THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 282, NO. 2, pp. 1456–1467, January 12, 2007

Sulfolobus solfataricus DNA Polymerase Dpo4 Is Partially Inhibited by "Wobble" Pairing between O^6-Methylguanine and Cytosine, but Accurate Bypass Is Preferred*[S]

Received for publication, October 13, 2006, and in revised form, November 10, 2006 Published, JBC Papers in Press, November 14, 2006, DOI 10.1074/jbc.M609661200

Robert L. Eoff, Adriana Irimia, Martin Egli, and F. Peter Guengerich[1]

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

We examined the effect of a single O^6-methylguanine (O^6-MeG) template residue on catalysis by a model Y family polymerase, Dpo4 from *Sulfolobus solfataricus*. Mass spectral analysis of Dpo4-catalyzed extension products revealed that the enzyme accurately bypasses O^6-MeG, with C being the major product (~70%) and T or A being the minor species (~20% or ~10%, respectively), consistent with steady-state kinetic parameters. Transient-state kinetic experiments revealed that k_{pol}, the maximum forward rate constant describing polymerization, for dCTP incorporation opposite O^6-MeG was ~6-fold slower than observed for unmodified G, and no measurable product was observed for dTTP incorporation in the pre-steady state. The lack of any structural information regarding how O^6-MeG paired in a polymerase active site led us to perform x-ray crystallographic studies, which show that "wobble" pairing occurs between C and O^6-MeG. A structure containing T opposite O^6-MeG was solved, but much of the ribose and pyrimidine base density was disordered, in accordance with a much higher $K_{m,dTTP}$ that drives the difference in efficiency between C and T incorporation. The more stabilized C:O^6-MeG pairing reinforces the importance of hydrogen bonding with respect to nucleotide selection within a geometrically tolerant polymerase active site.

Of the myriad forms that covalent modification of DNA can take, alkylation of the purine/pyrimidine bases is one of the most extensively studied (1, 2). The term "alkylating agent" encompasses a variety of known carcinogenic chemicals ranging from the spontaneously reactive nitrogen and sulfur mustards (*e.g.* mechlorethamine) and *N*-alkyl-*N*-nitrosoureas to metabolically activated compounds such as cyclophosphamide and *N*-nitrosamines (3). Guanine is generally considered the most easily oxidized of the bases, and the N7 position is the most nucleophilic atom of guanine. One prevalent form of guanine oxidation occurs at the O6 position, with a simple and widely studied process being addition of a methyl group to form O^6-MeG[2] (4). Methylation of the O6 atom results in alternate pairing schemes that include a O^6-MeG:C "wobble" pairing and a pseudo-"Watson-Crick" O^6-MeG:T pair (Fig. 1), and the relevance of O^6-MeG to mutagenesis is well established (5–7). Cells can repair O^6-MeG by recognition and/or removal of the lesion through either the mismatch repair pathway or through the actions of (O^6-alkylguanine DNA alkyltransferase) (8–10). O^6-MeG is even observed in the DNA of the general population, although the level measured between studies has varied (4, 11). Substantially increased levels of O^6-MeG are found in patients treated with chemotherapeutic regimes that include methylating agents (4, 12, 13).

Of the enzymes associated with what has commonly been referred to as "translesion synthesis," the Y-family DNA polymerases are thought to represent the major constituent present during post-replication repair of covalently modified DNA (14–16). Four human Y-family polymerases are known (η, ι, κ, and Rev1), and representatives also occur in other eukaryotic, archaeal, and prokaryotic systems (17). Current models for translesion synthesis across damaged DNA during replication propose a dynamic exchange between two general groups of polymerases, namely the high fidelity replicative polymerases that perform the vast majority of incorporation events and the Y-family enzymes (14, 16). In mammalian systems the coordination of the four Y-family polymerases, at sites of damage or otherwise, is less than clear at this point. For all of these reasons, the one or more mechanisms by which specialized polymerases bypass damaged DNA is an area of intense focus.

Several crystal structures of the Dpo4 DNA polymerase from *Sulfolobus solfataricus* in complex with covalently modified DNA have served as a major source of structural information regarding how Y-family polymerases bypass damaged DNA

* This work was supported by National Institutes of Health Grants R01 ES010375 (to F. P. G.), F32 CA119776 (to R. L. E.), P30 ES000267 (to F. P. G. and M. E.), and P01 ES05355 (to M. E.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The atomic coordinates and structure factors (code 2j6s (r2j6ssf), 2j6u (r2j6usf), and 2j6t (r2j6tsf)) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

[S] The on-line version of this article (available at http://www.jbc.org) contains supplemental Figs. S1–S10 and Tables S1–S3.

[1] To whom correspondence should be addressed: Dept. of Biochemistry and Center in Molecular Toxicology, Vanderbilt University School of Medicine, 638 Robinson Research Bldg., 23rd and Pierce Avenues, Nashville, TN 37232-0146. Tel.: 615-322-2261; Fax: 615-322-3141; E-mail: f.guengerich@vanderbilt.edu.

[2] The abbreviations used are: MeG, methylguanine; CID, collision-induced dissociation; dCTPαS, 2′-deoxycytidine 5′-*O*-(1-thiotriphosphate); DTT, dithiothreitol; LC, liquid chromatography; LC-MS/MS, liquid chromatography-tandem mass spectrometry; MALDI-TOF, matrix-assisted laser desorption ionization/time-of-flight; MS, mass spectrometry; pol, (DNA) polymerase; pol T7$^-$, bacteriophage pol T7 exonuclease-deficient; RT, reverse transcriptase; HIV, human immunodeficiency virus.

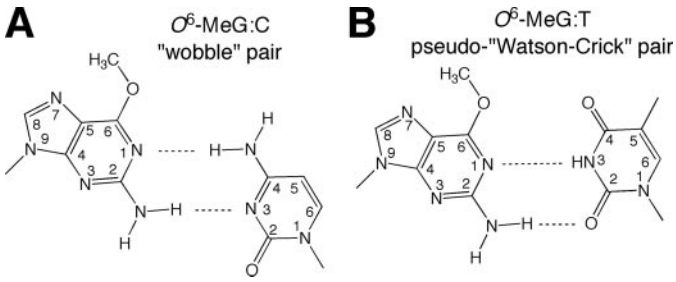

FIGURE 1. **Schematic representation of potential O^6-MeG base pairing orientations.** *A*, O^6-MeG:C wobble base pair. *B*, O^6-MeG:T pseudo-Watson-Crick base pair.

templates (18–23). Rigorous kinetic analysis of Dpo4 catalysis performed with unmodified DNA indicates that the enzyme bears all of the hallmarks of a "translesion" polymerase, namely low efficiency ("low" k_{pol} and "high" $K_{D,dCTP}$), low processivity (~16 incorporation events prior to dissociation), and low "fidelity" (one mistake every few thousand insertions) (24, 25). However, within the context of the cell these attributes are not at all surprising, because copying undamaged DNA does not appear to be the major function of these enzymes. An investigation of Dpo4-catalyzed bypass of a ubiquitous product of oxidative damage, 7,8-dihydro-8-oxodeoxyguanosine, revealed that Dpo4 efficiency is increased ~2-fold during lesion bypass (23). The increased catalytic efficiency is in direct contrast to results obtained with $T7^-$ and other high fidelity polymerases, where catalysis is, in general, greatly inhibited for both C and A incorporation events (26, 27).

In the present study, transient-state kinetic approaches were combined with mass spectral analysis of incorporation/extension products and x-ray crystallography. The results clearly illustrate that Dpo4 favors C incorporation followed by correct extension of at least 4 bp, with T and some A incorporations occurring as minor products. The rate-constant defining Dpo4-catalyzed incorporation of dCTP, k_{pol}, is ~6-fold slower for incorporation opposite O^6-MeG relative to G. The basis for the decreased rate was revealed by the crystal structure to be formation of a wobble base pairing between O^6-MeG and C. From these results some of the mechanistic distinctions between polymerase subfamilies and the subsequent influence of those distinctions upon whether C or T is paired opposite O^6-MeG become apparent.

EXPERIMENTAL PROCEDURES

Materials—Dpo4 was expressed in *Escherichia coli* and purified to electrophoretic homogeneity as described previously (25). All unlabeled dNTPs were obtained from Amersham Biosciences (Piscataway, NJ), S_p-dCTPαS was purchased from Biolog Life Science Institute (Bremen, Germany), and [γ-^{32}P]ATP was purchased from PerkinElmer Life Sciences. All oligonucleotides used in this work were synthesized by Midland Certified Reagent Co. (Midland, TX) and purified using high-performance liquid chromatography by the manufacturer, with analysis by matrix-assisted laser desorption time-of-flight MS. The 13-base primer sequence used in the kinetic and mass spectral analyses was 5′-GGGGGAAGGATTC-3′. The 14-base primer sequences used in the indicated kinetic assays and the crystal structures was 5′-GGGGGAAGGATTCC-3′ for the O^6-MeG:C structure and 5′-GGGGGAAGGATTCT-3′ for the O^6-MeG:T structure. The template DNA sequence used in all of the kinetic and mass spectral assays and in the O^6-MeG:C and O^6-MeG:dATP structures was 5′-TCAT*X*GAATCCTTCCCCC-3′, where X = G or O^6-MeG, as indicated. A second template sequence, used for the O^6-MeG:T structure, was 5′-TCAC*X*GAATCCTTCCCCC-3′, where $X = O^6$-MeG.

