jr. Memorial Laboratory for Cancer Research, Vanderbilt Institute of Chemical Biology, Center in Molecular Toxicology, and Vanderbilt-Ingram Cancer Center, Vanderbilt University School of Medicine, Nashville, Tennessee 37232-0146* [⊥]*These authors contributed equally to this work.*

Received March 2, 2009; Revised Manuscript Received May 9, 2009

ABSTRACT: Oxidative stress can induce the formation of reactive electrophiles, such as DNA peroxidation products, e.g., base propenals, and lipid peroxidation products, e.g., malondialdehyde. Base propenals and malondialdehyde react with DNA to form adducts, including 3-(2′-deoxy-β-D-*erythro*-pentofuranosyl) pyrimido[1,2-α]purin-10(3*H*)-one (M_1dG). When paired opposite cytosine in duplex DNA at physiological pH, M_1dG undergoes ring opening to form N^2-(3-oxo-1-propenyl)-dG (N^2-OPdG). Previous work has shown that M_1dG is mutagenic in bacteria and mammalian cells and that its mutagenicity in *Escherichia coli* is dependent on induction of the SOS response, indicating a role for translesion DNA polymerases in the bypass of M_1dG. To probe the mechanism by which translesion polymerases bypass M_1dG, kinetic and structural studies were conducted with a model Y-family DNA polymerase, Dpo4 from *Sulfolobus solfataricus*. The level of steady-state incorporation of dNTPs opposite M_1dG was reduced 260−2900-fold and exhibited a preference for dATP incorporation. Liquid chromatography−tandem mass spectrometry analysis of the full-length extension products revealed a spectrum of products arising principally by incorporation of dC or dA opposite M_1dG followed by partial or full-length extension. A greater proportion of −1 deletions were observed when dT was positioned 5′ of M_1dG. Two crystal structures were determined, including a "type II" frameshift deletion complex and another complex with Dpo4 bound to a dC·M_1dG pair located in the postinsertion context. Importantly, M_1dG was in the ring-closed state in both structures, and in the structure with dC opposite M_1dG, the dC residue moved out of the Dpo4 active site, into the minor groove. The results are consistent with the reported mutagenicity of M_1dG and illustrate how the lesion may affect replication events.

Covalent modification of DNA leads to a complex cellular response that includes altered DNA replication, DNA damage signaling, DNA repair, cell cycle arrest, and cell death (*1*). The pattern of the response depends upon the chemical nature and level of the damage. An important lesion that occurs in human genomic DNA is 3-(2′-deoxy-β-D-*erythro*-pentofuranosyl)pyrimido[1,2-α]purin-10(3*H*)-one (M_1dG) (Figure 1), which is produced by reaction of dG residues with the DNA peroxidation products, base propenals, or the lipid peroxidation product malondialdehyde (MDA)[1] (*2*−*4*). M_1dG adducts have been reported to be present at a density of 5400 molecules per cell in healthy human liver (*5*), and the levels have been reported to increase following exposure to carbon tetrachloride or polychlorinated biphenyls (*6*, *7*). Repair of M_1dG occurs by the nucleotide excision repair pathway (*8*), and M_1dG residues, possibly excised from the genome, have been detected in human urine (*9*).

Base propenals and MDA are mutagenic in *Salmonella typhimurium* (*3*, *10*), and reaction of single- or double-stranded

[†]This work was supported by the National Institutes of Health (R37 CA087819 to L.J.M., R01 ES010375 to F.P.G., P30 ES000267 to C.J.R., F.P.G., L.J.M., and M.E., P01 ES05355 to C.J.R. and M.E., and F32 CA119776 to R.L.E.).

[‡]The atomic coordinates and measured structure factor amplitudes for the 14c:M_1dG, and dGTP:M_1dG complexes (codes 2v4q (r2v4qsf), 2v4r (r2v4rsf)) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rscb.org/).

*To whom correspondence should be addressed: Department of Biochemistry, Vanderbilt University School of Medicine, Nashville, TN 37232-0146. E-mail: larry.marnett@vanderbilt.edu. Telephone: (615) 343-7329. Fax: (615) 343-7534.

[1]Abbreviations: BSA, bovine serum albumin; CID, collision-induced dissociation; ds, double-stranded; DTT, dithiothreitol; ESI, electrospray ionization; LC, liquid chromatography (HP, high-performance; UP, ultraperformance); LC−MS/MS, liquid chromatography−tandem mass spectrometry; MALDI-TOF, matrix-assisted laser desorption ionization time-of-flight; MDA, malondialdehyde; MOPS, 3-morpholinopropane-1-sulfonic acid; MS, mass spectrometry; PdG, 1,N^2-propanodeoxyguanosine; pol, (DNA) polymerase; pol T7$^-$, bacteriophage pol T7 (exonuclease deficient); ss, single-stranded; UDG, uracil DNA glycosylase; wt, wild type.

Published on Web 06/04/2009

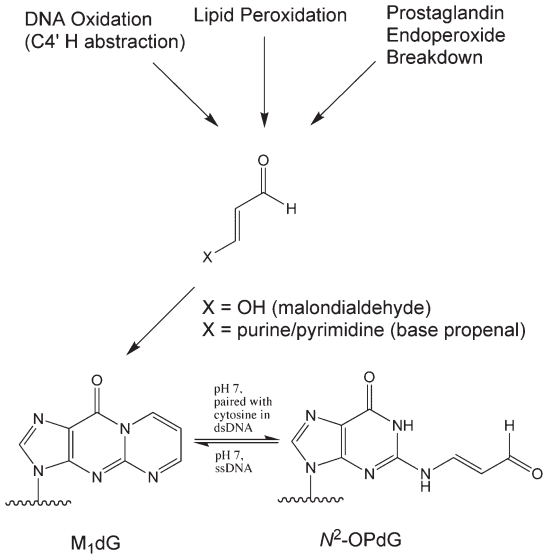

FIGURE 1: Schematic illustration of pathways generating M_1dG/N^2-OPdG and the equilibrium between the ring-open and ring-closed forms of the lesion.

DNA with malondialdehyde followed by transformation into *Escherichia coli* or human cells, respectively, leads to mutations at dG residues and frameshift mutations (mainly deletions) (*11*, *12*). Replication of viral genomes or shuttle vectors containing a single, site-specifically positioned M_1dG leads to the induction of base substitutions (primarily M_1dG to dT and M_1dG to dA), deletions, and untargeted mutations in the vicinity of the adduct (*13*). The type of mutation is sequence-dependent with a larger proportion of mutations occurring in reiterated sequences.

Mutations induced by M_1dG in *E. coli* are dependent on the induction of the SOS response and are significantly reduced by deletion of the *umuC* or *umuD* locus (*8*), indicating that error-prone translesion polymerases are important in M_1dG mutagenesis. The Y-family DNA polymerases represent a specialized class of enzymes that are thought to facilitate bypass of DNA adducts when the replicative polymerases are inhibited (*14*, *15*). To date, no crystal structures of *E. coli* pols IV (*dinB*) and V (*umuD′C*) are available. The Y-family polymerase Dpo4 from *Sulfolobus solfataricus*, which is a *dinB* homologue, has been very useful for determining the structural and kinetic properties of lesion bypass. The structural characteristics of the Y-family polymerase core are conserved in archaeal and eukaryotic systems (*16*), with minor differences probably accounting for variable biochemical properties between enzymes (*17*−*28*). To improve our understanding of the molecular basis for mutagenic profiles observed following exposure to the genotoxic agent MDA, we have studied the in vitro catalytic properties of the model Y-family DNA polymerase Dpo4 from *S. solfataricus* opposite M_1dG. Of particular interest was the possibility that Dpo4 enhances the ring opening of M_1dG during replication, because ring opening of M_1dG to N^2-OPdG occurs spontaneously in duplex DNA (*29*) and has been proposed to contribute to error-free bypass of M_1dG by certain polymerases (*13*). These results provide insights into the bypass of exocyclic DNA lesions and address the question of M_1dG ring opening in the Dpo4 active site.

MATERIALS AND METHODS

Materials. Wild-type Dpo4 was expressed in *E. coli* and purified to electrophoretic homogeneity as described previously (*27*). All unlabeled dNTPs were obtained from Amersham Biosciences (Piscataway, NJ), and [γ-^{32}P]ATP was purchased from PerkinElmer Life Sciences (Boston, MA). Oligonucleotides containing M_1dG were synthesized as described previously (*30*). Unmodified oligonucleotides used in this work were synthesized by Midland Certified Reagent Co. (Midland, TX) and purified using HPLC by the manufacturer, with analysis by matrix-assisted laser desorption time-of-flight MS. The modified 23-mer oligonucleotide for the replication studies was synthesized by a post-synthetic modification approach as previously described (*31*). The modified 18-mer oligonucleotides were prepared from the modified phosphoramidite reagent (*30*, *32*). The modified oligonucleotides were purified by HPLC. A total of six primer−template DNA substrate sequences were used in the kinetic analysis (Table 1). The template oligonucleotide sequence used in the mass spectral analysis was either 5′-TCACXGAATCCTTCCCCC-3′ or 5′-TCATXGAATCCTTCCCCC-3′, where X indicates M_1dG. The 13-base primer sequence used in the mass spectral analyses was 5′-GGGGGAAGGAUTC-3′. The mass of each of the M_1dG-modified oligonucleotides was determined experimentally by MALDI-TOF MS: 5′-TCACXGAATCCTTACGAGCCCCC-3′ [MALDI-TOF MS (HPA) m/z calcd for (M − H) 6965.6, found 6965.2], 5′-TCATXGAATCCTTACGAGCCCCC-3′ [MALDI-TOF MS (HPA) m/z calcd for (M − H) 6977.1, found 6976.9], 5′-TCACXGAATCCTTCCCCC-3′ [MALDI-TOF MS (HPA) m/z calcd for (M − H) 5392.0, found 5393.1], and 5′-TCATXGAATCCTTCCCCC-3′ [MALDI-TOF MS (HPA) m/z calcd for (M−H) 5407.0, found 5407.0], where X denotes M_1dG.

