

# Structure-Function Relationships in Miscoding by *Sulfolobus solfataricus* DNA Polymerase Dpo4

## GUANINE $N^2,N^2$ -DIMETHYL SUBSTITUTION PRODUCES INACTIVE AND MISCODING POLYMERASE COMPLEXES<sup>\*§</sup>

Received for publication, February 3, 2009, and in revised form, April 28, 2009. Published, JBC Papers in Press, May 1, 2009, DOI 10.1074/jbc.M109014274

Huidong Zhang<sup>‡1</sup>, Robert L. Eoff<sup>‡1</sup>, Ivan D. Kozekov<sup>§</sup>, Carmelo J. Rizzo<sup>§</sup>, Martin Egli<sup>‡</sup>, and F. Peter Guengerich<sup>‡2</sup>

From the <sup>‡</sup>Department of Biochemistry and Center in Molecular Toxicology, Vanderbilt University School of Medicine, Nashville, Tennessee 37232-0146 and the <sup>§</sup>Department of Chemistry and Center in Molecular Toxicology, Vanderbilt University, Nashville, Tennessee 37235

Previous work has shown that Y-family DNA polymerases tolerate large DNA adducts, but a substantial decrease in catalytic efficiency and fidelity occurs during bypass of  $N^2,N^2$ -dimethyl ( $\text{Me}_2$ )-substituted guanine ( $N^2,N^2\text{-Me}_2\text{G}$ ), in contrast to a single methyl substitution. Therefore, it is unclear why the addition of two methyl groups is so disruptive. The presence of  $N^2,N^2\text{-Me}_2\text{G}$  lowered the catalytic efficiency of the model enzyme *Sulfolobus solfataricus* Dpo4 16,000-fold. Dpo4 inserted dNTPs almost at random during bypass of  $N^2,N^2\text{-Me}_2\text{G}$ , and much of the enzyme was kinetically trapped by an inactive ternary complex when  $N^2,N^2\text{-Me}_2\text{G}$  was present, as judged by a reduced burst amplitude (5% of total enzyme) and kinetic modeling. One crystal structure of Dpo4 with a primer having a 3'-terminal dideoxycytosine ( $\text{C}_{\text{dd}}$ ) opposite template  $N^2,N^2\text{-Me}_2\text{G}$  in a post-insertion position showed  $\text{C}_{\text{dd}}$  folded back into the minor groove, as a catalytically incompetent complex. A second crystal had two unique orientations for the primer terminal  $\text{C}_{\text{dd}}$  as follows: (i) flipped into the minor groove and (ii) a long pairing with  $N^2,N^2\text{-Me}_2\text{G}$  in which one hydrogen bond exists between the O-2 atom of  $\text{C}_{\text{dd}}$  and the N-1 atom of  $N^2,N^2\text{-Me}_2\text{G}$ , with a second water-mediated hydrogen bond between the N-3 atom of  $\text{C}_{\text{dd}}$  and the O-6 atom of  $N^2,N^2\text{-Me}_2\text{G}$ . A crystal structure of Dpo4 with dTTP opposite template  $N^2,N^2\text{-Me}_2\text{G}$  revealed a wobble orientation. Collectively, these results explain, in a detailed manner, the basis for the reduced efficiency and fidelity of Dpo4-catalyzed bypass of  $N^2,N^2\text{-Me}_2\text{G}$  compared with mono-substituted  $N^2$ -alkyl G adducts.

Cellular DNA is continuously attacked by physical agents and by various endogenous and exogenous chemicals to pro-

duce DNA damage products, including abasic sites (1, 2), deamination products (3, 4), oxidized adducts (e.g. 8-oxoG<sup>3</sup> (5–7)), alkyl lesions (particularly  $O^6$ -alkyl G (8, 9)), UV-induced pyrimidine dimers, and 6–4 photoproducts (10, 11), *bis*-electrophile-induced exocyclic DNA products (12–15), and inter- and intra-strand DNA cross-links (9, 15). Even with all of the cellular DNA repair systems available, some DNA adducts are present in cells and lead to misincorporation, mutation, and blockage when they interact with DNA polymerases, potentially causing cell death, aging, and cancer (9).

The details of how DNA polymerases interact with these modifications in the DNA are important in understanding the biochemistry relevant to these adverse events. DNA polymerases are complex enzymes; they must be able to bind not only DNA but also all four of the canonical dNTPs. The dilemma is how polymerases “sense” which DNA base is in the active site and how the process selects the appropriate base prior to the phosphodiester bond formation step, which “seals” the choice unless exonuclease or repair processes intervene later. The details of this sensing, and the obligate conformational changes necessary to amplify the energy involved in recognition (16), are still controversial (17–20). The situation can be much more mechanistically complex when a modified base is present in the DNA substrate (19).

The N-2 atom of guanine is susceptible to attack by a wide variety of chemicals, including formaldehyde (21), acetaldehyde (22–25), styrene oxide (26),  $\alpha$ -hydroxy-*N*-nitrosopyrrolidine (27), oxidation products of heterocyclic aromatic amines (28), and various polycyclic aromatic hydrocarbons (29, 30), leading to the formation of various  $N^2$ -alkyl G adducts. Pyridyloxobutyl-derived  $N^2$ -alkyl G adducts have been detected in rats after chronic treatment with *N'*-nitrosornicotine (31).  $N^2$ -EtG was detected in granulocyte and lymphocyte DNA and urine of alcoholic patients (32, 33), and misincorporation opposite  $N^2$ -EtG by *Escherichia coli* DNA polymerase I/Klenow fragment (exonuclease<sup>−</sup>) has been reported (22). Crotonaldehyde- and acetaldehyde-derived 1, $N^2$ -propanodeoxyguanosine adducts have also been characterized in human tissue DNA (25). Even 2'-*O*-substitutions that differ by a single methylene

\* This work was supported, in whole or in part, by National Institutes of Health Grants R01 ES010375 (to F. P. G.), F32 CA119776 (to R. L. E.), P01 ES05355 (to I. D. K., C. J. R., and M. E.), and P30 ES000267 (to F. P. G., C. J. R., and M. E.). The atomic coordinates and structure factors (codes 2w9a, 2w9b and 2w9c) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (<http://www.rcsb.org/>).

§ The on-line version of this article (available at <http://www.jbc.org>) contains supplemental Tables S1–S12 and Figs. S1–S6.

<sup>1</sup> Both authors contributed equally to this work.