Polymerization Assays and Gel Electrophoresis—A ^{32}P-labeled primer, annealed to either an unmodified or adducted template oligonucleotide, was extended in the presence of the indicated dNTP(s). Each reaction was initiated by adding 2 μl of dNTP·Mg^{2+} (250 μM dNTP and 5 mM $MgCl_2$) solution to a preincubated Dpo4·DNA complex (25–100 nM). The reaction was carried out at 37 °C in Tris-HCl (pH 7.8 at 22 °C) buffer containing 50 mM NaCl, 1.0 mM DTT, and 50 μg μl^{-1} bovine serum albumin. At the indicated time, 5-μl aliquots were quenched with 50 μl of 500 mM EDTA, pH 9.0. The samples were then mixed with 100 μl of a 95% formamide/20 mM EDTA solution and were separated on a 20% polyacrylamide (w/v)/7 M urea gel. Products were visualized and quantified using a phosphorimaging screen and Quantity One™ software, respectively (Bio-Rad, Hercules, CA). Formation of an 18-base extension product from a 13-base primer was quantified by fitting the data to Equation 1,

$$f_{18mer}(t) = A\left(1 - \sum_{r=1}^{n}\frac{((k_{obs})t)^{r-1}}{(r-1)!}e^{-(k_{obs})t}\right) + k_2 t \qquad \text{(Eq. 1)}$$

where A is the amount of product formed during the first binding event between Dpo4 and DNA, k_{obs} is the an observed rate constant defining nucleotide incorporation, n is the number of incorporation events required to observe product formation, k_2 is the steady-state rate of nucleotide incorporation, and t is time. All statistical values given indicate the standard error.

Steady-state Kinetics—Dpo4-catalyzed single nucleotide incorporation was measured over a range of dNTP concentrations. All reactions were carried out at 37 °C in 50 mM Tris-HCl (pH 7.8 at 25 °C) buffer containing 50 mM NaCl, 1.0 mM DTT, 50 μg μl^{-1} bovine serum albumin, and 5% glycerol (v/v). Dpo4 (10 nM) was preincubated with DNA (100 nM), and the reaction was initiated by adding dNTP·Mg^{2+}. Aliquots were quenched with 500 mM EDTA (pH 9.0) after varying incubation times. The initial portion of the velocity curve was fit to a linear equation in the program GraphPad Prism (GraphPad, San Diego, CA). The resulting velocity was plotted as a function of dNTP concentration and then fit to a hyperbola, correcting for enzyme concentration to obtain an estimate of k_{cat} and $K_{m,dNTP}$.

Pre-steady-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 7.8 at 25 °C) buffer in the drive syringes. Initially, all RQF experiments were carried out at 37 °C in a buffer containing 50 mM Tris-HCl, pH 7.8 (at 25 °C), 50 mM NaCl, 5 mM DTT, 100 μg μl^{-1} bovine serum albumin, and 5% (v/v) glycerol.

Subsequent experiments indicated that increasing the concentration of glycerol in the reaction mixture resulted in considerably more product in the first binding event for Dpo4-catalyzed incorporation of dCTP opposite O^6-MeG (supplemental Fig. S4). Therefore, the pre-steady-state reactions were repeated using reaction buffer containing 35% glycerol (v/v). Polymerase catalysis was stopped via addition of 500 mM EDTA (pH 9.0). Where indicated, competitor primer/template DNA (1 μM 13/18-mer) was included in the right syringe as a trap for protein, thereby creating single-turnover conditions even under enzyme limiting conditions. Substrate and product DNA was separated by electrophoresis on a 20% polyacrylamide (w/v)/7 M urea gel. The products were then visualized using phosphorimaging and quantitated using Quantity One™ software (Bio-Rad). Results obtained under single-turnover conditions were fit to Equation 2,

$$y = A(1 - e^{-k_{obs}t}) \quad \text{(Eq. 2)}$$

where A is the product formed in first binding event, k_{obs} is the rate constant defining polymerization under the conditions used for the experiment being analyzed, and t is time. Results obtained under conditions that allowed a second round of Dpo4·DNA binding and polymerase action were fit to Equation 3,

$$y = A(1 - e^{-k_{obs}t}) + k_{ss}t \quad \text{(Eq. 3)}$$

where k_{ss} represents a steady-state velocity of nucleotide incorporation.

Liquid Chromatography Mass Spectrometry 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 Mg^{2+} (5 mM) for a final volume of 100 μl. Dpo4 catalysis was allowed to proceed at 37 °C for 4 h in 50 mM Tris-HCl (pH 7.8 at 25 °C) buffer containing 50 mM NaCl, 1 mM DTT, 50 μg μl^{-1} bovine serum albumin, and 5% glycerol (v/v). The reaction was terminated by extraction of the remaining dNTPs by using a size-exclusion chromatography column (Bio-Spin 6 chromatography column, Bio-Rad). Concentrated stocks of Tris-HCl, DTT, and EDTA were added to restore the concentrations to 50 mM, 5 mM, and 1 mM, respectively. Next, *E. coli* uracil DNA glycosylase (20 units, Sigma-Aldrich) was added and the solution was incubated at 37 °C for 6 h to hydrolyze the uracil residue on the extended primer (22). 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 re-suspended in 100 μl of H_2O for MS analysis.

LC-MS/MS analysis was performed on a Waters Aquity ultraperformance liquid chromatography system (Waters, Milford, MA) connected to a Finnigan LTQ mass spectrometer (ThermoElectron Corp., San Jose, CA), operating in the electrospray ionization negative ion mode. An Aquity ultraperformance liquid chromatography BEH octadecylsilane (C_{18}) column (1.7 μm, 1.0 mm × 100 mm) was used with the following LC conditions: buffer A contained 10 mM $NH_4CH_3CO_2$ plus 2% CH_3CN (v/v), and buffer B contained 10 mM $NH_4CH_3CO_2$ plus 95% CH_3CN (v/v). The following gradient program was used with a flow rate of 150 μl min^{-1}: 0–3 min, linear gradient from 100% A to 97%A/3% B (v/v); 3–4.5 min, linear gradient to 80% A/20% B (v/v); 4–5.5 min, linear gradient to 100% B; 5–5.5 min, hold at 100% B; 5.5–6.5 min, linear gradient to 100% A; 6.5–9.5 min, hold at 100% A. The temperature of the column was maintained at 50 °C. Samples were infused with an autosampler system. Electrospray ionization conditions were as follows: source voltage, 4 kV; source current, 100 μA; auxillary gas flow rate setting, 20; sweep gas flow rate setting, 5; sheath gas flow setting, 34; capillary voltage, −49 V; capillary temperature, 350 °C; and tube lens voltage, −90 V. MS/MS conditions were as follows: normalized collision energy, 35%; activation Q, 0.250; and activation time, 30 ms. Product ion spectra were acquired over the range m/z 345–2000. The doubly (negatively) charged species were generally used for CID analysis. The calculations of the CID fragmentations of oligonucleotide sequences were done using a program linked to the Mass Spectrometry Group of Medicinal Chemistry at the University of Utah (www.medlib.med.utah.edu/massspec). The nomenclature used in supplemental Tables S1–S3 has been described previously (28).

Crystallization of Dpo4·DNA Complexes—Dpo4 was concentrated to ~300–550 μM (~12–22 mg ml^{-1}) using a spin concentrator with a 10^4 M_r cutoff filter (Amicon) in 50 mM Tris-HCl (pH 7.4 at 25 °C) buffer containing 200 mM NaCl, 5 mM β-mercaptoethanol, and 10% glycerol (v/v). Dpo4 was then mixed with DNA (1:1.2 molar ratio), incubated at 37 °C for 10 min, centrifuged at 10^4 rpm for 5 min (Eppendorf, centrifuge 5415C) to remove insoluble material, and then placed on ice for 1 h prior to incubation with 1 mM d(N)TP and 5 mM $CaCl_2$. Crystals were grown using the sitting drop, vapor-diffusion method by mixing 1 μl of complex with 1 μl of solution containing 5–10% polyethylene glycol 3350 (w/v), and 100 mM $Ca(OAc)_2$, and equilibrated against a well solution containing 25 mM Tris-HCl (pH 7.4 at 25 °C) buffer, 5–10% polyethylene glycol 3350 (w/v), 100 mM $Ca(OAc)_2$, and 2.5% glycerol (v/v). Crystals were soaked in mother liquor containing an additional 25% polyethylene glycol 3350 (w/v) and 15% ethylene glycol (v/v), and then swiped through paratone-N (Hampton Research, Aliso Viejo, CA) and flash frozen in a stream of liquid nitrogen.

X-ray Diffraction Data Collection and Processing—Diffraction data sets for Dpo4 ternary O^6-MeG:C and O^6-MeG:dATP complexes were collected at 100 K using a radiation wavelength of 1.54 Å on a Bruker Microstar (Bruker AXS, Madison, WI) system housed in the Center for Structural Biology at Vanderbilt. Data sets for the O^6-MeG:T were collected at 110 K using synchrotron radiation wavelength of 0.98 Å on the X25 beamline at the National Synchrotron Light Source, Brookhaven, NY. Indexing and scaling were performed using HKL2000 (29). All three structures indexed to the same space group and had very similar unit cell parameters.

Structure Determination and Refinement—The refined Dpo4-dG model (23) was used as a starting model for the O^6-MeG:C structure, the refined O^6-MeG:C model was used and the starting model for the O^6-MeG:T structure, and the Dpo4–2 model (22) was used as a starting model for the

O^6-MeG:dATP structure. In each instance, several rounds of rigid body refinement of the diffraction data, with gradually increasing resolution, optimized the initial positions of the models. The model was refined further using the CNS Solve package (version 1.1) (30), including simulated annealing, gradient minimization, and refinement of individual isotropic temperature factors. Individual occupancy refinement was necessary to establish the final model for the active site thymidine residue in the O^6-MeG:T structure. Manual model building was performed using TURBO.[3]

RESULTS

Extension of Oligonucleotide Primers by Dpo4 in the Presence of All Four dNTPs—A time course performed under enzyme-limiting conditions provides a general measure of how Dpo4 catalysis is affected by O^6-MeG (Fig. 2). An observed rate constant defining five incorporation events can be measured by following the appearance of the fully extended 18-mer primer and fitting the data to Equation 1, where $n = 5$. The amount of product formed in the first binding event (*i.e.* the "burst" amplitude) is diminished by roughly 25% relative to what is observed with the DNA control indicating that fewer Dpo4 molecules are able to fully extend the primer in the first attempt (Fig. 2*C*). It is important to note that, in the absence of any other evidence (see below), the exact identity of the fully extended product is unknown.