Generation of Primer−*Template DNA for in Vitro Experiments*. The 18-base primer oligonucleotide 5′-GGGGGCTCGTAAGGATTC-3′ was 5′-phosphorylated with T4 polynucleotide kinase with 250 μCi of [γ-^{32}P]ATP (> 6000 Ci/mmol) in 50 mM MOPS buffer (pH 7.5) containing 10 mM $MgCl_2$ and 5 mM DTT for 1 h at 37 °C. The ^{32}P-labeled primers were mixed with M_1dG-containing or control template in a 1:1 molar ratio in the presence of 40 mM NaCl and heated to 95 °C for 2 min, followed by slow cooling overnight.

Steady-State Kinetics. Dpo4 stocks were diluted in 50 mM Tris-HCl buffer (pH 7.5) containing 1 mM DTT, 0.5 mM EDTA, and 10% glycerol (v/v). Reactions were conducted in buffer containing 50 mM Tris-HCl buffer (pH 7.5) containing 50 mM NaCl, 5 mM DTT, 5 mM $MgCl_2$, 100 μg/mL BSA, and 5% glycerol (v/v). Typical reaction volumes (of 10 μL) contained 50 nM annealed, 5′-end labeled primer−template, 5 nM Dpo4, and varying concentrations of individual nucleotide triphosphates (from 1 nM to 1 mM) in reaction buffer. Enzyme, buffer, and DNA template were preincubated at 37 °C for 5 min prior to the addition of $MgCl_2$ and dNTP to initiate the reaction. The reaction incubation periods were optimized to ensure that the data collected represented the initial velocities of the reactions. Reactions were terminated with the addition of 10 μL of 95% formamide, 10 mM EDTA, 0.03% bromophenol blue, and 0.03% xylene cyanol, and reaction mixtures were then heated to 95 °C for 2 min to ensure reaction quenching. Products were visualized on 20% polyacrylamide (w/v)/7 M urea gels by electrophoresis at a constant voltage (3000 V) for 3 h. The products were then visualized using a phosphorimager and quantitated using Quantity One.

Transient-State Kinetics. All pre-steady-state experiments were performed using a KinTek RQF-3 model chemical quench-flow apparatus (KinTek Corp., Austin, TX) with 50 mM Tris-HCl (pH 7.4) buffer in the drive syringes. All RQF experiments

Table 1: Sequences[a] of the DNA Substrates

Sequences of the DNA substrates

Single nucleotide incorporation assays:

5′-GGG GGC TCG TAA GGA TTC-3′ 3′-CCC CCG AGC ATT CCT AAG **XC**A CT-5′	substrate 1
5′-GGG GGC TCG TAA GGA TTC-3′ 3′-CCC CCG AGC ATT CCT AAG **XT**A CT-5′	substrate 2
5′-GGG GGC TCG TAA GGA TTC **C**-3′ 3′-CCC CCG AGC ATT CCT AAG **XC**A CT-5′	substrate 3
5′-GGG GGC TCG TAA GGA TTC **A**-3′ 3′-CCC CCG AGC ATT CCT AAG **XC**A CT-5′	substrate 4
5′-GGG GGC TCG TAA GGA TTC **C**-3′ 3′-CCC CCG AGC ATT CCT AAG **XT**A CT-5′	substrate 5
5′-GGG GGC TCG TAA GGA TTC **A**-3′ 3′-CCC CCG AGC ATT CCT AAG **XT**A CT-5′	substrate 6

Substrates used in the LC-MS analysis:

5′-GGG GGA AGG AUT C-3′ 3′-CCC CCT TCC TAA G**XC** ACT-5′	substrate 7
5′-GGG GGA AGG AUT C-3′ 3′-CCC CCT TCC TAA G**XT** ACT-5′	substrate 8

where X denotes G or M_1dG

[a] Where X denotes G or M_1dG.

were conducted at 37 °C in 50 mM Tris-HCl (pH 7.4) buffer containing 50 mM NaCl, 5 mM DTT, 100 μg/mL BSA, and 5% (v/v) glycerol. Dpo4 (350 nM) was incubated with DNA substrate 1 (200 nM) containing either G or M_1dG in the template at the insertion position, and the reaction was initiated upon mixing in the quench apparatus with a solution containing dNTP (1 mM) and $MgCl_2$ (5 mM). Polymerase catalysis was stopped by the addition of 500 mM EDTA (pH 9.0). Substrate and product DNA were separated by electrophoresis on a 20% polyacrylamide (w/v)/7 M urea gel. Results obtained under single-turnover conditions were fit to eq 1

$$y = A(1 - e^{-k_{obs}t}) \quad (1)$$

where A is the product formed in the first binding event, k_{obs} is the rate constant defining polymerization under the conditions used for the experiment being analyzed, and t is time.

LC−MS/MS Analysis of Oligonucleotide Products from Dpo4 Reactions. Dpo4 (5 μM) was preincubated with primer−template DNA (10 μM), and the reaction was initiated by addition of dNTP (1 mM each) and $MgCl_2$ (10 mM) in a final volume of 200 μL. Dpo4 catalysis was allowed to proceed at 37 °C for 5 h in 50 mM MOPS (pH 7.4) buffer containing 50 mM NaCl, 1 mM DTT, 100 μg/mL BSA, and 5% glycerol (v/v). The reaction was terminated by extraction of the remaining dNTPs using a size-exclusion chromatography column (Bio-Spin 6 chromatography column, Bio-Rad, Hercules, CA). Concentrated stocks of MOPS (pH 7.4), DTT, and EDTA were added to restore the concentrations to 50, 5, and 1 mM, respectively. Next, *E. coli* uracil DNA glycosylase (20 units) (New England Biolabs, Ipswich, MA) was added, and the solution was incubated at 37 °C for 6 h to hydrolyze the uracil residue on the extended primer. The reaction mixture was then heated at 95 °C for 1 h in the presence of 0.25 M piperidine, followed by removal of the solvent by centrifugation under vacuum. The dried sample was resuspended in 100 μL of H_2O for MS analysis.

LC−MS/MS analysis (*21*, *27*) was performed on a Waters Aquity UPLC system (Waters, Milford, MA) connected to a Finnigan LTQ mass spectrometer (Thermo Fisher Scientific, Waltham, MA), operating in the ESI negative ion mode, as described previously (*20*). The nomenclature used in Tables S1−S29 of the Supporting Information has been described previously (*33*). An estimate for the amount of each product in the reaction mixture was calculated by integrating the area of the corresponding peak in a selected ion trace for the ion of interest.

Crystallization of Dpo4−DNA Complexes. Dpo4 was crystallized in complex with two primer−template DNA sequences: 13-mer primer−18-mer template and 14-mer primer−18-mer template. In both instances, the 18-mer template sequence was 5′-TCA CXG AAT CCT TCC CCC-3′, where X is M_1dG. The 13-mer primer strand was 5′-GGG GGA AGG ATT C-3′, and the 14-mer primer strand was 5′-GGG GGA AGG ATT CC-3′. Dpo4 was mixed with DNA (1:1.2 molar ratio) in 20 mM Tris-HCl buffer (pH 7.4, 25 °C) containing 60 mM NaCl, 4% glycerol (v/v), and 5 mM β-mercaptoethanol and then placed on ice for 1 h prior to incubation with $CaCl_2$ (5 mM) and dGTP (1 mM). The final Dpo4 concentration was ∼10 mg/mL. The complexes were crystallized by sitting-drop vapor diffusion as described previously (*27*) with PEG 3350 (7.5−10%), $Ca(OAc)_2$ (100 mM), and glycerol (2.5%) in the reservoir solution.

X-ray Diffraction Data Collection and Processing. The X-ray diffraction data sets for the two complexes were collected at the Advanced Photon Source (APS, 21-ID, LS-CAT, Argonne, IL) at 110 K and using a synchrotron radiation wavelength of 1.0 Å. Indexing and scaling were performed using XDS (*34*). Both structures belong to space group $P2_12_12$. The resulting data sets for the 14C−M_1dG and M_1dG−dGTP complexes were of good quality, with R_{merge} values of 5.3 and 8.7%, respectively. CCP4 package programs, including TRUNCATE (*35*), were used for further processing of the data.