<sup>2</sup> To whom correspondence should be addressed: Dept. of Biochemistry and Center in Molecular Toxicology, Vanderbilt University School of Medicine, 638 Robinson Research Bldg., 2200 Pierce Ave., Nashville, TN 37232-0146. Tel.: 615-322-2261; Fax: 615-322-3141; E-mail: f.guengerich@vanderbilt.edu.

<sup>3</sup> The abbreviations used are: 8-oxoG, 7,8-dihydro-8-oxodeoxyguanosine; Et, ethyl; G, 2'-deoxyguanosine; LC, liquid chromatography; Me, methyl; MS, mass spectrometry; MS/MS, tandem mass spectrometry;  $N^2$ -NaphG,  $N^2$ -(2-naphthyl)methylG; pol, (DNA) polymerase; PDB, Protein Data Bank.

## DNA Polymerase Dpo4 Structure and Function

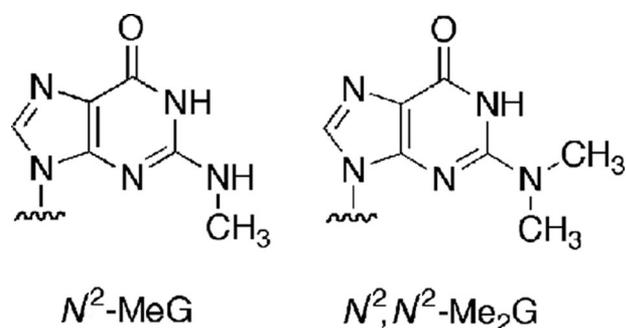


FIGURE 1.  $N^2\text{-MeG}$  and  $N^2,N^2\text{-Me}_2\text{G}$ .

moiety can exhibit very different RNA affinities in terms of modified DNAs (34). In addition to exogenous agents, S-adenosyl-L-methionine-dependent methyltransferases can both mono- and di-methylate the N-2 atom of guanine in RNA (35).  $N^2\text{-MeG}$  and  $N^2,N^2\text{-Me}_2\text{G}$  have been detected in tRNA, rRNA, and small nuclear RNA molecules from prokaryotic, archaeal, and eukaryotic systems (36, 37). The  $N^2,N^2\text{-Me}_2\text{G}$  modification occurs at position 26 in at least 103 eukaryotic tRNAs, with  $N^2\text{-MeG}$  occurring in at least 17 tRNAs (25, 37, 38). It has been postulated that the presence of  $N^2,N^2\text{-Me}_2\text{G}$  at positions 10 and 26 of tRNA helps prevent +1 frameshifting during translation events, but whether this is because  $N^2,N^2\text{-Me}_2\text{G}$  at position 26 stabilizes the tertiary structure of the region near the D-stem and the anti-codon stem or because the modifications influence tRNA positioning in the ribosome is unclear (39).  $N^2$ -Methylation of guanine is important in the context of genomic replication because some retroviruses can use host RNA to prime viral genomic synthesis (40), and  $N^2\text{-MeG}$  has been shown to inhibit avian myeloblastosis virus reverse transcriptase activity on *E. coli* 16 S rRNA (37).

Structure-function relationships with various  $N^2$ -alkyl G lesions have been studied in oligonucleotides in this laboratory with a number of DNA polymerases. Even a Me or Et group is quite inhibitory to the processive DNA polymerases HIV-1 reverse transcriptase and bacteriophage DNA polymerase T7 (41). The polymerization activity of Klenow fragment was only slightly attenuated, but misincorporation was observed (21, 22). With human Y-family DNA polymerases, the rates of dNTP incorporation with the four normal dNTPs are slower, and processivity is lower than with the processive A- and B-family DNA polymerases. However, the presence of alkyl (and aralkyl) moieties at the N-2 atom of G is much less inhibitory to Y-family DNA polymerases (42–46), and with some DNA polymerases even a bulky (6-benzo[*a*]pyrenyl)methyl moiety is tolerated reasonably well (43–45). Dpo4, a prototypic Y-family DNA polymerase from the crenarchaeon *Sulfolobus solfataricus*, was even less sensitive to the effect of increasing bulk than the human Y-family enzymes (46).

The presence of an Et group has only a limited effect in slowing the Y-family DNA polymerases (42–47), and small conformational changes occurring in pol  $\iota$  have been shown to accommodate the lesion (47). However, distributing the same amount of overall bulk into two 1-carbon units, *i.e.*  $N^2,N^2\text{-Me}_2\text{G}$  (Fig. 1), has a dramatic effect in slowing incorporation of dCTP by human Y-family pol  $\eta$  and  $\kappa$  (42, 44) but not (human) pol  $\iota$  or

TABLE 1

### Oligodeoxynucleotides used in this study

G\* indicates G,  $N^2\text{-MeG}$ , or  $N^2,N^2\text{-Me}_2\text{G}$ .

24-mer	5'-GCCTCGAGCCAGCCGAGACGCAGC
25C-mer	5'-GCCTCGAGCCAGCCGAGACGCAGC
25T-mer	5'-GCCTCGAGCCAGCCGAGACGCAGT
36-mer	3'-CGGAGCTCGGTCGGCGTCTCTCTCGGGCT
13C <sub>dd</sub> -mer	5'-GGGGGAAGGATTC <sub>dd</sub>
14C <sub>dd</sub> -mer	5'-GGGGGAAGGATTC <sub>dd</sub>
18G <sup>*</sup> C-mer	3'-CCCCCTCCTAAG( $N^2,N^2\text{-Me}_2\text{G}$ )CACT
18G <sup>*</sup> T-mer	3'-CCCCCTCCTAAG( $N^2,N^2\text{-Me}_2\text{G}$ )TACT

Rev1 (43, 45), the latter two of which use non-“Watson-Crick” pairing in alignments for phosphodiester bond formation (48–50). In the case of Dpo4 and the bulky lesion  $N^2\text{-NaphG}$ , we were able to characterize the crystal structures of two forms and provide evidence for the existence of multiple conformations of the polymerase·oligonucleotide·dNTP complex (46). The dramatic attenuation of dCTP incorporation opposite  $N^2,N^2\text{-Me}_2\text{G}$  compared with  $N^2\text{-MeG}$  with human pol  $\eta$  and  $\kappa$  (44) is striking and was also observed in preliminary experiments with Dpo4. The small size and strong blocking nature of this adduct rendered  $N^2,N^2\text{-Me}_2\text{G}$  useful in structural and kinetic investigations regarding why this simple adduct is so inhibitory. The results allow for some general conclusions that have relevance in the context of DNA polymerases bound to damaged DNA and their equilibria between catalytically active and nonproductive species.