Dpo4 Catalysis in the Presence of a Single dNTP—In the next set of experiments, Dpo4 catalysis was allowed to proceed in the presence of a single nucleotide. Dpo4 can incorporate each of the four dNTPs across from O^6-MeG (supplemental Fig. S1). Steady-state kinetic assays were then employed as a first quantitative measure of the preferential mechanism for Dpo4 insertion opposite O^6-MeG. The relative catalytic efficiency of nucleotide incorporation by Dpo4 was determined by varying the concentration of dNTP in the reaction solution (Table 1). Dpo4-catalyzed incorporation of dCTP opposite O^6-MeG is inhibited ~10^3-fold relative to unmodified DNA, but the enzyme is ~3-, ~6-, and ~14-fold more efficient at correct incorporation of C opposite O^6-MeG compared with incorrect incorporation of T, A, and G, respectively.

LC-MS/MS Analysis of Full-length Extension Products—Unambiguous identification of full-length extension products resulting from Dpo4 catalysis was carried out as described previously (23), with slight modifications that are described under "Experimental Procedures." MS results for Dpo4-catalyzed incorporation opposite and extension past O^6-MeG modified template DNA are summarized in Fig. 3. Two major ions were observed at m/z 1078.6 and 719.1, corresponding with the −2 and −3 ions, respectively (Fig. 3*B*). The total ion trace for the m/z 1079 ion MS/MS is shown in Fig. 3*C*, and CID analysis of the m/z 1079 ion resulted in the fragmentation pattern shown in Fig. 3*D*. The major ions in the fragmentation pattern are consistent with the sequence, 5′-pTCCATGA-3′ (supplemental Table S1), which corresponds to the insertion of C opposite O^6-MeG and accurate full-length extension of the primer.

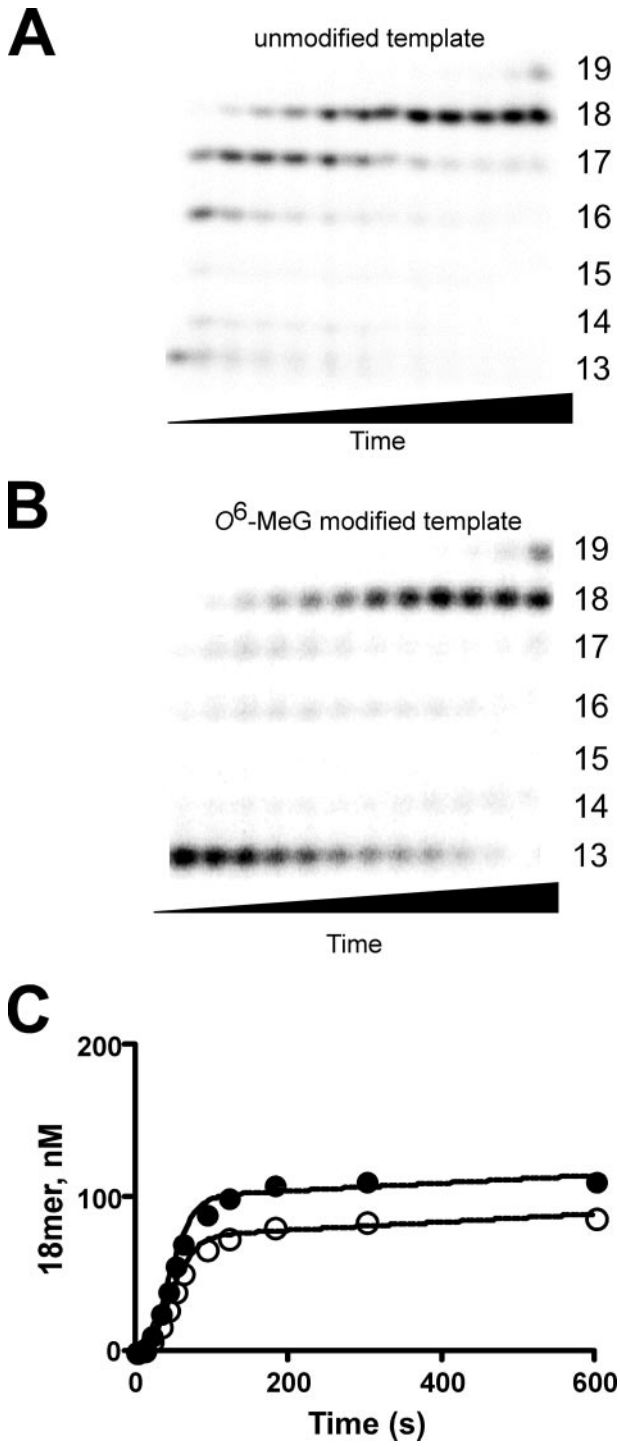

FIGURE 2. **Dpo4-catalyzed incorporation opposite and extension past O^6-MeG adducted DNA.** Dpo4-catalyzed (100 nM) full-length extension of primer/template DNA (200 nM) containing unmodified G (*A*) or O^6-MeG (*B*). The lengths of the oligonucleotide products are indicated on the *right*. *C*, plots of full-length product formation as a function of time with unmodified G (●) or O^6-MeG (○). Results were fit to the Equation 1 to yield the following kinetic parameters: unmodified G: $A = 99 \pm 4$ nM, $k_{obs} = 0.101 \pm 0.003$ s^{-1}, and $k_2 = 0.025 \pm 0.010$ nM s^{-1} (●); O^6-MeG: $A = 73 \pm 2$ nM, $k_{obs} = 0.097 \pm 0.002$ s^{-1}, and $k_2 = 0.027 \pm 0.006$ nM s^{-1} (○).

A second pair of ions was detected at m/z 1086.1 and 723.8, both of which are consistent with the −2 and −3 charge states of a parent ion representing T insertion opposite O^6-MeG followed by accurate full-length extension. CID provided a frag-

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

TABLE 1
Steady-state kinetic parameters for one-base incorporation by Dpo4

Oligomer pair	Primer-template pair	dNTP	k_{cat}	$K_{m,dNTP}$	ΔEfficiency relative to dCTP:G
			s^{-1}	μM	
13-mer 18-mer-1	-G-	dCTP	0.58 ± 0.01	3.0 ± 0.2	
13-mer 18-mer-1	-O^6MeG-	dCTP	0.071 ± 0.005	340 ± 70	950-fold less
13-mer 18-mer-1	-G-	dTTP	0.11 ± 0.01	2100 ± 500	3600-fold less
13-mer 18-mer-1	-O^6MeG-	dTTP	0.088 ± 0.003	1200 ± 120	2600-fold less
13-mer 18-mer-1	-G-	dATP	0.012 ± 0.001	225 ± 22	3600-fold less
13-mer 18-mer-1	-O^6MeG-	dATP	0.0022 ± 0.0001	71 ± 16	6000-fold less
13-mer 18-mer-1	-G-	dGTP	0.012 ± 0.001	760 ± 130	12000-fold less
13-mer 18-mer-1	-O^6MeG-	dGTP	0.008 ± 0.001	590 ± 70	13500-fold less

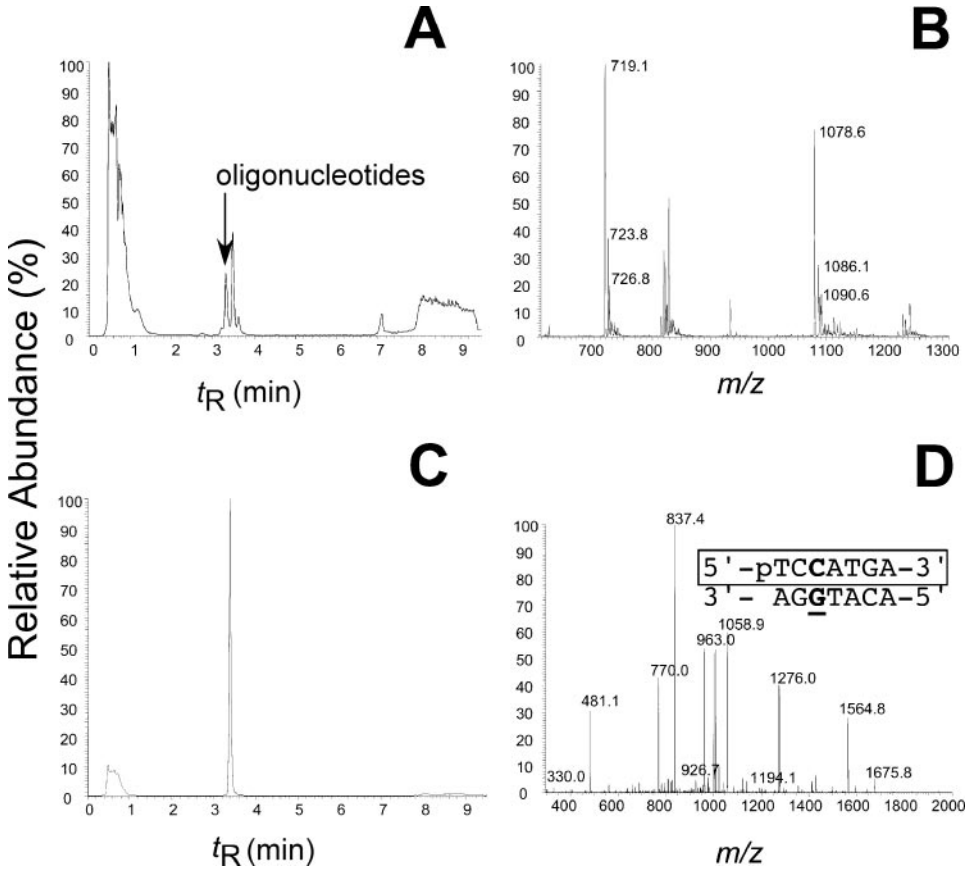

FIGURE 3. **Identification of Dpo4-catalyzed full-length extension products by LC-MS/MS.** *A*, total ion current trace of products derived from extension of 13/18-mer DNA containing O^6-MeG. *B*, electrospray ionization mass spectrum of the oligonucleotide peaks that elute at 3.4 min. *C*, total ion current trace of ion *m/z* 1079. *D*, CID mass spectrum of ion *m/z* 1079.

mentation pattern consistent with this sequence assignment (supplemental Fig. S2 and Table S2). The third ion pair, 1090.7 and 726.8, was identified as A insertion products, again followed by accurate full-length extension (supplemental Fig. S3 and Table S3). Comparison of the selected ion counts (for the ions corresponding to all three products) indicates that correct incorporation of C opposite O^6-MeG comprises roughly 70% of the full-length extension products observed in the reaction mixture. Misincorporation of T accounted for ~20% of the products, and A accounted for the remaining ~10%, consistent with the steady-state parameters.