Structure Determination and Refinement. A refined wild-type Dpo4−dG complex [Protein Data Bank entry 2bqr (*27*)] without solvent molecules, ions, and dGTP was used as the starting model for both data sets. The initial position of the model was optimized by several rounds of rigid body refinement while the resolution of the diffraction data was gradually increased. Manual model rebuilding was conducted with TURBO-FRODO.[2] The maps were computed using the σA modified coefficients (*36*). Clear positive density for the Ca^{2+} ions and the dGTP was observed in the initial difference Fourier electron density maps of both complexes. In the case of the cytosine at postion 14 in the primer of the 14C−M_1dG structure, there was clear positive density corresponding to the phosphate and the pyrimidine base following molecular replacement, but a discontinuity in the region corresponding to the ribose moiety required several rounds of refinement before the final position of the base was determined. The CNS package (*37*) was used for the refinement of the models by performing simulated annealing, gradient minimization, refinement of individual isotropic temperature factors, and individual occupancy. The crystallographic figures were prepared using Pymol.[3]

[2] Cambillau, C., and Roussel, A. (1997) *Turbo Frodo*, OpenGL 1 version, Université Aix-Marseille II, Marseille, France.

[3] DeLano, W. L. (2002) *The pyMOL Molecular Graphics System*, DeLano Scientific, San Carlos, CA.

Comparison of Dpo4-Catalyzed Bypass of M_1dG and N^2-OPdG. M_1dG-containing template DNA (substrate 1) was annealed to a complementary primer using two conditions. ^{32}P-radiolabeled primer was added to template DNA (1:1) in the presence of MOPS (50 mM at either pH 7.5 or 10) and NaCl (40 mM), heated to 95 °C, and slowly cooled to room temperature. Annealing at pH 10 results in formation of the ring-opened N^2-OPdG. M_1dG remains in the ring-closed form in ssDNA at pH 7.5 (*38*). Dpo4 catalysis of incorporation of either dCTP or dATP (1 mM) was then assessed under four conditions. Under the first condition, primer−template DNA annealed at pH 7.5 was used as a control to measure dNTP incorporation opposite M_1dG. Under the second condition, primer−template DNA annealed at pH 10 was added to assay buffer containing MOPS (50 mM, pH 7.5), NaCl (50 mM), 1 mM DTT, 100 μg/mL BSA, and 5% glycerol (v/v), and the reaction was started immediately by addition of dNTP (1 mM) and $MgCl_2$ (5 mM). Under these conditions, Dpo4 catalysis should encounter the ring-opened N^2-OPdG adduct. The third condition involved adding assay buffer (pH 7.5) to primer−template DNA that was annealed at pH 10 and incubating the sample for approximately five half-lives of the N^2-OPdG ring closure (2 h) before initiating polymerase catalysis. The fourth and final experimental setup involved incubating primer−template DNA (pH 10), Dpo4, and assay buffer (pH 7.5) for 2 h prior to initiation of polymerase activity. The 2 and 5 h incubations should both have allowed quantitative ring closure to occur for DNA annealed at pH 10. Only the second condition tested the ability of Dpo4 to bypass the ring-opened form of the adduct.

RESULTS

Steady-State Kinetic Analysis of Dpo4 Catalysis opposite M_1dG. Single-nucleotide incorporation assays were performed to measure steady-state parameters of Dpo4 catalysis. Two DNA substrates that differed only in the base to the 5′ side of the adduct were used (Table 1, substrates 1 and 2). Incorporation of dATP was favored for both substrates (Table 2). A higher k_{cat} was the primary reason dATP was favored when C was positioned 5′ of M_1dG (substrate 1). Conversely, a lower $K_{m,dNTP}$ led to more efficient incorporation of dATP when T was 5′ of M_1dG (substrate 2). The catalytic efficiency ($k_{cat}/K_{m,dNTP}$) was decreased 260- and 430-fold for incorporation of dATP for substrates 1 and 2, respectively. For substrate 1, dCTP was the second most efficient insertion event, and dCTP was the least favored event with substrate 2.

Transient-State Kinetic Analysis of Dpo4-Catalyzed dNTP Incorporation opposite M_1dG. Single-turnover experiments were performed, but incorporation of both dCTP and dATP was strongly inhibited in the presteady state (Figure 2A). Incorporation of dCTP yielded a small fraction of product (~7.5%) in the presteady state, but after 10 s, incorporation of dATP was clearly more abundant (Figure 2B). It is possible that a small fraction of dCTP binding events induce the ring-opened state of M_1dG in the polymerase active site, resulting in the small burst in product observed in the presteady state, but dATP is clearly favored over accurate bypass in every other kinetic measure of Dpo4 catalysis.

Analysis of Primer Incorporation/Extension Products Using LC−MS/MS. The ability of M_1dG to generate frameshift mutations in cells could play a significant role in disrupting coding sequences. To assess the propensity for frameshift mutations by Dpo4, we analyzed incorporation/extension products in vitro using LC−MS/MS. In contrast to single-nucleotide incorporation assays, analysis of full-length extension products in the presence of all four dNTPs can provide a measure of truncated, untargeted (i.e., random), mutations and/or frameshifted extension products.

Table 2: Steady-State Kinetic Parameters for One-Base Incorporation by Dpo4

template base	dNTP	k_{cat} (s^{-1})	$K_{m,dNTP}$ ($\times 10^3$ μM)	k_{cat}/K_M ($\times 10^3$ μM^{-1} s^{-1})	fold reduction from dCTP:dG
		Substrate 1 with dC 5′ of G/M_1dG			
dG	dCTP	160 ± 5	3.2 ± 0.6	49	–
	dATP	20 ± 2	260 ± 90	0.075	650
	dTTP	35 ± 1	200 ± 40	0.18	270
	dGTP	16 ± 2	70 ± 40	0.22	220
M_1dG	dCTP	9.3 ± 1.7	120 ± 90	0.080	610
	dATP	24 ± 1	130 ± 20	0.19	260
	dTTP	3.8 ± 0.3	110 ± 30	0.034	1400
	dGTP	1.1 ± 0.1	30 ± 5	0.035	1400
		Substrate 2 with dT 5′ of G/M_1Dg			
dG	dCTP	890 ± 30	4.3 ± 0.5	210	–
	dATP	22 ± 1	120 ± 20	0.18	1.200
	dTTP	58 ± 2	230 ± 40	0.25	840
	dGTP	19 ± 1	400 ± 120	0.048	4400
M_1dG	dCTP	29 ± 2	400 ± 80	0.073	2900
	dATP	15 ± 1	30 ± 5	0.49	430
	dTTP	15 ± 1	130 ± 30	0.12	1800
	dGTP	16 ± 2	180 ± 70	0.089	2400

In vitro extension was performed on two M_1dG-modified templates that differed in the identity of the base on the 5′ side of the M_1dG lesion (dC vs dT) (Figure 3). Bypass of M_1dG by Dpo4 was highly error-prone in both sequence contexts. In the 5′-C-(M_1dG)-G-3′ sequence context (substrate 7), 48% of the extension products were error-free (Figure 3A). In the 5′-T-(M_1dG)-G-3′ sequence context (substrate 8), 17% of the extension products were error-free (Figure 3B). For substrate 7, ~17 or 2% of the extension products was derived from a simple misinsertion of dATP or dTTP opposite M_1dG, respectively. Two frameshift deletion products were observed with substrate 7. The first product [5′-pTC-GTGA-3′ (7%)] results from a deletion at the adduct site, and a second product results from misinsertion of A followed by skipping of the next base [5′-pTC**A**-TGA-3′ (3%)]. The former can be rationalized by insertion of dGTP opposite the dC that is 5′ of M_1dG. A similar product was observed for the 1,N^2-ε-dG adduct in the same sequence (*27*). Crystallographic analysis of a ternary Dpo4 complex bound to 1,N^2-ε-dG-modified DNA revealed that the extended purine ring system of the two-carbon etheno bridge provided a π-stacking platform for the incoming dGTP, which formed a nascent Watson-Crick pair with the 5′-dC of the template. Three related products were identified in 6, 3, and 4% relative yield that result from misinsertion of dATP opposite M_1dG followed by error-prone extension.