## EXPERIMENTAL PROCEDURES

**Materials**—Unlabeled dNTPs were purchased from Amersham Biosciences; ( $S_p$ )-dCTP $\alpha$ S was obtained from Biolog Life Science Institute (Bremen, Germany), and [ $\gamma\text{-}^{32}\text{P}$ ]ATP (specific activity  $3 \times 10^3$  Ci mmol<sup>-1</sup>) was from PerkinElmer Life Sciences. T4 polynucleotide kinase was purchased from New England Biolabs (Beverly, MA). Dpo4 was expressed and purified as described elsewhere (51). Unmodified 13-mer, 14-mer, 24-mer, 25-mer, and 36-mer (Table 1) were purchased from Midland Certified Reagent Co. (Midland, TX). The 18-mer and 36-mer templates containing  $N^2\text{-MeG}$  or  $N^2,N^2\text{-Me}_2\text{G}$  were synthesized and characterized by capillary gel electrophoresis and matrix-assisted laser desorption ionization/time-of-flight mass spectrometry (43). Extinction coefficients for the oligonucleotides, estimated by the Borer method (52), were as follows: 13-mer,  $\epsilon_{260} = 112 \text{ mM}^{-1} \text{ cm}^{-1}$ ; 14-mer,  $\epsilon_{260} = 122 \text{ mM}^{-1} \text{ cm}^{-1}$ ; 18-mer,  $\epsilon_{260} = 157 \text{ mM}^{-1} \text{ cm}^{-1}$ ; 24-mer,  $\epsilon_{260} = 224 \text{ mM}^{-1} \text{ cm}^{-1}$ ; 25-mer,  $\epsilon_{260} = 232 \text{ mM}^{-1} \text{ cm}^{-1}$ ; and 36-mer,  $\epsilon_{260} = 310 \text{ mM}^{-1} \text{ cm}^{-1}$  (43).

**Preparation of Labeled Primer and Duplex**—The 5'-end of the 24-mer primer was labeled with [ $\gamma\text{-}^{32}\text{P}$ ]ATP using T4 polynucleotide kinase at 37 °C for 30 min. After removal of excess [ $\gamma\text{-}^{32}\text{P}$ ]ATP using a Bio-Spin 6 column (Bio-Rad), the labeled primer and unmodified (G) or modified ( $N^2\text{-MeG}$  or  $N^2,N^2\text{-Me}_2\text{G}$ ) template (molar ratio 1:1) were heated at 95 °C for 5 min and then slowly cooled to room temperature to form the 24-mer/36-mer duplexes, which were used for all steady-state and pre-steady-state kinetic experiments.

**Reaction Conditions for Assays and Product Analysis Methods**—Standard DNA polymerization reactions with Dpo4 were carried out in 50 mM Tris-HCl buffer (pH 7.5 at 25 °C)

containing 50 mM NaCl, 5 mM dithiothreitol, 100  $\mu\text{g}$  of bovine serum albumin  $\text{ml}^{-1}$ , and 5% (v/v) glycerol at 37 °C (42–45). All reactions were initiated by mixing dNTP/MgCl<sub>2</sub> (final MgCl<sub>2</sub> concentration of 5 mM) solution to preincubated Dpo4/DNA mixtures. After reaction, 5- $\mu\text{l}$  aliquots were quenched with EDTA/formamide solution (50  $\mu\text{l}$  of 20 mM EDTA in 95% formamide (v/v) with 0.5% bromophenol blue (w/v) and 0.05% xylene cyanol (w/v)). Products were resolved using 20% polyacrylamide (w/v) denaturing gel electrophoresis (containing 8 M urea) and visualized and quantitated by phosphorimaging analysis using a Bio-Rad Molecular Imager FX instrument and Quantity One software.

**Primer Extension Assays with All Four dNTPs**—The primer was extended by adding four dNTPs (100  $\mu\text{M}$  each) and 5 mM MgCl<sub>2</sub> to the incubated mixture of 100 nM DNA and Dpo4 (0, 0.5, 2, or 10 nM) under standard assay conditions. After 20 min, reactions were quenched and processed as described above.

**Steady-state Kinetic Assays**—Steady-state single-base incorporation experiments were performed by adding a single dNTP at varying concentrations (12 points) and MgCl<sub>2</sub> to the Dpo4·DNA complexes, incubated in the reaction buffer. The molar ratio of Dpo4 to DNA was <0.1, and primer conversion to product was kept <20% by adjusting the polymerase concentration and incorporation time (53). Reactions were quenched and products were analyzed and quantitated; graphs of the incorporation rates versus dNTP concentration were fit to a hyperbolic equation to yield  $k_{\text{cat}}$  and  $K_m$  values using nonlinear regression in GraphPad Prism version 3.0 (GraphPad, San Diego).

**Pre-steady-state Kinetic Analysis**—Rapid quench experiments were performed in a model RQF-3 KinTek Quench Flow apparatus (KinTek, Austin, TX) with 50 mM Tris-HCl (pH 7.4) aqueous solution in the drive syringes. Reactions with excess DNA (including phosphorothioate analysis) were initiated by rapid mixing of 70 nM Dpo4, 120 nM DNA mixtures (12.5  $\mu\text{l}$ ) with 1 mM dNTP (or (*S*<sub>p</sub>)-dCTP $\alpha$ S), 5 mM MgCl<sub>2</sub> (10.9  $\mu\text{l}$ ). Reactions with excess Dpo4 were initiated by mixing 200 nM Dpo4, 100 nM DNA mixtures (12.5  $\mu\text{l}$ ) with dCTP (varying concentrations)·5 mM MgCl<sub>2</sub> complex (10.9  $\mu\text{l}$ ). After the reactions were quenched by the addition of 0.6 M EDTA from the central syringe line after varying times, the products were analyzed and quantitated. The reactions with excess DNA or with excess Dpo4 were fit to Equations 1 or 2, respectively (where  $t$  is time), to obtain a burst amplitude  $A$ , burst rate  $k_p$ , and steady-state velocity  $k_{\text{ss}}$ ,

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

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

Plots of  $k_p$  versus dCTP concentration were fit to hyperbolic Equation 3,

$$k_p = k_{\text{pol}}[\text{dCTP}]/([\text{dCTP}] + K_{d,\text{dCTP}}) \quad (\text{Eq. 3})$$

to estimate  $k_{\text{pol}}$  and  $K_{d,\text{dCTP}}$ , where  $k_{\text{pol}}$  is the maximal rate of nucleotide incorporation, and  $K_{d,\text{dCTP}}$  is an equilibrium dissociation constant for dCTP in the active form of the polymerase (54, 55). All nonlinear regression analysis used GraphPad Prism version 3.0.