Transient-state Kinetic Analysis for Dpo4 Bypass of O^6-MeG—Pre-steady-state experiments were performed under enzyme-limiting conditions in the presence of dCTP alone. The presence of O^6-MeG in the template strand reduces the amount of product generated in the first catalytic turnover by roughly 3-fold under the experimental conditions used here (Fig. 4). The k_{obs} value is 4.4-fold slower for incorporation opposite O^6-MeG compared with undamaged DNA at this particular concentration of nucleoside triphosphate (1 mM dCTP).

Previous studies suggest that phosphoryl transfer (*i.e.* the "chemistry" step) is not the rate-limiting step that defines correct dNTP incorporation by Dpo4 (24). One approach to determining the overall contribution of "chemistry" to the polymerase catalytic cycle involves substituting sulfur for one of the oxygen atoms in the α-phosphate group. In principle, substitution of oxygen with a less electronegative sulfur atom makes bond breakage (and subsequent phosphoryl transfer) more difficult. If phosphoryl transfer is the rate-limiting step in the multistep polymerase catalytic cycle then the sulfur substitution experiment will exhibit a

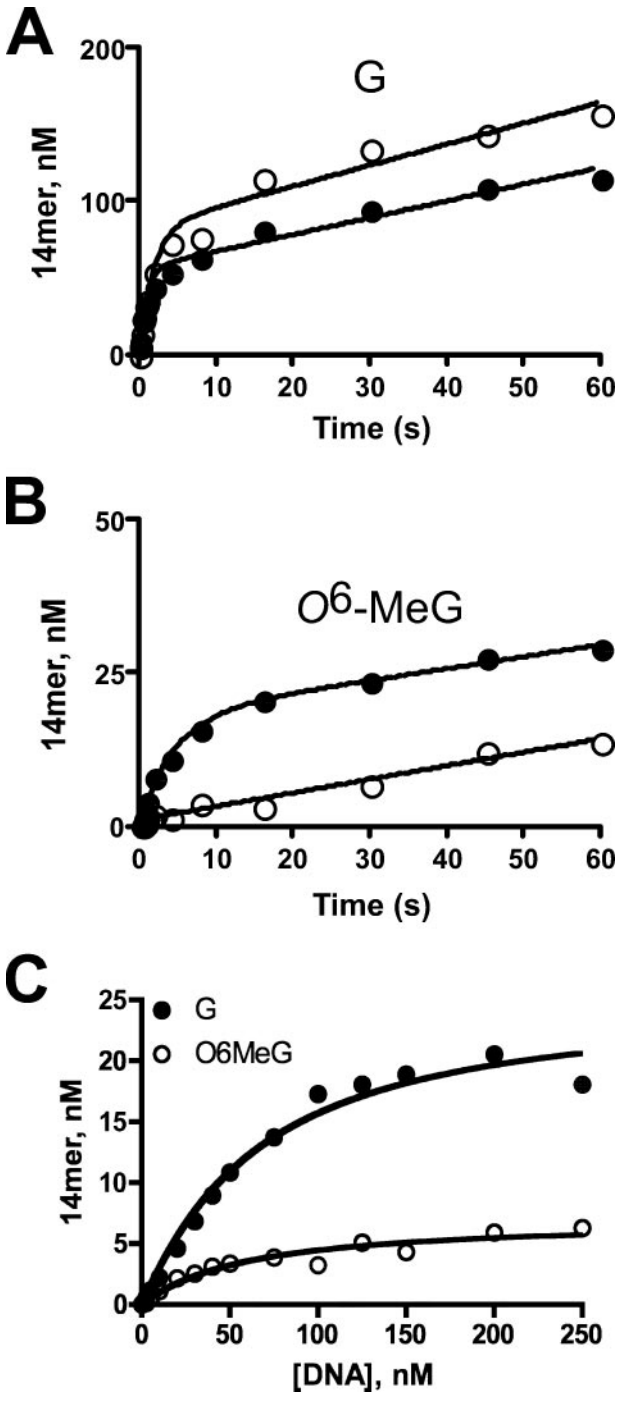

FIGURE 4. **Pre-steady-state analysis of Dpo4-catalyzed incorporation of dCTP.** *A*, Dpo4 (100 nM) was incubated with primer/template DNA (200 nM) containing unmodified G and 1 mM (●) dCTP or (○) S_p-dCTPαS. The results were then fit to Equation 3, and the following parameters were obtained: dCTP (●): $A = 56 \pm 4$ nM, $k_{obs} = 0.98 \pm 0.21$ s^{-1}, $k_{ss} = 1.10 \pm 0.13$ nM s^{-1}. S_p-dCTPαS (○): $A = 82 \pm 7$ nM, $k_{obs} = 0.52 \pm 0.11$ s^{-1}, $k_{ss} = 1.37 \pm 0.18$ nM s^{-1}. *B*, Dpo4 (100 nM) was incubated with primer/template DNA (200 nM) containing O^6-MeG and 1 mM (●) dCTP or (○) S_p-dCTPαS. The results for dCTP incorporation (●) were then fit to Equation 2, and the following parameters were obtained: dCTP (●): $A = 18 \pm 1$ nM, $k_{obs} = 0.22 \pm 0.02$ s^{-1}, $k_{ss} = 0.19 \pm 0.02$ s^{-1}. The results for S_p-dCTPαS incorporation (○) were fit to a linear equation, which provides a steady-state rate of nucleotide incorporation, $k_{ss} = 0.23 \pm 0.01$ s^{-1}. *C*, single-turnover product amplitude was plotted as a function of DNA concentration (2.5–250 nM). The results were then fit to a quadratic equation to obtain a measure of $K_{D,DNA}$: G (●): $A = 25 \pm 1$ nM, $K_{D,DNA} = 51 \pm 9$ nM; O^6-MeG (○): $A = 7 \pm 1$ nM, $K_{D,DNA} = 55 \pm 14$ nM. The concentrations of Dpo4, dCTP, and unlabeled primer/template DNA trap were 25 nM, 1 mM, and 1 mM, respectively.

decreased rate of nucleotide incorporation, although interpretation of such changes are a matter of some debate (32). The measured reduction in k_{obs} upon substitution of sulfur for oxygen ("thio" effect) for unmodified G was ~1.9 (Fig. 4*A*), which is similar to previous phosphorothioate substitution experiments with Dpo4 that resulted in a thio effect of ~1.4 (24). For incorporation opposite O^6-MeG-modified DNA, substitution of dCTP with S_p-dCTPαS eliminates any observable burst in product formation (Fig. 4*B*), suggesting that phosphoryl transfer by Dpo4 is partially inhibited in the presence of O^6-MeG-modified template DNA. Additional evidence in favor of such an interpretation includes the observation that the amount of product formed in the first binding event increases as a function of glycerol concentration (supplemental Fig. S4). Any conformational change in the enzyme structure is likely to be slowed by a more viscous reaction solution, as has been observed for pol β (33). The rate of phosphoryl transfer should not, however, be greatly affected by increasing viscosity. Thus, the increasing product amplitudes observed at higher glycerol concentrations may be attributable to a higher overall residence time for Dpo4 on any given substrate, which allows a greater fraction of Dpo4 to complete the chemistry step prior to dissociation. It is possible that the Dpo4·DNA complex is inherently less stable in the presence of O^6-MeG-adducted DNA. Active site titration experiments were performed to assess the relative stability of Dpo4·G and Dpo4·O^6-MeG complexes when the enzyme is required to proceed through the phosphoryl transfer step (Fig. 4*C*). The measured $K_{D,DNA}$ values for both unmodified and O^6-MeG-modified DNA substrates are similar, indicating that O^6-MeG has no effect upon the intrinsic stability of the ternary complex.

To further assess the Dpo4 mechanism of O^6-alkylG bypass, the concentration of dCTP in the reaction mixture was varied to measure the maximum forward rate constant describing polymerization, k_{pol}, and the equilibrium dissociation constant, $K_{D,dCTP}$, describing dCTP binding affinity toward the Dpo4·DNA complex (Fig. 5 and Table 2). The overall kinetic efficiency of the control reaction, $k_{pol}/K_{D,dCTP}$, is similar to that measured previously (23). Dpo4-catalyzed incorporation of dCTP opposite O^6-MeG-modified DNA proceeds at a rate approximately ~6-fold slower, as evidenced by a decreased value for k_{pol}, and the apparent nucleotide binding affinity is ~2-fold less than that of unmodified G (*i.e.* $K_{D,dCTP}$ is roughly 2-fold higher for O^6-MeG). The overall catalytic efficiency is decreased 14-fold when Dpo4 attempts to insert dCTP opposite O^6-MeG. Pre-steady-state kinetic analysis of dTTP misincorporation could not be determined due to an absence of any measurable amount of product formed in the first Dpo4·DNA binding event under all conditions tested.