A variety of error-prone bypass products were identified from the insertion and extension of substrate 8 (Figure 3B). As stated above, the extent of error-free bypass was much reduced (17%) with substrate 8. Misinsertion of dATP and dGTP followed by accurate extension accounted for only 9 and 1% of the product distribution, respectively. The most prevalent products were those derived from a one-base deletion opposite the lesion

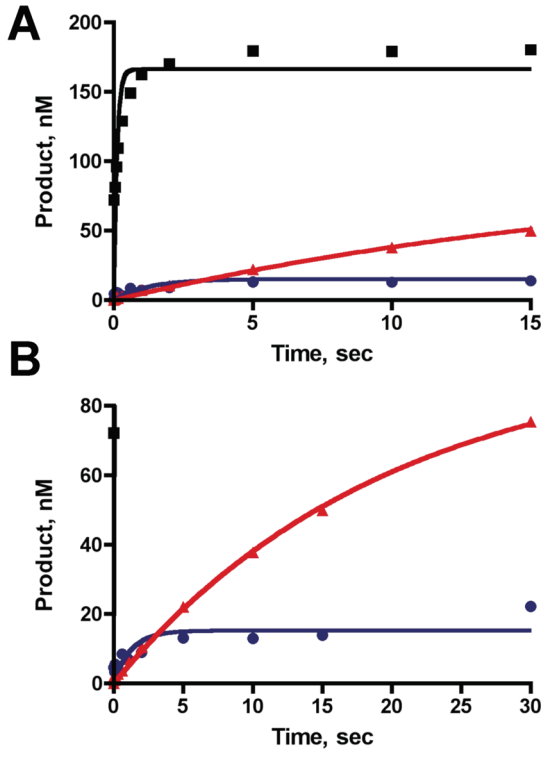

FIGURE 2: Pre-steady-state analysis of Dpo4-catalyzed incorporation opposite M_1dG. (A) Dpo4 (350 nM) was incubated with radiolabeled 18/23-mer DNA (200 nM), the indicated dNTP (1 mM), and $MgCl_2$ (5 mM). The data were fit to eq 1 to yield the following kinetic parameters: for dCTP·G (■), $A = 170 \pm 9$ nM and $k_{obs} = 8.9 \pm 2.3\ s^{-1}$; for dCTP·$M_1$dG (blue ●), $A = 15 \pm 2$ nM and $k_{obs} = 0.81 \pm 0.35\ s^{-1}$; for dATP·$M_1$dG (red ▲), $A = 95 \pm 4$ nM and $k_{obs} = 0.051 \pm 0.004\ s^{-1}$. (B) The results from panel A are shown to illustrate the small amount of product formed rapidly during dCTP insertion opposite M_1dG.

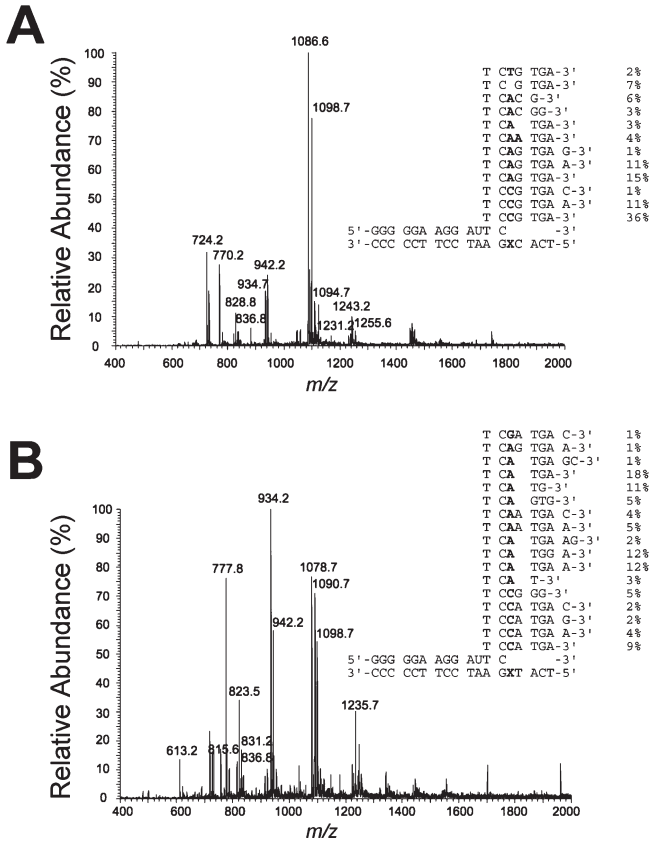

FIGURE 3: LC−MS analysis of Dpo4-catalyzed full-length extension products. (A) ESI mass spectrum of products derived from Dpo4-catalyzed extension of 13/18-mer DNA containing M_1dG with cytosine to the 5′ side of the lesion. The products identified in the spectrum are summarized. (B) ESI mass spectrum of products derived from Dpo4-catalyzed extension of 13/18-mer DNA containing M_1dG with thymidine to the 5′ side of the lesion. The products identified in the spectrum are summarized.

(59%); these seven products differed in the progress of the extension steps. Two related deletion products resulted from error-prone extension (6%). An additional product was characterized as resulting from the correct insertion of dCTP followed by highly error-prone extension (5%).

In both sequence contexts, the steady-state catalytic efficiency was best for the insertion of dATP (Table 2). The degree of insertion of dATP opposite M_1dG was 260- and 430-fold lower than the degree of insertion of dCTP opposite G in the 5′-C-(X)-G-3′ (substrate 1) and 5′-T-(X)-G-3′ (substrate 2) sequences, respectively. Extension products arising from the initial insertion of dATP into the 5′-C-(M_1dG)-G-3′ sequence (substrate 7) accounted for 43% of the observed LC−MS products, while 48% of the observed LC−MS products resulted from initial insertion of dCTP in the same sequence. Kinetic analysis shows that the catalytic efficiency for dCTP insertion was ~2.4-fold lower than for insertion of dATP in the 5′-C-(M_1dG)-G-3′ sequence context (Table 2). On the basis of the kinetic results, we would expect to see a greater proportion of products resulting from dATP insertion in the LC−MS analysis of substrate 7 than what is actually observed. The larger-than-expected fraction of dCTP insertion products in the LC−MS analysis suggests that extension from the M_1dG·dC template−primer terminus is favored over extension from an M_1dG·dA terminus, at least in the 5′-C-(M_1dG)-G-3′ sequence.

Dpo4 inserts dATP opposite the M_1dG lesion in the 5′-T-(M_1dG)-G-3′ sequence 6.7-fold more efficiently than dCTP (Table 2). The misinsertion of dATP opposite M_1dG represents 73% of the total extension products observed by LC−MS in the 5′-T-(M_1dG)-G-3′ sequence context (Figure 3B). The small amount of error-free products observed in the 5′-T-(M_1dG)-G-3′ sequence (17%) results from the fact that (i) Dpo4 favors insertion of dATP and (ii) Dpo4 can readily accommodate a pairing between the inserted dA and the thymidine on the 5′ side of M_1dG by skipping over the lesion. The kinetic efficiency for insertion of dTTP or dGTP opposite M_1dG was slightly better than that for insertion of dCTP using the 5′-T-(M_1dG)-G-3′ template; however, in the 5′-T-(M_1dG)-G-3′ sequence context, only 1% of the observed products resulted from initial misinsertion dGTP, and no products were characterized that were derived from initial insertion of dTTP, suggesting that extension of a template−primer terminus containing dG or dT opposite M_1dG is highly disfavored.

Steady-State Kinetic Analysis of Next-Nucleotide Extension past M_1dG. The identities of the full-length extension products did not correspond to our expectations based on the results of the single-nucleotide incorporation experiments. Accurate insertion of dC into the full-length products in both sequence contexts was more efficient than anticipated from single-nucleotide incorporation experiments. We determined the steady-state kinetic parameters for single-nucleotide extension from M_1dG·dC and M_1dG·dA mispairs (Table 3). The results indicated that M_1dG·dC mispairs (substrate 3) are extended 7.5-fold more rapidly than M_1dG·dA mispairs (substrate 4) in the

Table 3: Steady-State Kinetic Parameters for Next-Base Extension past M_1dG by Dpo4

primer pair	dNTP	k_{cat} (s^{-1})	$K_{m,dNTP}$ ($\times 10^3$ μM)	k_{cat}/K_M ($\times 10^3$ μM^{-1} s^{-1})	fold reduction from dC:dG extension
Substrates Containing dC 5′ of G/M_1dG					
dC·dG	dGTP	420 ± 10	5.0 ± 0.7	84	–
dA·dG	dGTP	24 ± 1	52 ± 5	0.47	180
dC·M_1dG	dGTP	6.6 ± 0.3	2.2 ± 0.5	2.9	29
dA·M_1dG	dGTP	8.3 ± 0.5	13 ± 4	0.64	130
Substrates Containing dT 5′ of G/M_1dG					
dC·dG	dATP	1100 ± 30	14 ± 1	78	–
dA·dG	dATP	2.1 ± 0.1	64 ± 18	0.033	2400
dC·M_1dG	dATP	28 ± 1.0	5.6 ± 1.0	5.0	16
dA·M_1dG	dATP	1.4 ± 0.1	56 ± 13	0.025	3100
dA·M_1dG	dTTP	2.6 ± 0.3	13 ± 7	0.2	390

5′-C-(M_1dG)-G-3′ sequence context. Thus, even though dC is not incorporated opposite M_1dG more readily than dA, the M_1dG·dC mispair is extended more efficiently, resulting in a higher-than-expected fraction (48%) of error-free products observed during LC−MS analysis of the extension products.

Extension of a primer containing dA opposite M_1dG was found to proceed most efficiently when dTTP was present in the reaction mixture (Table 3) even though the template base should be dT (substrate 6), which would indicate that dA is not paired with M_1dG but is stably paired with the template thymidine to the 5′ side of M_1dG in what is often termed a "type II" conformation (*39*). With a stable type II orientation, the template base in substrate 6 would be dA, which explains the preference for dTTP during extension. The favorable insertion of dATP opposite M_1dG combined with the increased kinetic propensity to extend the type II complex is consistent with the larger fraction of −1 deletion products observed in the LC−MS/MS analysis.