**LC-MS/MS Analysis of Primer Extension Products (51, 56)**—The primer was extended by incubating Dpo4 (5  $\mu\text{M}$ ), unlabeled DNA (10  $\mu\text{M}$ ), a mixture of all four dNTPs (1 mM each), and MgCl<sub>2</sub> (5 mM) in Tris-HCl buffer (pH 7.5, final volume 100  $\mu\text{l}$ ) at 37 °C for 4 h. Each reaction was terminated by extraction of the remaining dNTPs using a Bio-Spin 6 chromatography column, and concentrated Tris-HCl, dithiothreitol, and EDTA were added to restore the concentrations to 50, 5, and 1 mM, respectively. *E. coli* uracil DNA glycosylase (20 units, Sigma) was then added; the solution was incubated at 37 °C for 6 h to hydrolyze the uracil residues on the extended primer and then heated at 95 °C for 1 h in the presence of 0.25 M piperidine, followed by removal of the solvent by *in vacuo* centrifugation. The dried sample was resuspended in 100  $\mu\text{l}$  of H<sub>2</sub>O for mass spectrometry analysis.

LC-MS/MS analysis was performed using a Waters Acquity UPLC system (Waters, Milford, MA) connected to a Finnigan LTQ mass spectrometer (Thermo Fisher Scientific, Waltham, MA) operating in the electrospray ionization negative ion mode. An Acquity UPLC BEH octadecylsilane (C<sub>18</sub>) column (1.7  $\mu\text{m}$ , 1.0  $\times$  100 mm) was used with the following LC conditions (all at 50 °C) with Buffer A (10 mM NH<sub>4</sub>CH<sub>3</sub>CO<sub>2</sub> plus 2% CH<sub>3</sub>CN (v/v)) and Buffer B (10 mM NH<sub>4</sub>CH<sub>3</sub>CO<sub>2</sub> plus 95% CH<sub>3</sub>CN (v/v)). The conditions used were similar to those reported previously (46, 57). The calculations of the collision-induced dissociation fragmentations of oligonucleotide sequences were done using a program linked to the Mass Spectrometry Group (Medicinal Chemistry) of the University of Utah.

**Crystallization and Collection of X-ray Diffraction Data**—Attempts to crystallize Dpo4 in the presence of *N*<sup>2</sup>,*N*<sup>2</sup>-Me<sub>2</sub>G-substituted primer-template DNA and an incoming dCTP were unsuccessful. However, we did obtain diffraction quality crystals for three complexes. For correct pairing opposite *N*<sup>2</sup>,*N*<sup>2</sup>-Me<sub>2</sub>G (Dpo4 + *N*<sup>2</sup>,*N*<sup>2</sup>-Me<sub>2</sub>G-<sup>14</sup>C<sub>dd</sub> + dGTP, post-insertion complex, *DMG-1*), the DNA duplex contained the 18G\**C*-mer template and the <sup>14</sup>C<sub>dd</sub>-mer primer, with C<sub>dd</sub> being a dideoxy residue (Table 1). The Dpo4/DNA mixture (1:1.2 molar ratio, in 20 mM Tris-HCl buffer (pH 7.5, 25 °C) containing 60 mM NaCl, 4% glycerol (v/v), and 5 mM  $\beta$ -mercaptoethanol) was placed on ice for 1 h prior to incubation with 5 mM MgCl<sub>2</sub> and 1 mM dGTP. The final Dpo4 concentration was 8–9 mg ml<sup>-1</sup> for all three setups. The crystals were grown using a sitting drop, vapor-diffusion method with the reservoir solution containing 10–15% polyethylene glycol 3350 (w/v), 30 mM NaCl, 100 mM MgCl<sub>2</sub>, and 3% glycerol (v/v). Droplets consisted of a 1:0.5 or 1:1.5 (v/v) mixture of the Dpo4·DNA·Mg<sup>2+</sup>·dGTP complex and the reservoir solutions and were equilibrated against the reservoir solutions.

A complex of Dpo4 bound to *N*<sup>2</sup>,*N*<sup>2</sup>-Me<sub>2</sub>G-substituted primer-template DNA (Dpo4-*N*<sup>2</sup>,*N*<sup>2</sup>-Me<sub>2</sub>G-<sup>14</sup>C<sub>dd</sub>, binary complex, *DMG-2*) was obtained in the presence of dGTP (1 mM). For correct pairing opposite *N*<sup>2</sup>,*N*<sup>2</sup>-Me<sub>2</sub>G (Dpo4 + *N*<sup>2</sup>,*N*<sup>2</sup>-Me<sub>2</sub>G-<sup>14</sup>C<sub>dd</sub>, *DMG-2*), the DNA duplex contained the 18G\*T-mer template and the <sup>14</sup>C<sub>dd</sub>-mer primer, with C<sub>dd</sub> being a dideoxy residue (Table 1). Crystals were grown using the same conditions described for *DMG-1*.