Incorporation Beyond O^6-MeG Is Unperturbed and Retains Normal Fidelity—To study the catalytic cycle that occurs immediately following nucleotide incorporation opposite O^6-MeG, two synthetic 14-mer primers were used. One contained C-paired opposite O^6-MeG and a second contained T-paired opposite O^6-MeG. Correct incorporation of dATP proceeds at a maximal forward rate quite comparable to unmodified G when C is paired opposite O^6-MeG (Table 2 and supplemental Fig. S5). Notably, the next base extension of the

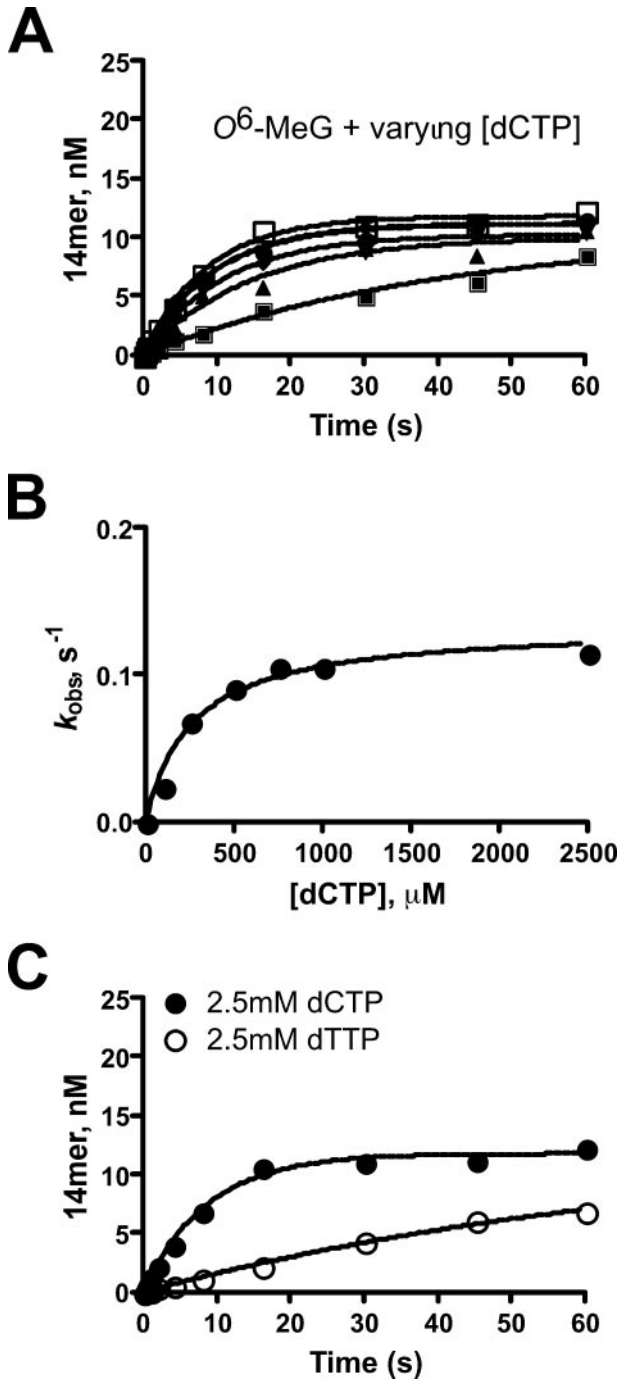

FIGURE 5. **Determination of k_{pol} and K_D, dCTP for Dpo4-catalyzed incorporation of dCTP opposite O^6-MeG.** *A*, measurement of Dpo4-catalyzed incorporation opposite O^6-MeG at varying concentrations of dCTP. *B*, the observed rates of nucleotide incorporation were plotted as a function of dCTP concentration and fit to a quadratic equation to yield kinetic parameters (Table 2). *C*, comparison of Dpo4-catalyzed incorporation of 2.5 mM (●) dCTP and (○) dTTP.

mispaired T:O^6-MeG proceeded at a 2.7-fold slower forward rate than extension beyond the correct pair of C opposite O^6-MeG (Table 2 and supplemental Fig. S5). The $K_{D,dCTP}$ was similar, in both instances, to what is observed for unmodified G. Comparing the efficiency of next base extension reveals that extension of a correct C:O^6-MeG pairing is basically identical to unmodified G, but the efficiency for next base extension of a

TABLE 2
Pre-steady-state kinetic parameters for one-base incorporation by Dpo4

Oligomer pair	Primer-template pair	dNTP	k_{pol}	$K_{D,dCTP}$	$k_{pol}/K_{D,dCTP}$
			s^{-1}	μM	$s^{-1}\mu M^{-1}$
13-mer 18-mer	-G-	dCTP	0.78 ± 0.08	114 ± 54	0.0068
13-mer 18-mer	-O^6MeG-	dCTP	0.13 ± 0.01	263 ± 60	0.00049
13-mer 18-mer	-O^6MeG-	dTTP	ND[a]	ND	ND
14-mer 18-mer	-C -O^6MeGT-	dATP	0.67 ± 0.04	72 ± 24	0.0093
14-mer 18-mer	-T -O^6MeGT-	dATP	0.25 ± 0.01	88 ± 16	0.0028

[a] ND, not determined due to lack of sufficient product formation in first binding event.

T:O^6-MeG mispair is decreased ~2- to 3-fold. Thus, once Dpo4 has managed to incorporate C opposite O^6-MeG, extension beyond the point of damage appears to be quite normal. In contrast, misincorporation of dTTP is strongly impeded at the site of modification, and this inhibition is retained somewhat during next base extension beyond the T:O^6-MeG mispair.

Based on the pre-steady-state kinetics, it is apparent that Dpo4-catalyzed incorporation of both dCTP and dTTP opposite the O^6-MeG lesion is impeded, with misincorporation events suffering a greater degree of inhibition. The slower forward rate and the complete loss of any detectable burst for S_p-dCTPαS incorporation are both consistent with inhibition of phosphoryl transfer.

X-ray Crystal Structures of Dpo4 Complexes—Crystallogenesis experiments were performed for all potential pairings opposite O^6-MeG to provide a structural understanding of the kinetic phenomena. Crystals with C and T primer DNA residues paired opposite O^6-MeG were readily obtained, but crystallization of the incoming dCTP or dTTP using a 13-nucleotide primer proved more difficult. Crystals were obtained using a 13-nucleotide primer and either dCTP or dTTP, but none of these crystals diffracted to high resolution. Additionally, using the dNTP that pairs with the base to the 5′-side of O^6-MeG resulted in large numbers of crystals. All crystals were obtained using Ca^{2+} as the metal ion cofactor. Three complexes were obtained that provided a structural basis for what was observed in both the mass spectral and kinetic analyses (Table 3). In the first complex (O^6-MeG:C), C is paired opposite O^6-MeG in what would be considered a post-insertion context. In the second complex (O^6-MeG:T) T is paired opposite O^6-MeG, again in a post-insertion register. The third complex (O^6-MeG:dATP) included a dATP molecule.

The initial O^6-MeG:C structure was solved by molecular replacement using the previously reported Dpo4-dG structure (PDB accession code 2c22) as a search model (23). In contrast to the Dpo4-dG structure, the O^6-MeG:C structure is of the Type I form (19), with only one template base accommodated directly in the active site of the enzyme. The O^6-MeG lesion is well stacked within the template strand, and the methyl group at position 6 is placed in a proximal orientation relative to the remainder of the purine ring system (Fig. 6*A*). The cytosine base (pC14) positioned opposite O^6-MeG has a shifted base-pairing

TABLE 3
Crystal data and refinement parameters

Parameter	*O*⁶MeG:C	*O*⁶MeG:T	*O*⁶MeG:dATP
Crystal data			
X-ray source	VCSB[a]	NSLS	VCSB
Beamline	Microstar	X25	Microstar
Detector	PT135 CCD	Quantum CCD	PT135 CCD
Wavelength (Å)	1.54	0.98	1.54
Temperature (K)	100	110	100
No. of crystals	1	1	1
Space group	$P2_12_12$	$P2_12_12$	$P2_12_12$
Unit cell (*a, b, c;* Å)	93.94, 104.9, 52.57	92.32, 102.28, 52.48	95.06, 102.98, 53.02
Resolution range (Å)	18.1–2.50	30.0–2.50	22.5–2.60
Highest resolution shell[b]	(2.60–2.50)	(2.60–2.50)	(2.70–2.60)
No. of measurements	122,742	267,671	72,824
No. of unique reflections	18,215 (1773)	17,652 (1517)	16,650 (1539)
Redundancy	6.69 (5.10)	15.10 (9.50)	4.35 (3.26)
Completeness (%)	99.5 (100)	98.8 (92.1)	99.6 (98.6)
R-merge[c] (%)	10.1	7.8	10.7
Signal to noise ()	14.36 (2.51)	32.17 (3.18)	9.94 (1.40)
Solvent content (%)	51.54	49.18	53.61
Model composition			
No. of amino acid residues	341	344	341
No. of water molecules	125	147	89
No. of Ca^{2+} ions	3	3	3
No. of template nucleotides	16	17	17
No. of primer nucleotides	14	14	13
No. of dATPs	1		1
No. of dGTPs		1	
R_f (%)[d]	23.4	21.7	23.7
R-free (%)[e]	28.0	26.5	26.9
Estimated coordinate error (Å)			
From Luzatti plot	0.36	0.35	0.39
From Luzatti plot (*c*-v)[f]	0.43	0.43	0.46
From σA plot	0.43	0.34	0.47
From σA plot (*c*-v)[f]	0.46	0.43	0.49
Temperature factors			
From Wilson plot (Å²)	42.8	57.7	45.8
mean isotropic (Å²)	38.3	56.0	43.8
r.m.s.d. in temperature factors			
Bonded main chain atoms (Å²)	1.30	1.72	1.23
Bonded side chain atoms (Å²)	1.89	2.74	1.81
r.m.s. standard deviation from ideal values			
Bond lengths (Å)	0.008	0.007	0.008
Bond angles (°)	1.2	1.4	1.3
Dihedral angles (°)	22.6	22.8	22.2
Improper angles (°)	0.98	3.46	1.03

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

orientation that has been referred to as a wobble pairing scheme (Fig. 7*A*). A wobble pair between *O*⁶-MeG and pC14 yields two hydrogen-bonding partners. The N1 atom of *O*⁶-MeG and the N4 exocyclic amino group of C are situated 2.8 Å apart. The second pair includes the N2 exocyclic amino group of *O*⁶-MeG and the N3 atom of C, which are situated 2.4 Å apart. The overall structure represents a post-insertion complex with Dpo4 poised to insert dATP. As with the other structures (*O*⁶-MeG:T and *O*⁶-MeG:dATP), three Ca^{2+} ions were found in or near the active site of the enzyme. Two Ca^{2+} ions are coordinated directly in the active site of Dpo4. One of these ions is considered to be the catalytic ion, and the second ion appears to stabilize the tri-phosphate moiety of the incoming dNTP. A third Ca^{2+} ion is located near the phosphate group of pC13.