Single-Nucleotide Incorporation of dNTPs opposite Ring-Opened N^2-OPdG. Primer−template DNA containing the M_1dG modification was prepared using conditions that should generate the ring-opened form of the MDA-derived adduct. At pH 10, the ring-opened form (N^2-OPdG) is favored (*40*). The half-life of ring closing is relatively slow so that upon addition of pH 7.4 assay buffer a comparison can be made between Dpo4-catalyzed bypass of M_1dG and N^2-OPdG (*41*). Experiments analogous to those described above for single-nucleotide incorporation opposite M_1dG were conducted. The results showed no clear difference in the rate of nucleotide incorporation opposite either M_1dG or N^2-OPdG (Figure S1 of the Supporting Information). Dpo4 polymerization was strongly inhibited by both forms of the lesion, as evidenced by the extremely slow rate of dCTP incorporation. As with M_1dG, dATP incorporation occurred at a faster rate than dCTP incorporation. Similar experiments performed with the Klenow fragment of DNA polymerase revealed a qualitative and quantitative difference in bypass of M_1dG and N^2-OPdG, so it is unlikely that the similarity in bypass observed in these experiments is due to rapid ring closing of N^2-OPdG to M_1dG in solution (*38*). However, we cannot rule out the possibility that Dpo4 promotes ring closing of N^2-OPdG to M_1dG when the modified template is in complex with the enzyme.

Dpo4·M_1dG·DNA Ternary Complex Structures. Previous studies have shown that placing cytosine opposite M_1dG in duplex DNA at pH 6.8 results in ring opening to form N^2-OPdG (*29*). Ring opening is reversible ($t_{1/2}$ = 23 min at neutral pH) and does not occur in ssDNA at physiological pH (*40*, *41*). It is of interest to know if DNA polymerases facilitate ring opening during bypass of M_1dG-containing DNA because the nature of the lesion during replication could be a strong determinant of whether bypass is error-free or error-prone. We determined two Dpo4·M_1dG·DNA·dGTP·Ca^{2+} ternary structures to ascertain the structural properties of M_1dG that are relevant to translesion DNA synthesis (Table 4).

Table 4: Crystal Data and Refinement Parameters for the Ternary (protein·DNA·dGTP) Complexes of Dpo4

	14C·M_1dG	M_1dG·dGTP
X-ray source	APS (LS-CAT)	APS (LS-CAT)
beamline	ID-21	ID-21
detector	MAR CCD	MAR CCD
wavelength (Å)	0.98	0.98
temperature (K)	110	110
no. of crystals	1	1
space group	$P2_12_12$	$P2_12_12$
unit cell [*a*, *b*, *c* (Å)]	92.61, 103.14, 52.00	93.54, 102.54, 52.18
resolution range (Å)	30.0−2.60	30.0−2.45
highest-resolution shell[a]	2.69−2.60	2.50−2.45
no. of measurements	87070 (8060)	65537 (7112)
no. of unique reflections	15716 (1443)	17600 (1776)
redundancy	5.5 (5.5)	3.7 (4.0)
completeness (%)	99.5 (96.8)	97.8 (99.8)
R_{merge}[b]	5.3 (45.5)	8.7 (34.4)
signal to noise ratio ($\langle I/\sigma I \rangle$)	23.10 (4.38)	10.89 (3.93)
solvent content (%)	59.2	59.5
model composition (asymmetric unit)		
no. of amino acid residues	341	342
no. of water molecules	100	137
no. of Ca^{2+} ions	3	3
no. of template nucleotides	16	17
no. of primer nucleotides	14	13
no. of dGTP molecules	1	1
R_f[c] (%)	20.5	21.7
R_{free}[d] (%)	25.2	27.2
estimated coordinate error (Å)		
from Luzatti plot	0.33	0.33
from Luzatti plot (c-v[e])	0.43	0.45
from sA plot	0.38	0.39
from sA plot (c-v[e])	0.43	0.53
temperature factor		
from Wilson plot (Å^2)	61.8	47.1
mean isotropic (Å^2)	59.2	45.0
root-mean-square deviation in temperature factor		
two bonded main chain atoms (Å)	1.25	1.37
two bonded side chain atoms (Å)	1.89	1.99
root-mean-square deviation from ideal values		
bond lengths (Å)	0.008	0.007
bond angles (deg)	1.3	1.3
dihedral angles (deg)	21.6	21.8
improper angles (deg)	1.56	1.94

[a] Values in parentheses correspond to the highest-resolution shells. [b] $R_{merge} = \sum_{hkl} \sum_{j=1,N} |\langle I_{hkl} \rangle - I_{hklj}| / \sum_{hkl} \sum_{j=1,N} |I_{hklj}|$, where the outer sum (*hkl*) is taken over the unique reflections. [c] $R_f = \sum_{hkl} |F_{ohkl}| - k|F_{chkl}| / \sum_{hkl} |F_{ohkl}|$, where $|F_{ohkl}|$ and $|F_{chkl}|$ are the observed and calculated structure factor amplitudes, respectively. [d] R_{free} *idem*, for the set of reflections (5% of the total) omitted from the refinement process. [e] Cross-validation.

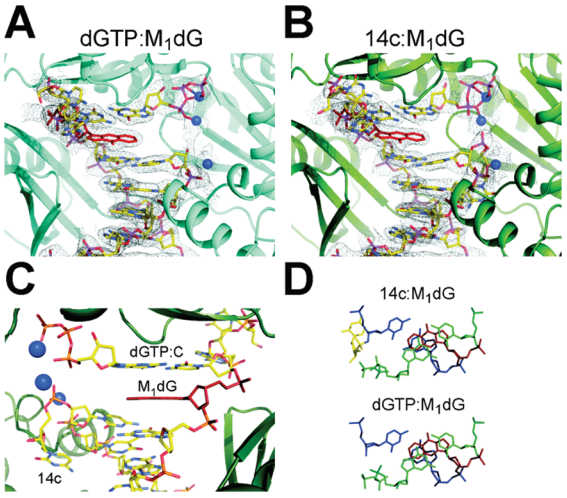

FIGURE 4: Configuration of bases in the active site of Dpo4·M_1dG ternary complexes. (A) The quality of the electron density from the dGTP·M_1dG crystal structure is shown (gray wire mesh) with the $3F_o-2F_c$ map contoured to 1σ. Dpo4 (cyan) is shown in schematic form in complex with primer−/template DNA (yellow carbons) and calcium ions (blue spheres). (B) The quality of the electron density from the 14C·M_1dG crystal structure is shown (gray wire mesh) with the $3F_o - 2F_c$ map contoured to 1σ. Dpo4 (green) is shown in schematic form in complex with primer−template DNA (yellow carbons) and calcium ions (blue spheres). (C) The orientation of the primer terminus (14C) relative to M_1dG from the 14C·M_1dG stucture is shown as viewed from the palm domain of Dpo4. (D) Stacking interactions between the DNA residues in both structures are shown. The C·G pair at the primer terminus (blue) is shown in the background, with M_1dG and the dGTP·C pair (green) stacked in the foreground along the helical axis. The 14th primer residue from the 14C·M_1dG crystal structure (14C) is colored yellow. The M_1dG moiety is colored red in all panels for the sake of clarity.

The first structure determined (M_1dG·dGTP) contained what represents formation of a −1 frameshift deletion product in the active site of Dpo4 (Figure 4A). The M_1dG moiety is in the ring-closed form and base-stacked between adjacent template bases. Similar to the case for the other so-called type II Dpo4 complexes (*39*), nascent pairing between the incoming dGTP and the dC to the 5′ side of M_1dG is observed, leaving an extended (6 Å) gap between the α-phosphate of dGTP and the 3′-hydroxyl group at the primer terminus. An ordered water molecule bridges the 3′-hydroxyl group and the phosphate. Three Ca^{2+} ions are coordinated in or near the active site, with two serving to coordinate dGTP binding. Other structural features, e.g., overall protein conformation, are quite similar to those observed in previous Dpo4 structures (*21*, *28*, *39*). Attempts to crystallize Dpo4 and M_1dG in the presence of an incoming dCTP or dATP were unsuccessful.

A second structure was determined using a 14-mer primer sequence in which a 3′ terminal cytosine (position 14) is designed to pair with the template M_1dG (14C·M_1dG). In stark contrast to what is observed in duplex DNA, M_1dG remains in the ring-closed form even when cytosine is covalently attached to the primer DNA opposite M_1dG (Figure 4B). The reason M_1dG is in the ring-closed state is apparent from the crystal structure. Instead of pairing with the lesion to stabilize the ring-opened N^2-OPdG, the cytosine at the 3′ terminus of the primer is displaced into the growing minor groove. Such a conformation has been observed previously during Dpo4-catalyzed bypass of O^6-benzylguanine and is considered to be nonproductive (*42*). The 14C·M_1dG structure suggests that the Dpo4 active site may not be restrictive enough to force cytosine to stabilize M_1dG ring opening. Another interpretation would consider that the level of cytosine hydration is most likely reduced in the active site of the polymerase relative to DNA in solution. Desolvation at the N3 atom of a dCTP molecule as it enters the polymerase active site may result in a reduced rate of ring opening. Failure to generate the ring-opened N^2-OPdG leaves an exocyclic DNA adduct that is well-stacked in the template strand, presenting a poor instructional site for Dpo4.