## DNA Polymerase Dpo4 Structure and Function

For mispairing opposite  $N^2,N^2$ -Me<sub>2</sub>G (Dpo4 +  $N^2,N^2$ -Me<sub>2</sub>G·13C<sub>dd</sub> + dTTP, insertion complex, *DMG-3*), the DNA was annealed by mixing the 18G\**C*-mer and the 13C<sub>dd</sub>-mer primer (Table 1), with C<sub>dd</sub> being a dideoxy residue. The Dpo4·DNA complex (1:1.2 molar ratio, in 20 mM Tris-HCl buffer (pH 7.5, 25 °C) containing 60 mM NaCl, 4% glycerol (v/v), and 5 mM β-mercaptoethanol) was placed on ice for 1 h prior to incubation with 5 mM MgCl<sub>2</sub> and 1 mM dGTP. Crystals were grown using the same method as *DMG-1* with the reservoir solution containing 20 mM Tris-HCl (pH 7.5 at 25 °C), 15% polyethylene glycol 3350 (w/v), 100 mM calcium diacetate, and 3% glycerol (v/v). Droplets consisted of a 1:2 (v/v) mixture of the Dpo4·DNA·Ca<sup>2+</sup>·dTTP complex and equilibrated against the reservoir solutions. All 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**—X-ray diffraction data for the Dpo4- $N^2,N^2$ -Me<sub>2</sub>G·<sup>14</sup>C<sub>dd</sub>-dGTP (*DMG-1*), Dpo4- $N^2,N^2$ -Me<sub>2</sub>G·<sup>14</sup>C<sub>dd</sub> (*DMG-2*), and Dpo4- $N^2,N^2$ -Me<sub>2</sub>G·13C<sub>dd</sub>-dTTP (*DMG-3*) complex crystals were collected at the Advanced Photon Source (Argonne National Laboratory, Argonne, IL) on the ID-21 beamline/station (Life Sciences Collaborative Access Team). The data sets were recorded from cryoprotected crystals using a wavelength of 0.98 Å at 110 K. The crystals diffracted to between 2.3 and 2.9 Å (Table 5). Individual frames were indexed and scaled with the program XDS (58). The *DMG-1* crystal belongs to space group P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub>, and the *DMG-2* and *DMG-3* crystals belong to space group P2<sub>1</sub>.

**Structure Determination and Refinement**—A refined wild-type Dpo4 complex (PDB accession code 2jef) (57) minus solvent molecules, template residue 5, metal ions, and dGTP was used as the initial search model for *DMG-1*. Another refined wild-type Dpo4 complex (PDB accession code 2asd) (59) minus solvent molecules, template residue 5, metal ions, and dGTP was used as the initial search model for *DMG-2* and *DMG-3*. Molecular replacement was performed using MOLREP as a part of the CCP4 program suite (60).

Manual model rebuilding was done with the program Turbo-Frodo (80). The maps were computed using the σ<sub>A</sub>-modified coefficients (61). Clear positive density for three Mg<sup>2+</sup> ions and the dGTP was observed in the initial difference Fourier electron density maps of the *DMG-1* complex. The *DMG-2* and *DMG-3* structures each indexed to the P2<sub>1</sub> space group and possess two molecules in the asymmetric unit. In the case of *DMG-2*, one molecule displayed better density maps near the DNA residues in the active site of the enzyme. Two Mg<sup>2+</sup> ions were positioned into unambiguous, positive difference Fourier electron density at the active site of both monomers per asymmetric unit. In each monomer there was only one Mg<sup>2+</sup> ion in the active site of the polymerase. The second Mg<sup>2+</sup> ion was positioned in a region of the thumb domain that resides near the -1 position of the primer DNA. It was clear from positive difference Fourier electron density that the terminal C<sub>dd</sub> existed in two alternate conformations in both molecules of the *DMG-2* complex. No conclusive density was observed for the incoming dGTP molecule. The *DMG-2* complex represents a pre-translocation state

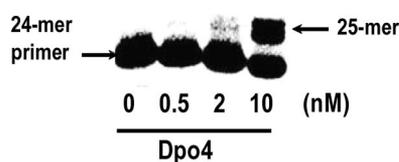


FIGURE 2. Extension of <sup>32</sup>P-labeled primer opposite  $N^2,N^2$ -Me<sub>2</sub>G by Dpo4 in the presence of all four dNTPs. A 24-mer primer/36-mer  $N^2,N^2$ -Me<sub>2</sub>G template complex (100 nM) was extended using Dpo4 (0, 0.5, 2, and 10 nM).

because there is no incoming dNTP and the primer terminus is still found to occupy much of the dNTP binding cleft.

In the *DMG-3* structure, the dTTP molecule and three Mg<sup>2+</sup> ions were positioned into unambiguous, positive difference Fourier electron density at the active site of both monomers per asymmetric unit. Two Mg<sup>2+</sup> ions were positioned in the active site of the polymerase, and the third Mg<sup>2+</sup> ion was positioned in a region of the thumb domain near the phosphate backbone of the primer strand. The CNS package (62) was used for the refinement of the models by performing simulated annealing, gradient minimization, and refinement of individual isotropic temperature, and occupancy factors. The statistics of the refined models for all structures are summarized (Table 5). The crystallographic figures were prepared using PyMol (81).

**Kinetic Simulations**—Fits were developed utilizing the program DynaFit (BioKin, Watertown, MA) (63), run on an iMac computer using a Macintosh OS 8.5.1 operating system (Apple Computer, Cupertino, CA). Kinetic simulations were performed using a minimal mechanism with an additional step involving reversible formation of a nonproductive ternary complex (64, 65). The experimental data came from the chemical quench experiments done with excess DNA. Initial parameters were chosen based upon previous work (66).

## RESULTS

**Primer Extension with All Four dNTPs**—In preliminary assays with Dpo4 and all four dNTPs, the primer was fully extended beyond template  $N^2$ -MeG or  $N^2$ -EtG (46). Only one base was incorporated opposite the lesion  $N^2,N^2$ -Me<sub>2</sub>G under the conditions tested (Fig. 2). Under identical extension conditions,  $N^2$ -alkyl G adducts with bulk up to that of a (2-naphthyl)methyl moiety can be readily and accurately extended by Dpo4 (46). Similar to Dpo4, bacteriophage DNA polymerase T7 (exonuclease<sup>-</sup>)-catalyzed extension past  $N^2,N^2$ -Me<sub>2</sub>G was blocked, and only one base was incorporated opposite the lesion site (results not shown) (oligonucleotides containing unmodified G or  $N^2$ -MeG can be readily extended to full length by this polymerase (41)).

**Steady-state dNTP Incorporation Opposite  $N^2,N^2$ -Me<sub>2</sub>G**—The steady-state efficiency ( $k_{cat}/K_m$ ) for dCTP incorporation opposite  $N^2$ -MeG was decreased 3-fold compared with opposite unmodified G (Table 2), but the efficiency for incorporation of dCTP opposite  $N^2,N^2$ -Me<sub>2</sub>G was drastically decreased (16,000-fold compared with G). Misincorporation efficiencies (dATP, dGTP, or dTTP opposite  $N^2,N^2$ -Me<sub>2</sub>G) were similar to each other and to those seen opposite G or  $N^2$ -MeG. The misincorporation frequencies for incorporation opposite  $N^2,N^2$ -Me<sub>2</sub>G were high, in the range of 0.36 to 2.3, because of the low rate of dCTP incorporation, and indicate that Dpo4 has essen-

TABLE 2

Steady-state kinetic parameters for one-base incorporation opposite unmodified G, N<sup>2</sup>-MeG, and N<sup>2</sup>,N<sup>2</sup>-Me<sub>2</sub>G by Dpo4

The extent of conversion of primer to the product was kept &lt;20% by adjustment of the enzyme concentration and reaction time.