The initial *O*⁶-MeG:T model did not contain a base opposite the templating lesion, and the initial $F_o - F_c$ map failed to reveal convincing positive electron density for much of the thymidine base (supplemental Fig. S6*A*). The next round of refinement included a thymidine at position 14 in the primer DNA (pT14), but the resulting difference maps showed negative density for much of the glycosidic bond region of the thymidine base (supplemental Fig. S6*B*). Subsequent rounds of refinement and manual model building also failed to improve the electron density near pT14. The remainder of the *O*⁶-MeG:T active site is very similar to the *O*⁶-MeG:C structure in that the *O*⁶-MeG lesion is well stacked in the template DNA (again located in a proximal position relative to the purine ring), and the incoming dGTP is paired with the cytosine located 5′ to *O*⁶-MeG residue (Fig. 6*B*). The final *O*⁶-MeG:T model places pT14 in three alternate conformations (Fig. 7*B*). Using multiple conformations was the only means of eliminating positive electron density from those regions of the map near pT14 (supplemental Fig. S7). The absence of unambiguous electron density near pT14 most likely results from disorder in this particular region of the crystal.

In the third structure, MeG-dATP, the incoming dATP was found to pair with the thymidine base to the 5′-side of *O*⁶-MeG (Fig. 6*C*). Such a structure is most likely representative of a

severely inhibited catalytic complex. As with the other structures, the O^6-MeG is well stacked in the template strand, and the methyl group is placed in a proximal orientation relative to the purine ring. The orientation of the methyl group in the O^6-MeG-dATP structure is somewhat surprising given the apparent preference for a distal orientation in the free nucleoside (34, 35). Without a base paired opposite, a reorientation to the more sterically more favorable distal position would seem logical. In fact, using the experimental electron density as a guide, the O6 methyl group can be manually positioned in the distal orientation. However, after every round of refinement the O6 methyl group was located in the proximal orientation and is considered to be the most favored orientation for the structures observed here. Overall, the final refined structures showed very little deviation (root mean square deviation = 0.78) from one another, except in certain salient aspects of the active site region (supplemental Fig. S8).

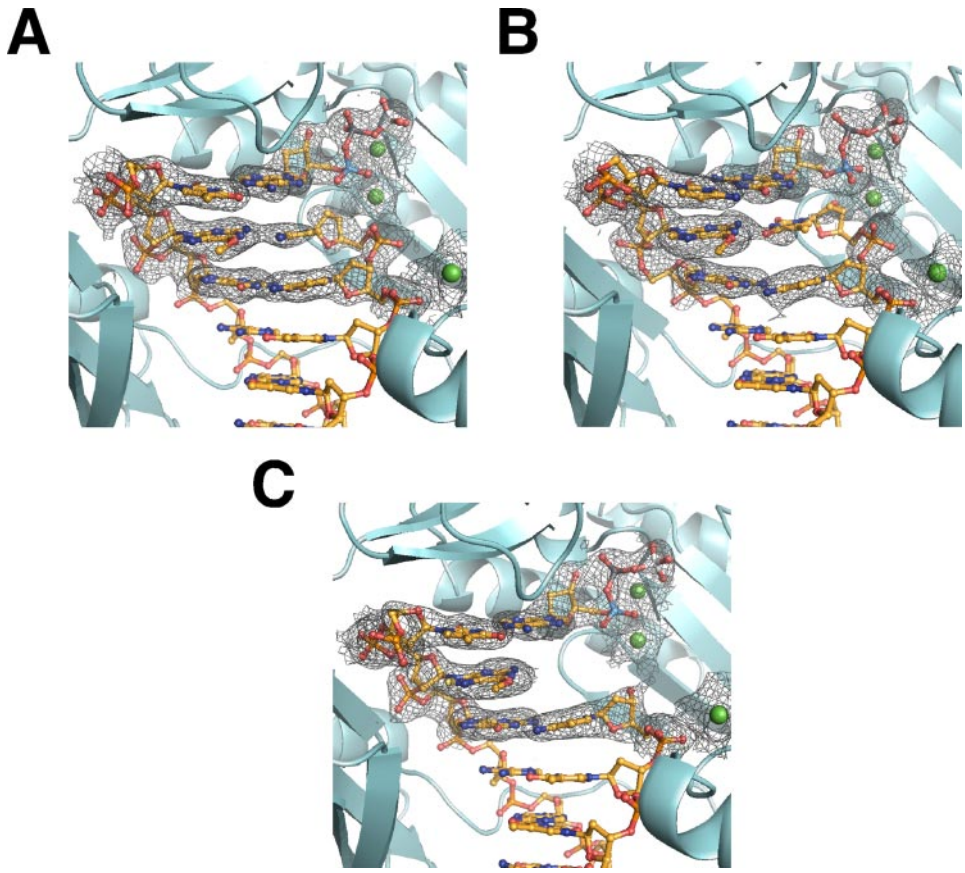

FIGURE 6. **Electron density and DNA duplex conformations observed at the active site of ternary complexes.** *A*, O^6-MeG:C; *B*, O^6-MeG:T; and *C*, O^6-MeG:dATP. Dpo4 is shown in schematic form (*cyan*). The DNA duplex is shown in ball-and-stick representation. Calcium ions are shown as *green spheres*. The electron density map (*black mesh*) $3F_o - 2F_c$ is contoured at the 1σ level.

DISCUSSION

Alkylation of DNA can take many forms. One of the simplest forms of DNA alkylation is addition of a methyl group at the O^6-position of guanine. The work presented herein involved use of a model Y-family polymerase, namely Dpo4 from *S. solfataricus*, as a means of clearly defining how this specialized polymerase processes the O^6-MeG lesion.

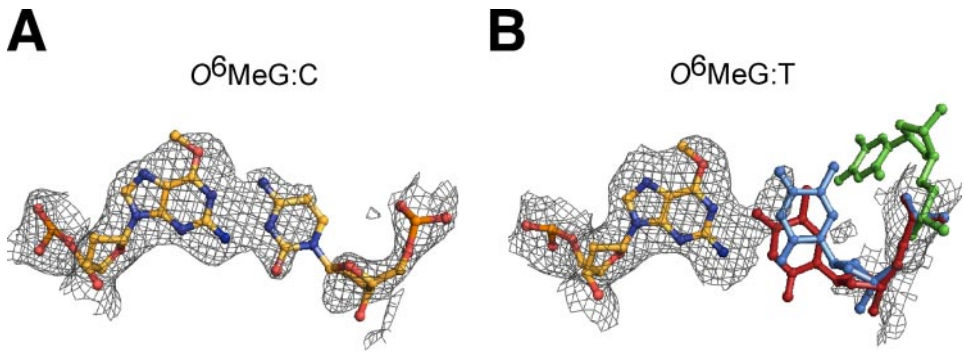

FIGURE 7. **Comparative C and T base pairing orientations opposite O^6-MeG at the catalytic center of the ternary complexes.** *A*, the 14th nucleotide C forms a wobble base pair with template O^6-MeG in the O^6-MeG:C structure. *B*, the 14th nucleotide T is positioned in three conformations opposite O^6-MeG, owing to the disorder observed in the region near thymidine in the O^6-MeG:T structure. The electron density map (*black mesh*) $3F_o - 2F_c$ is contoured at the 1σ level.

The mutagenic potential of O^6-MeG is apparent from the structure of the lesion, in that the modified base loses some of its ability to form the hydrogen bonding pattern normally observed within a G:C pair (Fig. 1). Upon addition of the methyl group, formation of the enol moiety at the O6 position redistributes some of the electron density into the aromatic ring, thereby reducing the ability of the N1 atom to act as a general base and become protonated. Without a protonated N1 atom, the most obvious predicted hydrogen-bonding pattern between O^6-MeG and C would then consist of two hydrogen bonds: the first occurring between the N1 of O^6-MeG and the exocyclic amine at position N4 of C, and the second occurring between the N2 exocyclic amine of O^6-MeG and the N3 atom of C (Fig. 1*A*). The existence of a wobble base pair between O^6-MeG and C was predicted 20 years ago based on NMR studies of oligonucleotides (in the absence of polymerase) (36). Two other types of O^6-alkylG:C pairing schemes have been reported in separate crystallographic studies of the modified oligonucleotides in isolation, including a split-hydrogen bond "bifurcated" pairing and a Watson-Crick type pairing in Z-form DNA (37, 38). Nevertheless, the O^6-MeG:C structure solved here provides conclusive evidence that a wobble pairing occurs in the active site of Dpo4 and very likely represents what occurs

within the context of post-replication repair by the Y-family DNA polymerases (Fig. 7*A*).