DISCUSSION

Previous work has shown M_1dG to be a mutagenic lesion (*13*), but the mechanism of polymerase bypass opposite M_1dG remains largely undefined. To improve our understanding of the basis for M_1dG-induced mutagenesis, the mechanism of translesion DNA synthesis opposite M_1dG by the model Y-family polymerase Dpo4 was studied at a molecular level using kinetic and structural approaches. Steady-state and transient-state kinetic results both indicate that Dpo4 catalysis is inhibited by M_1dG (260−2900-fold), with dATP being the favored insertion event for both sequences tested (Table 2 and Figure 2). However, kinetic analysis of the next-base insertion event showed that extension of the M_1dG·dC pair, while inhibited ~30-fold relative to that of unmodified substrates, is ~8-fold more efficient than extension of the M_1dG·dA pair when dC is 5′ to M_1dG (Table 3). In the 5′-T-(M_1dG)-G-3′ sequence context, kinetic analysis predicts a shift toward higher levels of −1 deletion products (Table 3). The increase in −1 deletion products results from the ability of Dpo4 to accommodate two template bases in the polymerase active site. The efficient insertion of dA opposite M_1dG then leads to a large portion of substrate molecules being extended as −1 deletions when T is 5′ to the adduct site. The −1 deletion products are not as abundant in the 5′-C-(M_1dG)-G-3′ sequence context because dGTP is not readily inserted opposite M_1dG (Table 2). Analysis of the full-length extension products by LC−MS/MS is consistent with the kinetic results but also reveals unpredictable error-prone bypass events (Figure 3).

Both the kinetic and LC−MS/MS results obtained with Dpo4 are largely consistent with the mutagenic profiles obtained for replication of site-specifically modified shuttle vectors in *E. coli* and COS-7 cells (*13*). Frameshift mutations have been observed during replication of ssDNA containing M_1dG in wild-type (LM102) and NER-deficient (LM103) *E. coli* strains (*13*). Our data clearly show that frameshift deletions are generated during Dpo4-catalyzed bypass of M_1dG (Figure 3). In *E. coli* (LM102), insertion of M_1dG resulted in a targeted mutation frequency of 18% (when corrected for strand bias) (*8*). In NER-deficient *E. coli* strains (LM103, LM115, and NR10148), the mutation frequency of M_1dG is increased 2−3-fold (*8*), which is relatively consistent with the levels of misinsertion obtained with Dpo4 (Table 2 and Figure 3). The primary mutations observed for M_1dG replication in either *E. coli* or COS-7 cells were either G to T transversions or G to A transitions (*8*, *13*). Dpo4 is more efficient at the insertion of dATP opposite M_1dG, and this result differs from the mutagenesis results because very little dTTP incorporation either is observed in the extension products determined by LC−MS (Figure 3) or is predicted by kinetic analysis (Table 2). However, multiple untargeted mutations were identified during Dpo4-catalyzed bypass of M_1dG (Figure 3), similar to what was observed in bacterial and mammalian cells.

We have determined two crystal structures of Dpo4 in a ternary complex with M_1dG-modified DNA and an incoming dGTP. M_1dG exists in the ring-closed form in both of the crystal structures (Figure 4). The first structure determined (dGTP·M_1dG) reveals a typical type II ternary complex in which the incoming dGTP is paired with the base to the 5′ side of the M_1dG lesion. In an attempt to address whether the ring-opened derivative of M_1dG (N^2-OPdG) is observed in the active site of Dpo4, a cytosine was placed in the primer strand opposite M_1dG. The Dpo4 active site is apparently not conducive to stabilization of the ring-opened form of the M_1dG adduct even when cytosine is placed in a postinsertion context, as evidenced by the 14C·M_1dG structure (Figure 4B). The lack of stable pairing between dC and M_1dG in the 14C·M_1dG structure precludes any ring opening.

As has been noted before, Dpo4 has no significant hydrogen bonding contacts with the purine/pyrimidine ring systems but rather relies upon van der Waals interactions and hydrogen bonding with the phosphate backbone and glycosidic moieties to effectively manipulate the DNA template (*39*). Stabilization of the incoming dNTP is primarily derived from hydrogen bonding among the triphosphate moiety, the 3′-hydroxyl group, and positive centers in the finger and palm domain. In the absence of any restrictive forces, the cytosine at the primer−template junction is moved into a pocket between the DNA and the palm domain of Dpo4 and does not readily catalyze ring opening of M_1dG to N^2-OPdG. Such a conclusion is further supported by the observation that dCTP insertion opposite M_1dG is strongly inhibited and by the large fraction of error-prone bypass events identified by LC−MS.

NMR studies with M_1dG-modified oligonucleotides have shown conclusively that ring-opened N^2-OPdG exists when placed opposite cytosine and the pair is located in the center of an 8 bp duplex (*29*). It seems reasonable to assume that ring opening for a dC·M_1dG pair located at a ssDNA−dsDNA junction is less likely because of thermal fluctuations (i.e., fraying) at the end of a duplex. We wanted to determine whether Dpo4 could act to facilitate ring opening during bypass of M_1dG. It has been suggested that ring opening occurs through a mechanism that involves hydration at the N3 atom of cytosine (*29*, *40*). It is possible that a similar mechanism occurs in the presence of Dpo4. The small burst in product observed for dCTP insertion opposite M_1dG in the presteady state (Figure 2) and the relatively large proportion of dC-containing products measured by LC−MS/MS (Figure 3) are consistent with the view that M_1dG ring opening occurs during insertion of dC. However, on the basis of the crystal structure of Dpo4 bound to primer−template junction-containing DNA (14C·M_1dG), the mechanism of ring opening for M_1dG-containing DNA in isolation may not be the most favored pathway in the polymerase active site. Instead, the thermodynamic stabilization of the M_1dG·dC pair that occurs following polymerase bypass, when the pair is located in duplex DNA, may be more important for driving the ring opening event, as judged by the extension kinetics (Table 3). It remains unclear whether desolvation of dC or a spacious active site would more greatly influence the failure to open the M_1dG ring. The crystal structure would suggest that alternative conformations do exist during accurate bypass of M_1dG, which would point to the contribution of a more open active site. Although the ring-opened dC·N^2-OPdG conformation is not observed with Dpo4 (Figure 4), we cannot exclude the possibility that the ring-opened derivative occurs in the active sites of other polymerases.

Similar to the results with M_1dG, the exocyclic DNA adduct 1, N^2-ε-G has been shown to strongly block Dpo4 synthesis (*27*). However, Dpo4-catalyzed bypass of 1,N^2-ε-G yields almost entirely −1 frameshift products by LC−MS/MS analysis, whereas bypass of M_1dG resulted in a substantial fraction of the products representing accurate bypass (Figure 3). Bypass of M_1dG yielded a greater diversity of error-prone products than 1,N^2-ε-G, with a substantial population of "untargeted" misincorporations. In fact, several of the products appear to lack any guidance from the template sequence. However, the presence of accurate bypass products indicates the possibility that at least some fraction of the M_1dG-adducted templates may exist in the ring-opened form. Such a mechanism of accurate bypass would be consistent with results showing that Dpo4 can accurately bypass a number of N^2-alkylG substrates, with a relatively modest decrease in catalytic efficiency. The 14C·M_1dG structure cannot be catalytically competent because the terminal 3′-hydroxyl group is too far removed from the α-phosphate of the incoming dNTP. It is quite reasonable to assume that successful catalytic events occur when the ring-opened N^2-OPdG derivative is paired with dC.