Template	dNTP	$k_{\text{cat}}$ $s^{-1}, \times 10^{-3}$	$K_m$ $\mu\text{M}$	$k_{\text{cat}}/K_m$ $\mu\text{M}^{-1} s^{-1}$	Decrease relative to G:C	Misinsertion ratio <sup>a</sup>
G	C	960 ± 45	1.2 ± 0.3	0.80		
	A	5.7 ± 0.2	490 ± 60	$1.2 \times 10^{-5}$		$1.5 \times 10^{-5}$
	G	6.3 ± 0.1	190 ± 18	$3.3 \times 10^{-5}$		$4.1 \times 10^{-5}$
	T	34 ± 2	780 ± 140	$4.4 \times 10^{-5}$		$5.4 \times 10^{-5}$
N <sup>2</sup> -MeG <sup>b</sup>	C	500 ± 21	1.9 ± 0.5	0.26	3-Fold	
	A	7.6 ± 0.3	650 ± 73	$1.2 \times 10^{-5}$		$4.5 \times 10^{-5}$
	G	23 ± 2	380 ± 77	$6.0 \times 10^{-5}$		$2.3 \times 10^{-4}$
	T	40 ± 5	320 ± 93	$1.3 \times 10^{-4}$		$4.8 \times 10^{-5}$
N <sup>2</sup> ,N <sup>2</sup> -Me <sub>2</sub> G	C	3.9 ± 0.1	78 ± 8	$5.0 \times 10^{-5}$	16,000-Fold	
	A	2.6 ± 0.1	140 ± 19	$1.9 \times 10^{-5}$		0.37
	G	4.2 ± 0.2	130 ± 19	$3.2 \times 10^{-5}$		0.64
	T	11 ± 0.3	93 ± 10	$1.2 \times 10^{-4}$		2.4

<sup>a</sup> The misinsertion ratio =  $(k_{\text{cat}}/K_m)_{\text{incorrect}}/(k_{\text{cat}}/K_m)_{\text{correct}}$ .<sup>b</sup> See 46.

TABLE 3

Steady-state kinetic parameters for next base extension from G; N<sup>2</sup>-MeG; or N<sup>2</sup>,N<sup>2</sup>-Me<sub>2</sub>G):C or :T termini by Dpo4

The extent of conversion of primer to the product was kept &lt;20% by adjustment of the enzyme concentration and reaction time.

Template	Primer base	$k_{\text{cat}}$ $s^{-1}, \times 10^{-3}$	$K_m$ $\mu\text{M}$	$k_{\text{cat}}/K_m$ $\mu\text{M}^{-1} s^{-1}$	Mis-extension ratio <sup>a</sup>
G	C	530 ± 30	23 ± 3	0.023	
	T	17 ± 1	175 ± 20	$9.7 \times 10^{-5}$	$4.2 \times 10^{-3}$
N <sup>2</sup> -MeG <sup>b</sup>	C	190 ± 10	185 ± 30	$1.0 \times 10^{-3}$	
	T	6.9 ± 0.2	73 ± 9	$9.5 \times 10^{-5}$	0.094
N <sup>2</sup> ,N <sup>2</sup> -Me <sub>2</sub> G	C	4.2 ± 0.1	3.7 ± 0.4	$1.1 \times 10^{-3}$	
	T	1.5 ± 0.1	86 ± 16	$1.7 \times 10^{-5}$	0.015

<sup>a</sup> "Mis-extension ratio" for N<sup>2</sup>,N<sup>2</sup>-Me<sub>2</sub>G is defined as  $(k_{\text{cat}}/K_m)_{\text{mismatch}}/(k_{\text{cat}}/K_m)_{\text{matched pair}}$ .<sup>b</sup> See Ref. 46.

tially no ability to discriminate between incoming dNTPs when N<sup>2</sup>,N<sup>2</sup>-Me<sub>2</sub>G is present.

**Steady-state Next-base Extension Following C or T Paired with G, N<sup>2</sup>-MeG, or N<sup>2</sup>,N<sup>2</sup>-Me<sub>2</sub>G**—Steady-state extension from N<sup>2</sup>,N<sup>2</sup>-Me<sub>2</sub>G pairing with C or T (incorporation of dGTP opposite the neighboring 5'-template base, C) was performed, and the results were compared with those obtained for G and N<sup>2</sup>-MeG (Table 3). Only extension of N<sup>2</sup>,N<sup>2</sup>-Me<sub>2</sub>G:T mismatches was studied because the three misincorporations, A, G, and T, showed similar efficiencies (Table 2). Extension efficiencies ( $k_{\text{cat}}/K_m$ ) for next base extension were decreased ~20-fold for N<sup>2</sup>-MeG or N<sup>2</sup>,N<sup>2</sup>-Me<sub>2</sub>G compared with G. The "mis-extension ratio" for N<sup>2</sup>,N<sup>2</sup>-Me<sub>2</sub>G, defined as  $(k_{\text{cat}}/K_m)_{\text{mismatch}}/(k_{\text{cat}}/K_m)_{\text{matched pair}}$  (53, 67) was increased by 20-fold compared with G, suggesting that the miscoding effect of the N<sup>2</sup>,N<sup>2</sup>-Me<sub>2</sub>G adduct is largely diminished at the extension step.

**Sequence Analysis of Primer Incorporation/Extension Products**—LC-MS/MS was utilized for analysis of the nucleotide sequences of the extension products, with higher concentrations of Dpo4 (5 μM), oligonucleotide substrate (10 μM), and four dNTPs (1 mM each) and longer reaction time (4 h) used to extend the primer beyond N<sup>2</sup>,N<sup>2</sup>-Me<sub>2</sub>G. Initial LC-MS analysis showed a complex mixture eluting at  $t_R$  2.9 min for full-length extension (past N<sup>2</sup>,N<sup>2</sup>-Me<sub>2</sub>G) products formed by Dpo4 (supplemental Fig. S1A), with nine main peaks at  $m/z$  1099.2, 1102.9, 1026.8, 1105.2, 1109.2, 1177.7, 1181.3, 1183.5, and 1175.3 (−4 charged ions) (supplemental Figs. S1–S4) (plus peaks at  $m/z$  884.1, 941.9, and 1023.0 with −5 charge, and some higher  $m/z$  species corresponding to the −3 charged species derived from the same products). LC-MS/MS analyses of full-

TABLE 4

Products of extension of N<sup>2</sup>,N<sup>2</sup>-Me<sub>2</sub>G template-primer complexes by Dpo4Reaction conditions are as follows: 5 μM Dpo4; 10 μM 24/36 DNA primer-template complex; 1 mM each dNTP; 5 mM MgCl<sub>2</sub>; 4-h incubation time, charge: −4; peak area calculated from  $m/z$  limit ± 0.2 atomic mass unit.