As indicated by the steady-state kinetic and mass spectral analyses, Dpo4-catalyzed bypass of O^6-MeG is not perfectly accurate (Table 1 and supplemental Figs. S2 and S3). Misincorporation of T was the second most observed product of Dpo4-catalyzed bypass. Indeed, typical mutation spectra observed when cells are treated with methylating agents are dominated by GC to AT transitions (4). The O^6-MeG:T structure solved here provides a rather ambiguous answer to the question of how O^6-MeG and T pair in the active site of a polymerase, because much of the corresponding thymidine base is disordered. However, analysis of the three conformations describing the position of pT14 reveals a possible rationale for why the base is disordered and how thymidine proceeds to pair opposite O^6-MeG-modified DNA.

The first conformation (Fig. 7*B*, *red* residue) places T in a position where a single hydrogen bond may occur between the N2 exocyclic amino group of O^6-MeG and the O4 atom of pT14. The second conformation (Fig. 7*B*, *blue* residue) assumes a pseudo-Watson-Crick geometry, but the closest pair of hydrogen-bonding partners (namely the exocyclic amino portion of O^6-MeG and the O2 atom of pT14) are 3.1 Å apart. Finally, the third conformation swings the sugar and base portion of pT14 out of the active site completely (Fig. 7*B*, *green* residue). Taken together, these three conformations are likely to represent the course of events undertaken by any T residue attempting to pair opposite O^6-MeG at the terminus of a DNA helix, a reasonable conclusion supported by the kinetic and mass spectral data.

The pairing between O^6-MeG and T was predicted by NMR to retain Watson-Crick geometry with a single hydrogen bond occurring between the N2 exocyclic amine of O^6-MeG and the O2 carbonyl group of T, and either no bonding or possibly a "long" bond occurring between N1 of O^6-MeG and the imino proton at position N3 of T (Fig. 1*B*) (39). Crystallographic analysis of a self-complementary dodecamer, which pairs O^6-MeG and T, revealed a Watson-Crick type pairing geometry (40). In the crystal structure of the dodecamer, the hydrogen bond between N1 of O^6-MeG and N3 of T is 2.9 Å in length, slightly closer than what was proposed by NMR (40). In our O^6-MeG:T structure the N1 atom of O^6-MeG and N3 atom of T are separated by 3.4 Å (although caution should be taken in placing too much emphasis upon the distances observed between O^6-MeG and T because of the sparse density found near that region and the resolution of our data). Further, the difference observed between the NMR studies and the crystallographic analysis of the modified DNA alone is an interesting point to consider within the context of the O^6-MeG:T structure observed here. In NMR studies with O^6-MeG-modified oligonucleotides, the O^6-MeG:T pair is separated from the end of the double-helix by two base pairs, and in the crystallographic analysis the O^6-MeG:T pair is separated from the end of the double-helix by three base pairs, possibly affecting the strength of the hydrogen bond between N1 of O^6-MeG and the imino proton at position N3 of T because of thermal fraying at the end of the dodecamer (39, 40). Melting studies with the dodecamer revealed that, in thermodynamic terms, both the C and T pairings with O^6-MeG pairing have a large destabilizing effect upon the B-form helix (40). In our structure, the pairing between T and O^6-MeG occurs at the very end of a DNA double-helix, in an environment that places fewer restraints upon the orientation of pT14. The potential hydrogen-bonding partners in the pseudo-Watson-Crick conformation (Fig. 7*B*, *blue* residue) are placed at distances of ~3.1 Å or greater, in general agreement with the view that thermal fraying occurs more readily near the ends of double-stranded DNA.

The literature defining how polymerases handle O^6-alkylguanines is somewhat limited in scope. Until now, there has been no structural information available for any of the polymerase families regarding how an O^6-alkylG lesion is accommodated in a polymerase active site and what types of pairing orientations are observed during catalysis. The "high fidelity" polymerases typically favor inserting a T opposite the O^6-MeG lesion (41–46). Studies with yeast polymerase η have revealed a more accurate mechanism of bypass in which C is the favored insertion product, at least in the steady-state (41). Interestingly however, mass spectral analysis of the full-length extension products indicated that human pol η inserts more T during full-length extension in the presence of all four dNTPs (47).

Previous work with Klenow fragment of *E. coli* DNA polymerase I revealed that neither T nor C incorporation was very facile opposite O^6-MeG (46). Steady-state experiments from our own group showed that both bacteriophage polymerase $T7^-$ and HIV-RT preferentially incorporate dTTP opposite O^6-MeG (44, 45). Notably, the catalytic efficiency for multiple turnovers was decreased substantially for both enzymes (100- to 1,000-fold for $T7^-$ and 3,000- to 6,000-fold for HIV-RT) relative to correct incorporation opposite unmodified DNA (44, 45). In pre-steady-state experiments, HIV-RT-catalyzed incorporation of both C and T opposite O^6-MeG was inhibited ~7-fold relative to correct incorporation opposite G (44, 45). Interestingly, in the pre-steady-state HIV-RT incorporated both C and T equally well opposite the lesion (44, 45). A general conclusion from both studies was that the Watson-Crick geometry of the T:O^6-MeG pairing was favored by Klenow fragment and HIV-RT over the more distorted backbone of the C:O^6-MeG pairing (44–46). Dpo4 follows the kinetic trend of HIV-RT in that the level of inhibition is greater when the enzyme is allowed to undergo multiple rounds of binding and dissociation, but the ability of Dpo4 to bypass O^6-MeG (in what is preferably an accurate manner) distinguishes it from most polymerases studied to date.

Pre-steady-state results for Dpo4-catalyzed insertion of C opposite O^6-MeG suggest that phosphoryl transfer is the rate-limiting step in the reaction mechanism. In a previous kinetic analysis a step prior to chemisty was proposed to be the rate-determining step during "correct" nucleotide incorporation by Dpo4 (24), but what that kinetic step represents in physical terms remains unclear. In the same study, phosphoryl transfer was proposed to limit progression of the reaction during "incorrect" insertion events (24). In kinetic terms, correct bypass of O^6-MeG is similar to incorrect insertion events with unmodified DNA. Both events are largely defined by slow forward rates, although binding affinities are also reduced to varying degrees. The O^6-MeG:C pair is in a "wobble" conformation, not the nor-

mal Watson-Crick geometry. Likewise, incorrect insertion events with unmodified DNA would be expected to adopt non-Watson-Crick geometry (excluding those events where the incoming dNTP is paired 5′ to the template base), but it remains unclear why chemistry becomes the rate-limiting kinetic step. Superimposition of the O^6-MeG:C structure with the original Type I (19) structure for Dpo4 reveals very few changes in the amino acid contacts between enzyme and substrate (supplemental Fig. S9). In comparing the original Type I (19) structure (PDB accession code 1jx4) with our O^6-MeG:C structure it appears that the enzyme reorganizes a few amino acid residues near the template to reach the free energy minimum necessary to accommodate the wobble base pair (supplemental Fig. S10). The side-chain reorientations include movement of Arg-332 and Ile-145 away from the templating base and an alternate conformation observed with Arg-247 that breaks a hydrogen bonding contact with the deoxyribose moiety of the O^6-MeG template residue (supplemental Fig. S10). It is plausible that non-Watson-Crick base pairs must simply "settle in" to the Dpo4 active site, consistent with the subtle side-chain rearrangements observed in the little finger domain. Such a passive geometric determinant for polymerase activity is possible because of the flexible nature of the Dpo4 active site. The substrate specificity of polymerases, *e.g.* pol T7$^-$ (48, 49), that place relatively rigid constraints upon base-pairing geometry within the active site and that are intolerant toward any deviation from such conformations may be more inclined to use Watson-Crick-like O^6-MeG:T pairing. Conversely, the substrate specificity of polymerases with a relatively open or flexible active site may be influenced by parameters such as the number and stability of hydrogen bonds between bases.

Four genes that are thought to code for polymerases in the *S. solfataricus* genome, and only one of these (Dpo4) is a confirmed "translesion" polymerase. The *S. solfataricus* replicative polymerase may be as (or more) efficient at bypass of O^6-MeG than Dpo4. If this is the case then insertion of T is likely to be the result if the *S. solfataricus* polymerase follows the trend of other family B polymerases. Consistent with such an idea, human pol δ can readily bypass and extend across from O^6-MeG-modified DNA, and steady-state results suggest an equal propensity to C or T incorporation (47). Human pol δ catalysis is inhibited ~10-fold opposite O^6-MeG. The human Y-family polymerases are also inhibited 10- to 100-fold, with the exception of pol ι which actually inserts T opposite O^6-MeG with greater efficiency than it does C opposite G (47). Another question that remains unanswered is the role of archaeal proliferating cell nuclear antigen-like proteins in Dpo4-catalyzed lesion bypass. It is possible that an interaction between archaeal proliferating cell nuclear antigen and Dpo4 serves to increase the efficiency of bypass observed for several types of lesions, including O^6-MeG.

Another interesting comparison can be made between the results obtained here and Dpo4-catalyzed bypass of 8-oxoG (23). Dpo4 is highly efficient at C insertion opposite 8-oxoG, whereas most if not all other polymerases studied are inhibited (even if they still insert more C than A). In the case of O^6-MeG, Dpo4 catalysis is inhibited as with all other polymerases studied to date. It is the ratio of C to T incorporation that sets Dpo4 apart. The 8-oxoG modification does not perturb the hydrogen bonding face of guanine. Several of the lesions studied to date that interrupt hydrogen bonding patterns tend to strongly inhibit Dpo4 catalysis (*e.g.* O^6-MeG, 1,N^2-etheno-G), pointing toward the relevance of incoming nucleotide stability, as effected by hydrogen bonding interactions, to Dpo4 catalysis. In the case of 8-oxoG, a hydrogen bond between Arg-332 and the O8 atom appears to help Dpo4 "grip" the lesion and stabilize the 8-oxoG:C base pair, which makes Dpo4-catalyzed insertion of dCTP opposite 8-oxoG ~20-fold more efficient than Dpo4-catalyzed insertion of dCTP opposite G (23).