Comparing our results with M_1dG to those obtained with the saturated 1,N^2-propanodeoxyguanosine (PdG) provides some insight into how dATP insertion might occur. PdG is a structural analogue of γ-HOPdG and M_1dG, but PdG does not undergo ring opening (*43*). PdG is, therefore, a useful model of what might happen during Dpo4-catalyzed bypass of an exclusively ring-closed M_1dG adduct. NMR studies have shown that PdG is found in the *syn* conformation when placed opposite cytosine in duplex DNA (*44*). Several type II structures of Dpo4 bound to PdG containing template DNA have been determined (*45*), and all of them show that PdG remains in the *anti* conformation when complexed with Dpo4, similar to the structures for M_1dG reported here. Single-nucleotide insertion experiments reveal that Dpo4 readily inserts dATP opposite PdG (*45*). Though we did not determine the structure of an incoming dATP paired with M_1dG, it is possible that both PdG and M_1dG assume the *syn* orientation during insertion of dATP to form a Hoogsteen base pair similar to what is observed during Dpo4-catalyzed insertion of dATP opposite 7,8-dihydro-8-oxo-deoxyguanosine (*28*). Previously, steady-state kinetic analysis with human pol ι was interpreted as supporting a *syn* conformation for PdG in the active site of pol ι during bypass of the lesion (*46*). Furthermore, a *syn*-oriented template guanosine with an N^2-methylnaphthyl moiety has been observed in the active site of Dpo4 (*47*). Thus, Y-family pols, including Dpo4, can accommodate *syn*-oriented guanosine adducts. The only other base inserted by Dpo4 opposite PdG is dGTP, and this is likely due to the fact that dGTP can pair with dC to the 5′ side of the adduct to form a type II structure. Importantly, Dpo4 does not appear to insert dCTP opposite PdG, based on the experiments reported. It seems likely that Dpo4 can insert dCTP opposite M_1dG more effectively than PdG because ring opening can occur with the former but not the latter moiety. Collectively, these results support the view that a small fraction of M_1dG may undergo ring opening during Dpo4-catalyzed insertion of dCTP, as well as the idea that misinsertion of dATP opposite M_1dG proceeds through a Hoogsteen-like base pair. We have previously reported that human pol η bypasses M_1dG in a highly error-prone fashion, inserting primarily dATP opposite the adduct with a smaller proportion of −1 frameshift events (*48*). Human pol η is more efficient at insertion of dATP opposite M_1dG than Dpo4. In addition, human pol η appears to prefer dATP insertion opposite PdG,

although the efficiency appears to be reduced relative to that of unmodified DNA (*49*). It is of interest to determine if human pol κ bypasses M_1dG in a manner different from that of pol η for at least two reasons: (i) it has been suggested that pol κ may be important during bypass of N^2-dG adducts (*17*, *50*, *51*), and (ii) the N-clasp domain of pol κ could conceivably be important for catalyzing ring opening of M_1dG by encircling the polymerase active site.

In conclusion, our results show that Dpo4 can bypass the exocyclic M_1dG adduct in what is largely an error-prone fashion. Similar to pol η, Dpo4 preferentially inserts dATP opposite M_1dG, but Dpo4 can also generate several other types of products, including −1 frameshift deletions and error-prone extension events. Importantly, Dpo4-catalyzed insertion of dCTP does not readily induce ring opening of M_1dG to form N^2-OPdG based on the efficiency and fidelity measured in kinetic assays and the active site orientation observed in the crystal structure. Ultimately, it is important to know which DNA polymerases replicate M_1dG in vivo and how often they are granted access to the site of damage, as there are undoubtedly multiple enzymes at work during bypass of this lesion, but the results obtained thus far indicate that the mutagenic nature of M_1dG is due in part to the nonfacile nature of ring opening during translesion DNA synthesis.

ACKNOWLEDGMENT

We are grateful to Dr. Z. Wawrzak (Northwestern University, Evanston, IL) for assistance with X-ray diffraction data collection at the Advanced Photon Source (APS). Vanderbilt University is a member institution of the Life Sciences Collaborative Access Team (LS-CAT) at sector 21 of the APS. Use of the APS was supported by the U.S. Department of Energy, Basic Energy Sciences, Office of Science, under Contract W-31-109-Eng-38.