	5'-GCCUCGAGUCAGCCGUAGACGUAG 3'-CGGAGCTCGGTCCGCG TCTGCG TCG CTC CTG CGG CT	%	Total
G	AGC GAG GAC GCC GA	17	100%
	AGC GAG GAC GCC GAC	24	C
	AGC GAG GAC GCC GAA	59	
N <sup>2</sup> ,N <sup>2</sup> -Me <sub>2</sub> G	AGC GAG GAC GCC GA	19	24%
	AGC GAG GAC GCC GAT	2	C
	AGC GAG GAC GCC GAA	3	
	AGA GAG GAC GCC G	5	47%
	AGA GAG GAC GCC GA	24	A
	AGA GAG GAC GCC GAA	9	
	AGA GAG GAC GCC GAC	9	
	AGT GAG GAC GCC GA	10	25%
	AGT GAG GAC GCC GAA	11	T
	AGT GAG GAC GCC GAC	4	
	AGG GAG GAC GCC GA	4	4% G

length extension products of N<sup>2</sup>,N<sup>2</sup>-Me<sub>2</sub>G and unmodified G are shown in supplemental Figs. S1–S4, and assignments of collision-induced dissociation ions for products are compiled in supplemental Tables S1–S11.

The products formed in full-length extension beyond G and N<sup>2</sup>,N<sup>2</sup>-Me<sub>2</sub>G by Dpo4 were quantified using previously described approaches (Table 4). In the case of G, all of the detected products correspond to correct incorporation and accurate extension, with some blunt-end additions of A or C. The extension products formed with N<sup>2</sup>,N<sup>2</sup>-Me<sub>2</sub>G included

TABLE 5

Crystal data and refinement parameters for the complexes containing Dpo4 and  $N^2,N^2$ -Me<sub>2</sub>G

Parameter	DMG-1	DMG-2	DMG-3
X-ray source	APS (LS-CAT)	APS (LS-CAT)	APS (LS-CAT)
Beam line	ID-21	ID-21	ID-21
Detector	MARCCD	MARCCD	MARCCD
Wavelength	0.98 Å	0.98 Å	0.98 Å
Temperature	110 K	110 K	110 K
No. of crystals	1	1	1
Space group	P2 <sub>1</sub> ,2 <sub>1</sub> ,2	P2 <sub>1</sub>	P2 <sub>1</sub>
Unit cell ( <i>a</i> , <i>b</i> , <i>c</i> )	94.67, 103.76, 52.75 Å	51.79, 101.78, 97.27 Å	51.71, 101.62, 96.75 Å
Resolution range	19.94 to 2.55 Å	28.75 to 2.28 Å	28.33 to 2.87 Å
Highest resolution shell <sup>a</sup>	(2.69 to 2.55)	(2.42 to 2.28)	(3.05 to 2.87)
No. of measurements	111,598 (10,905)	173,773 (44,372)	71,564 (21,646)
No. of unique reflections	16,285 (1,767)	45,625 (6,507)	20,782 (2,892)
Redundancy	6.8 (6.2)	3.9 (6.8)	3.4 (7.5)
Completeness	98.2% (97.4%)	96.4% (89.7%)	97.3% (79.7%)
$R_{\text{merge}}^b$	6.0 (20.5)	10.1 (26.4)	8.6 (31.0)
Signal to noise ( $\langle I/\sigma I \rangle$ )	23.3 (5.3)	10.3 (3.8)	12.3 (4.2)
Solvent content	53.74%	53.21%	52.83%
<b>Model composition (asymmetric unit)</b>			
No. of amino acid residues	341	683	682
No. of water molecules	65	207	112
No. of Mg <sup>2+</sup> ions	4	4	6
No. of template nucleotides	16	32	32
No. of primer nucleotides	14	28	26
No. of dGTPs	1	0	0
No. of dTTPs	0	0	2
$R_{\text{free}}^c$	21.3%	23.9%	25.2%
$R_{\text{free}}^d$	24.6%	25.1%	28.8%
<b>Estimated coordinate error (Å)</b>			
From Luzatti plot	0.33	0.34	0.43
From Luzatti plot ( <i>c-v</i> <sup>e</sup> )	0.40	0.36	0.50
From $\sigma A$ plot	0.35	0.40	0.38
From $\sigma A$ plot ( <i>c-v</i> <sup>e</sup> )	0.40	0.43	0.41
<b>Temperature factors</b>			
From Wilson plot	42.4 Å <sup>2</sup>	41.0 Å <sup>2</sup>	37.4 Å <sup>2</sup>
Mean isotropic	42.7 Å <sup>2</sup>	36.3 Å <sup>2</sup>	51.9 Å <sup>2</sup>
Root mean square deviation in temperature factors			
Bonded main chain atoms	1.23 Å <sup>2</sup>	1.40 Å <sup>2</sup>	1.14 Å <sup>2</sup>
Bonded side chain atoms	1.85 Å <sup>2</sup>	2.75 Å <sup>2</sup>	1.59 Å <sup>2</sup>
Root mean square standard deviation from ideal values			
Bond lengths	0.009 Å	0.034 Å	0.010 Å
Bond angles	1.4°	3.2°	1.6°
Dihedral angles	21.9°	25.5°	24.1°
Improper angles	1.83°	2.39°	1.27°

<sup>a</sup> Values in parentheses correspond to the highest resolution shells.<sup>b</sup>  $R_{\text{merge}} = \sum_{hkl} \sum_j |I_{hklj} - \langle I_{hkl} \rangle| / \sum_{hkl} \sum_j I_{hklj}$ , where the outer sum (*hkl*) is taken over the unique reflections.<sup>c</sup>  $R_{\text{free}} = \sum_{hkl} |F_{o,hkl} - k|F_{c,hkl}| / \sum_{hkl} |F_{o,hkl}|$ , where  $|F_{o,hkl}|$  and  $|F_{c,hkl}|$  are the observed and calculated structure factor amplitudes, respectively.<sup>d</sup>  $R_{\text{free}} = \sum_{hkl} |F_{o,hkl} - k|F_{c,hkl}| / \sum_{hkl} |F_{o,hkl}|$ , where  $|F_{o,hkl}|$  and  $|F_{c,hkl}|$  are the observed and calculated structure factor amplitudes, respectively, for the set of reflections (5% of the total) omitted from the refinement process.<sup>e</sup> *c-v* indicates cross-validation.