The resulting picture for Dpo4-catalyzed bypass of O^6-MeG is clear. Correct incorporation of C opposite O^6-MeG is the major product, followed by accurate extension of at least 4 bp. The mechanism by which Dpo4 accomplishes correct incorporation is ~14-fold less efficient than what is observed with unmodified DNA but remains more efficient than misincorporation by any measure tested here. Based on kinetic and structural data, proper alignment of the tri-phosphate moiety in the catalytic center is most likely the step at which inhibition occurs. Misalignment appears to result from a shift in the hydrogen-bonding pattern between C and O^6-MeG to form a wobble base pair, as observed in the crystal structure (Fig. 7*A*). Dpo4 does catalyze insertion of T (and to a lesser extent A) but an elevated $K_{m,\mathrm{dTTP}}$ and a disordered T in the crystal structure indicate that the O^6-MeG:T pair is less stable in the Dpo4 active site, which results in an enzyme that favors accurate bypass. These findings provide insight into the factors that directly influence the preference for C over T incorporation when Dpo4 bypasses O^6-MeG.

In conclusion, the major points of interest involve the ascertainment of how O^6-MeG pairs within a biologically relevant context. Previous biophysical studies of the nucleic acids alone have revealed several modes of pairing between C (or T) and O^6-alkylG with the oligonucleotides in isolation (36, 40); however, none of these studies are instructive regarding how the lesion would be handled by proteins. A crystal structure of O^6-alkylguanine-DNA alkyltransferase in complex with O^6-MeG, although informative about DNA repair, fails to provide any information regarding how O^6-MeG pairs because of the elegant "base-flipping" mechanism utilized as a means of damage recognition (31). The structural enzymology presented here provides what appears to be the first structural insight into how O^6-MeG might be processed during replication by a uniquely fitted group of enzymes. Whether the structural results observed with an archaeal member of the Y-family hold true for the eukaryotic enzymes, *e.g.* pol η and pol κ, is a question for further studies.

Acknowledgments—We thank H. Robinson for data collection support at the National Synchrotron Light Source, L. M. Manier for assistance in the Vanderbilt Mass Spectrometry Center, J. Harp for assistance in the Vanderbilt Center for Structural Biology, and H. Zang for initial LC-MS results.

REFERENCES

1. Loveless, A. (1969) *Nature* **223,** 206–207
2. Lawley, P. D. (1984) in *Chemical Carcinogens* (Searle, C. E., ed) 2nd Ed., pp. 325–484, American Chemical Society, Washington, D. C.

3. Wheeler, G. P. (1962) *Cancer Res.* **22,** 651–688
4. Margison, G. P., Santibanez Koref, M. F., and Povey, A. C. (2002) *Mutagenesis* **17,** 483–487
5. Delaney, J. C., and Essigmann, J. M. (2001) *Biochemistry* **40,** 14968–14975
6. Dodson, L. A., Foote, R. S., Mitra, S., and Masker, W. E. (1982) *Proc. Natl. Acad. Sci. U. S. A.* **79,** 7440–7444
7. Dosanjh, M. K., Singer, B., and Essigmann, J. M. (1991) *Biochemistry* **30,** 7027–7033
8. Pegg, A. E., Roberfroid, M., von Bahr, C., Foote, R. S., Mitra, S., Bresil, H., Likhachev, A., and Montesano, R. (1982) *Proc. Natl. Acad. Sci. U. S. A.* **79,** 5162–5165
9. Pegg, A. E., Dolan, M. E., Scicchitano, D., and Morimoto, K. (1985) *Environ. Health Perspect.* **62,** 109–114
10. Yoshioka, K., Yoshioka, Y., and Hsieh, P. (2006) *Mol. Cell* **22,** 501–510
11. Georgiadis, P., Samoli, E., Kaila, S., Katsouyanni, K., and Kyrtopoulos, S. A. (2000) *Cancer Epidemiol. Biomarkers Prev.* **9,** 299–305
12. Kyrtopoulos, S. A., Souliotis, V. L., Valavanis, C., Boussiotis, V. A., and Pangalis, G. A. (1993) *Environ. Health Perspect.* **99,** 143–147
13. Middleton, M. R., Lee, S. M., Arance, A., Wood, M., Thatcher, N., and Margison, G. P. (2000) *Int. J. Cancer* **88,** 469–473
14. Langston, L. D., and O'Donnell, M. (2006) *Mol. Cell* **23,** 155–160
15. Ohmori, H., Friedberg, E. C., Fuchs, R. P., Goodman, M. F., Hanaoka, F., Hinkle, D., Kunkel, T. A., Lawrence, C. W., Livneh, Z., Nohmi, T., Prakash, L., Prakash, S., Todo, T., Walker, G. C., Wang, Z., and Woodgate, R. (2001) *Mol. Cell* **8,** 7–8
16. Plosky, B. S., and Woodgate, R. (2004) *Curr. Opin. Genet. Dev.* **14,** 113–119
17. Prakash, S., Johnson, R. E., and Prakash, L. (2005) *Annu. Rev. Biochem.* **74,** 317–353
18. Ling, H., Boudsocq, F., Plosky, B. S., Woodgate, R., and Yang, W. (2003) *Nature* **424,** 1083–1087
19. Ling, H., Boudsocq, F., Woodgate, R., and Yang, W. (2001) *Cell* **107,** 91–102
20. Ling, H., Sayer, J. M., Plosky, B. S., Yagi, H., Boudsocq, F., Woodgate, R., Jerina, D. M., and Yang, W. (2004) *Proc. Natl. Acad. Sci. U. S. A.* **101,** 2265–2269
21. Ling, H., Boudsocq, F., Woodgate, R., and Yang, W. (2004) *Mol. Cell* **13,** 751–762
22. 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) *J. Biol. Chem.* **280,** 29750–29764
23. Zang, H., Irimia, A., Choi, J. Y., Angel, K. C., Loukachevitch, L. V., Egli, M., and Guengerich, F. P. (2006) *J. Biol. Chem.* **281,** 2358–2372
24. Fiala, K. A., and Suo, Z. (2004) *Biochemistry* **43,** 2116–2125
25. Fiala, K. A., and Suo, Z. (2004) *Biochemistry* **43,** 2106–2115
26. Furge, L. L., and Guengerich, F. P. (1997) *Biochemistry* **36,** 6475–6487
27. Furge, L. L., and Guengerich, F. P. (1998) *Biochemistry* **37,** 3567–3574
28. Christian, N. P., Reilly, J. P., Mokler, V. R., Wincott, F. E., and Ellington, A. D. (2001) *J. Am. Soc. Mass Spectrom.* **12,** 744–753
29. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol.* **276,** 307–326
30. 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) *Acta Crystallogr. D Biol. Crystallogr.* **54,** 905–921
31. Daniels, D. S., Woo, T. T., Luu, K. X., Noll, D. M., Clarke, N. D., Pegg, A. E., and Tainer, J. A. (2004) *Nat. Struct. Mol. Biol.* **11,** 714–720
32. Joyce, C. M., and Benkovic, S. J. (2004) *Biochemistry* **43,** 14317–14324
33. Bakhtina, M., Lee, S., Wang, Y., Dunlap, C., Lamarche, B., and Tsai, M. D. (2005) *Biochemistry* **44,** 5177–5187
34. Parthasarathy, R., and Fridey, S. M. (1986) *Carcinogenesis* **7,** 221–227
35. Yamagata, Y., Kohda, K., and Tomita, K. (1988) *Nucleic Acids Res.* **16,** 9307–9321
36. Patel, D. J., Shapiro, L., Kozlowski, S. A., Gaffney, B. L., and Jones, R. A. (1986) *Biochemistry* **25,** 1027–1036
37. Sriram, M., van der Marel, G. A., Roelen, H. L., van Boom, J. H., and Wang, A. H. (1992) *Biochemistry* **31,** 11823–11834
38. Ginell, S. L., Kuzmich, S., Jones, R. A., and Berman, H. M. (1990) *Biochemistry* **29,** 10461–10465
39. Patel, D. J., Shapiro, L., Kozlowski, S. A., Gaffney, B. L., and Jones, R. A. (1986) *Biochemistry* **25,** 1036–1042
40. Leonard, G. A., Thomson, J., Watson, W. P., and Brown, T. (1990) *Proc. Natl. Acad. Sci. U. S. A.* **87,** 9573–9576
41. Haracska, L., Prakash, S., and Prakash, L. (2000) *Mol. Cell. Biol.* **20,** 8001–8007
42. Singh, J., Su, L., and Snow, E. T. (1996) *J. Biol. Chem.* **271,** 28391–28398
43. Voigt, J. M., and Topal, M. D. (1995) *Carcinogenesis* **16,** 1775–1782
44. Woodside, A. M., and Guengerich, F. P. (2002) *Biochemistry* **41,** 1027–1038
45. Woodside, A. M., and Guengerich, F. P. (2002) *Biochemistry* **41,** 1039–1050
46. Tan, H. B., Swann, P. F., and Chance, E. M. (1994) *Biochemistry* **33,** 5335–5346
47. Choi, J. Y., Chowdury, G., Zang, H., Angel, K. C., Choua, C. V., Peterson, L. A., and Guengerich, F. P. (2006) *J. Biol. Chem.* **281,** 38244–38256
48. Kim, T. W., Brieba, L. G., Ellenberger, T., and Kool, E. T. (2006) *J. Biol. Chem.* **281,** 2289–2295
49. Doublie, S., Tabor, S., Long, A. M., Richardson, C. C., and Ellenberger, T. (1998) *Nature* **391,** 251–258