SUPPORTING INFORMATION AVAILABLE

Figures S1 and Tables S1−S29 detailing a large portion of the LC−MS/MS results used to analyze full-length extension by Dpo4 and single-nucleotide insertion experiments that sought to distinguish M_1dG and N^2-OPdG bypass events. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Friedberg, E. C., Walker, G. C., Siede, W., Wood, R. D., Shultz, R. A., and Ellenberger, T. (2006) DNA Repair and Mutagenesis , 2nd ed., ASM Press, Washington, DC.
2. Basu, A. K., O'Hara, S. M., Valladier, P., Stone, K., Mols, O., and Marnett, L. J. (1988) Identification of adducts formed by reaction of guanine nucleosides with malondialdehyde and structurally related aldehydes. *Chem. Res. Toxicol. 1*, 53–59.
3. Dedon, P. C., Plastaras, J. P., Rouzer, C. A., and Marnett, L. J. (1998) Indirect mutagenesis by oxidative DNA damage: Formation of the pyrimidopurinone adduct of deoxyguanosine by base propenal. *Proc. Natl. Acad. Sci. U.S.A. 95*, 11113–11116.
4. Seto, H., Okuda, T., Takesue, T., and Ikemura, T. (1983) Reaction of malonaldehyde with nucleic acid. I. Formation of fluorescent pyrimido[1,2a]purin-10(3H)-one nucleosides. *Bull. Chem. Soc. Jpn. 56*, 1799–1802.
5. Chaudhary, A. K., Nokubo, M., Reddy, G. R., Yeola, S. N., Morrow, J. D., Blair, I. A., and Marnett, L. J. (1994) Detection of endogenous malondialdehyde-deoxyguanosine adducts in human liver. *Science 265*, 1580–1582.
6. Jeong, Y. C., Walker, N. J., Burgin, D. E., Kissling, G., Gupta, M., Kupper, L., Birnbaum, L. S., and Swenberg, J. A. (2008) Accumulation of M1dG DNA adducts after chronic exposure to PCBs, but not from acute exposure to polychlorinated aromatic hydrocarbons. *Free Radical Biol. Med. 45*, 585–591.
7. Sharma, R. A., Ireson, C. R., Verschoyle, R. D., Hill, K. A., Williams, M. L., Leuratti, C., Manson, M. M., Marnett, L. J., Steward, W. P., and Gescher, A. (2001) Effects of dietary curcumin on glutathione S-transferase and malondialdehyde-DNA adducts in rat liver and colon mucosa: relationship with drug levels. *Clin. Cancer Res. 7*, 1452–1458.
8. Fink, S. P., Reddy, G. R., and Marnett, L. J. (1997) Mutagenicity in *Escherichia coli* of the major DNA adduct derived from the endogenous mutagen malondialdehyde. *Proc. Natl. Acad. Sci. U.S.A. 94*, 8652–8657.
9. Hoberg, A. M., Otteneder, M., Marnett, L. J., and Poulsen, H. E. (2004) Measurement of the malondialdehyde-2′-deoxyguanosine adduct in human urine by immuno-extraction and liquid chromatography/atmospheric pressure chemical ionization tandem mass spectrometry. *J. Mass Spectrom. 39*, 38–42.
10. Mukai, F. H., and Goldstein, B. D. (1976) Mutagenicity of malonaldehyde, a decomposition product of peroxidized polyunsaturated fatty acids. *Science 191*, 868–869.
11. Benamira, M., Johnson, K., Chaudhary, A., Bruner, K., Tibbetts, C., and Marnett, L. J. (1995) Induction of mutations by replication of malondialdehyde-modified M13 DNA in *Escherichia coli*: Determination of the extent of DNA modification, genetic requirements for mutagenesis, and types of mutations induced. *Carcinogenesis 16*, 93–99.
12. Niedernhofer, L. J., Daniels, J. S., Rouzer, C. A., Greene, R. E., and Marnett, L. J. (2003) Malondialdehyde, a product of lipid peroxidation, is mutagenic in human cells. *J. Biol. Chem. 278*, 31426–31433.
13. VanderVeen, L. A., Hashim, M. F., Shyr, Y., and Marnett, L. J. (2003) Induction of frameshift and base pair substitution mutations by the major DNA adduct of the endogenous carcinogen malondialdehyde. *Proc. Natl. Acad. Sci. U.S.A. 100*, 14247–14252.
14. Friedberg, E. C., and Gerlach, V. L. (1999) Novel DNA polymerases offer clues to the molecular basis of mutagenesis. *Cell 98*, 413–416.
15. Yang, W., and Woodgate, R. (2007) What a difference a decade makes: Insights into translesion DNA synthesis. *Proc. Natl. Acad. Sci. U.S.A. 104*, 15591–15598.
16. Yang, W. (2003) Damage repair DNA polymerases Y. *Curr. Opin. Struct. Biol. 13*, 23–30.
17. Choi, J. Y., Angel, K. C., and Guengerich, F. P. (2006) Translesion synthesis across bulky N^2-alkyl guanine DNA adducts by human DNA polymerase κ. *J. Biol. Chem. 281*, 21062–21072.
18. Choi, J. Y., Chowdhury, G., Zang, H., Angel, K. C., Vu, C. C., Peterson, L. A., and Guengerich, F. P. (2006) Translesion synthesis across O^6-alkylguanine DNA adducts by recombinant human DNA polymerases. *J. Biol. Chem. 281*, 38244–38256.
19. Choi, J. Y., and Guengerich, F. P. (2005) Adduct size limits efficient and error-free bypass across bulky N^2-guanine DNA lesions by human DNA polymerase η. *J. Mol. Biol. 352*, 72–90.
20. Eoff, R. L., Irimia, A., Angel, K. C., Egli, M., and Guengerich, F. P. (2007) Hydrogen bonding of 7,8-dihydro-8-oxodeoxyguanosine with a charged residue in the little finger domain determines miscoding events in *Sulfolobus solfataricus* DNA polymerase Dpo4. *J. Biol. Chem. 282*, 19831–19843.
21. Eoff, R. L., Irimia, A., Egli, M., and Guengerich, F. P. (2007) *Sulfolobus solfataricus* DNA polymerase Dpo4 is partially inhibited by "wobble" pairing between O^6-methylguanine and cytosine, but accurate bypass is preferred. *J. Biol. Chem. 282*, 1456–1467.
22. Haracska, L., Prakash, L., and Prakash, S. (2002) Role of human DNA polymerase κ as an extender in translesion synthesis. *Proc. Natl. Acad. Sci. U.S.A. 99*, 16000–16005.
23. Haracska, L., Washington, M. T., Prakash, S., and Prakash, L. (2001) Inefficient bypass of an abasic site by DNA polymerase η. *J. Biol. Chem. 276*, 6861–6866.
24. Haracska, L., Yu, S. L., Johnson, R. E., Prakash, L., and Prakash, S. (2000) Efficient and accurate replication in the presence of 7,8-dihydro-8-oxoguanine by DNA polymerase η. *Nat. Genet. 25*, 458–461.
25. Washington, M. T., Johnson, R. E., Prakash, L., and Prakash, S. (2001) Accuracy of lesion bypass by yeast and human DNA polymerase η. *Proc. Natl. Acad. Sci. U.S.A. 98*, 8355–8360.
26. Washington, M. T., Johnson, R. E., Prakash, S., and Prakash, L. (2000) Accuracy of thymine-thymine dimer bypass by *Saccharomyces cerevisiae* DNA polymerase η. *Proc. Natl. Acad. Sci. U.S.A. 97*, 3094–3099.
27. 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,N^2-ethenoguanine. *J. Biol. Chem. 280*, 29750–29764.
28. 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.
29. Mao, H., Schnetz-Boutaud, N. C., Weisenseel, J. P., Marnett, L. J., and Stone, M. P. (1999) Duplex DNA catalyzes the chemical rearrangement of a malondialdehyde deoxyguanosine adduct. *Proc. Natl. Acad. Sci. U.S.A. 96*, 6615–6620.
30. Schnetz-Boutaud, N. C., Saleh, S., Marnett, L. J., and Stone, M. P. (2001) The exocyclic 1,N^2-deoxyguanosine pyrimidopurinone M1G is a chemically stable DNA adduct when placed opposite a two-base deletion in the (CpG)3 frameshift hotspot of the *Salmonella typhimurium* hisD3052 gene. *Biochemistry 40*, 15638–15649.
31. Wang, H., Kozekov, I. D., Kozekova, A., Tamura, P. J., Marnett, L. J., Harris, T. M., and Rizzo, C. J. (2006) Site-specific synthesis of oligonucleotides containing malondialdehyde adducts of deoxyguanosine and deoxyadenosine via a postsynthetic modification strategy. *Chem. Res. Toxicol. 19*, 1467–1474.
32. Szekely, J., Wang, H., Peplowski, K. M., Knutson, C. G., Marnett, L. J., and Rizzo, C. J. (2008) "One-pot" syntheses of malondialdehyde adducts of nucleosides. *Nucleosides, Nucleotides Nucleic Acids 27*, 103–109.
33. Christian, N. P., Reilly, J. P., Mokler, V. R., Wincott, F. E., and Ellington, A. D. (2001) Elucidation of the initial step of oligonucleotide fragmentation in matrix-assisted laser desorption/ionization using modified nucleic acids. *J. Am. Soc. Mass Spectrom. 12*, 744–753.
34. Kabsch, W. (1988) Evaluation of Single Crystal X-Ray Diffraction Data from a Position Sensitive Detector. *J. Appl. Crystallogr. 21*, 916–924.
35. French, G. S., and Wilson, K. S. (1978) On the Treatment of Negative Intensity Observations. *Acta Crystallogr. A34*, 517–525.
36. Vellieux, F. M. D., and Dijkstra, B. W. (1997) Computation of Bhat's OMIT maps with different coefficients. *J. Appl. Crystallogr. 30*, 396–399.
37. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D54*, 905–921.
38. Hashim, M. F., Riggins, J. N., Schnetz-Boutaud, N., Voehler, M., Stone, M. P., and Marnett, L. J. (2004) *In vitro* bypass of malondialdehyde-deoxyguanosine adducts: Differential base selection during extension by the Klenow fragment of DNA polymerase I is the critical determinant of replication outcome. *Biochemistry 43*, 11828–11835.
39. Ling, H., Boudsocq, F., Woodgate, R., and Yang, W. (2001) Crystal structure of a Y-family DNA polymerase in action: A mechanism for error-prone and lesion-bypass replication. *Cell 107*, 91–102.
40. Riggins, J. N., Daniels, J. S., Rouzer, C. A., and Marnett, L. J. (2004) Kinetic and thermodynamic analysis of the hydrolytic ring-opening of the malondialdehyde-deoxyguanosine adduct, 3-(2'-deoxy-β-D-erythro-pentofuranosyl)- pyrimido[1,2-alpha]purin-10(3H)-one. *J. Am. Chem. Soc. 126*, 8237–8243.
41. Riggins, J. N., Pratt, D. A., Voehler, M., Daniels, J. S., and Marnett, L. J. (2004) Kinetics and mechanism of the general-acid-catalyzed ring-closure of the malondialdehyde-DNA adduct, N2-(3-oxo-1-propenyl)deoxyguanosine (N2OPdG-), to 3-(2′-deoxy-β-D-erythro-pentofuranosyl)pyrimido[1,2-α]purin-10(3H)-one (M1dG). *J. Am. Chem. Soc. 126*, 10571–10581.
42. Eoff, R. L., Angel, K. C., Egli, M., and Guengerich, F. P. (2007) Molecular basis of selectivity of nucleoside triphosphate incorporation opposite O^6-benzylguanine by *Sulfolobus solfataricus* DNA polymerase Dpo4: Steady-state and pre-steady-state kinetics and X-ray crystallography of correct and incorrect pairing. *J. Biol. Chem. 282*, 13573–13584.
43. Marinelli, E. R., Johnson, F., Iden, C. R., and Yu, P. L. (1990) Synthesis of 1,N^2-(1,3-propano)-2′-deoxyguanosine and incorporation into oligodeoxynucleotides: A model for exocyclic acrolein-DNA adducts. *Chem. Res. Toxicol. 3*, 49–58.
44. Singh, U. S., Moe, J. G., Reddy, G. R., Weisenseel, J. P., Marnett, L. J., and Stone, M. P. (1993) ^{1}H NMR of an oligodeoxynucleotide containing a propanodeoxyguanosine adduct positioned in a (CG)3 frameshift hotspot of *Salmonella typhimurium* hisD3052: Hoogsteen base-pairing at pH 5.8. *Chem. Res. Toxicol. 6*, 825–836.
45. Wang, Y., Musser, S. K., Saleh, S., Marnett, L. J., Egli, M., and Stone, M. P. (2008) Insertion of dNTPs opposite the 1,N^2-propanodeoxyguanosine adduct by *Sulfolobus solfataricus* P2 DNA polymerase IV. *Biochemistry 47*, 7322–7334.
46. Wolfle, W. T., Johnson, R. E., Minko, I. G., Lloyd, R. S., Prakash, S., and Prakash, L. (2005) Human DNA polymerase ι promotes replication through a ring-closed minor-groove adduct that adopts a *syn* conformation in DNA. *Mol. Cell. Biol. 25*, 8748–8754.
47. Zhang, H., Eoff, R. L., Kozekov, I. D., Rizzo, C. J., Egli, M., and Guengerich, F. P. (2008) Versatility of Y-family *Sulfolobus solfataricus* DNA polymerase Dpo4 in translesion synthesis past bulky N^2-alkyl guanine adducts. *J. Biol. Chem. 284*, 3563–3576.
48. Stafford, J. B., Eoff, R. L., Kozekova, A., Rizzo, C. J., Guengerich, F. P., and Marnett, L. J. (2009) Translesion DNA Synthesis by Human DNA Polymerase η on Templates Containing a Pyrimidopurinone Deoxyguanosine Adduct, 3-(2′-Deoxy-β-D-erythro-pentofuranosyl) pyrimido-[1,2-a]purin-10(3H)-one. *Biochemistry 48*, 471–480.
49. Minko, I. G., Washington, M. T., Kanuri, M., Prakash, L., Prakash, S., and Lloyd, R. S. (2003) Translesion synthesis past acrolein-derived DNA adduct, γ-hydroxypropanodeoxyguanosine, by yeast and human DNA polymerase η. *J. Biol. Chem. 278*, 784–790.
50. Lemée, F., Bavoux, C., Pillaire, M. J., Bieth, A., Machado, C. R., Pena, S. D., Guimbaud, R., Selves, J., Hoffmann, J. S., and Cazaux, C. (2007) Characterization of promoter regulatory elements involved in downexpression of the DNA polymerase κ in colorectal cancer. *Oncogene 26*, 3387–3394.
51. Minko, I. G., Harbut, M. B., Kozekov, I. D., Kozekova, A., Jakobs, P. M., Olson, S. B., Moses, R. E., Harris, T. M., Rizzo, C. J., and Lloyd, R. S. (2008) Role for DNA polymerase κ in the processing of N^2-N^2-guanine interstrand cross-links. *J. Biol. Chem. 283*, 17075–17082.