47% A, 24% C, 25% T, and 4% G inserted opposite  $N^2,N^2$ -Me<sub>2</sub>G. Extension beyond  $N^2,N^2$ -Me<sub>2</sub>G was found to be error-free, and no frameshift products were detected.

**Crystal Structures of  $N^2,N^2$ -Me<sub>2</sub>G-<sup>14</sup>C and  $N^2,N^2$ -Me<sub>2</sub>G-dTTP Complexes**—Three x-ray structures were solved for Dpo4 in complex with DNA duplex containing  $N^2,N^2$ -Me<sub>2</sub>G-modified template (Table 5 and Figs. 3–5). All three of the refined structures containing  $N^2,N^2$ -Me<sub>2</sub>G possess a distinct arrangement of the DNA in the active site, and the  $N^2,N^2$ -Me<sub>2</sub>G moiety exhibits well defined electron density (Fig. 3). *DMG-1*, *DMG-2*, and *DMG-3* diffracted to 2.6, 2.3, and 2.9 Å, with reasonable  $R_{\text{merge}}$  values of 6.0, 10.1, and 8.6%, respectively (Table 5). The presence of  $N^2,N^2$ -Me<sub>2</sub>G group did not disrupt the DNA duplex in any of the structures. The relatively unperturbed duplex is different from crystal structures containing the  $N^2$ -NaphG adduct, in which there was noticeable distortion of the helix several base pairs away from the actual site of damage in the presence of  $N^2$ -NaphG (46).

*DMG-1* is a post-insertion complex with an incoming dGTP paired with 5'-neighboring base C near to  $N^2,N^2$ -Me<sub>2</sub>G. Similar to two previously reported structures from our group (46, 57), the primer terminus is flipped out of the active site into a cleft between the protein and the growing minor groove (Fig. 3C). The complex cannot be productive because of the position of the C<sub>dd</sub> primer terminus and most likely represents at least one of the nonproductive entities described in the kinetic modeling section (see below). The modified  $N^2,N^2$ -Me<sub>2</sub>G template residue is positioned in the *anti* orientation and is well stacked between neighbor bases, as is the case for *DMG-2* and *DMG-3* (Figs. 3–5). The incoming dGTP was paired opposite the 5'-C next to the  $N^2,N^2$ -Me<sub>2</sub>G in a standard Watson-Crick mode. The distances for the formed three hydrogen bonds were 2.8, 2.9, and 3.2 Å between the O-4, N-3, and O-2 atoms of the template C and the O-6 atom, the N-1 atom, and 2-position exocyclic amino group of the incoming dGTP, respectively. Two Mg<sup>2+</sup> ions (Fig. 3B, labeled A and B) are coordinated in the

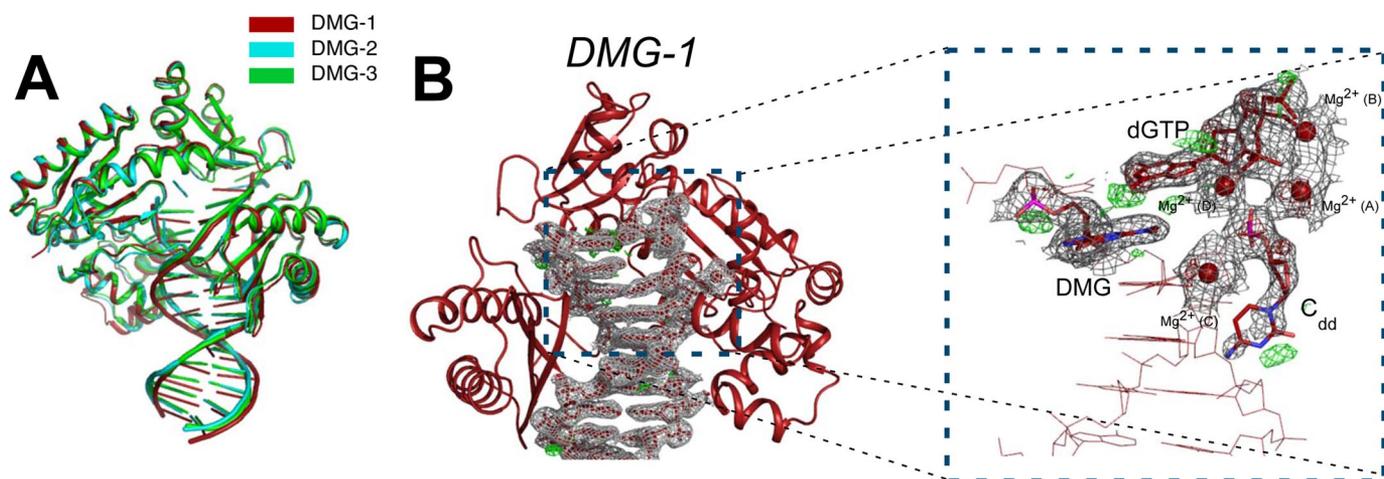


FIGURE 3. **Crystal structures of Dpo4 bound to  $N^2,N^2$ -Me<sub>2</sub>G-modified DNA.** A, superimpositions of DMG-1 (red), DMG-2 (molecule A, cyan), and DMG-3 (molecule A, green) reveal overall similarity in Dpo4 structure. B, representative electron density near the active site of Dpo4 in the DMG-1 structure. The  $3F_o - 2F_c$  map (gray mesh) is shown contoured to the  $1\sigma$  level. The  $F_o - F_c$  difference maps are shown contoured to  $3\sigma$  and  $-3\sigma$  for positive (red mesh) and negative (green mesh) density, respectively. The terminal C<sub>dd</sub> residue is flipped out of base-stacking orientation, but the incoming dGTP forms a Watson-Crick pair with the cytosine to the 5'-side of  $N^2,N^2$ -Me<sub>2</sub>G (DMG).

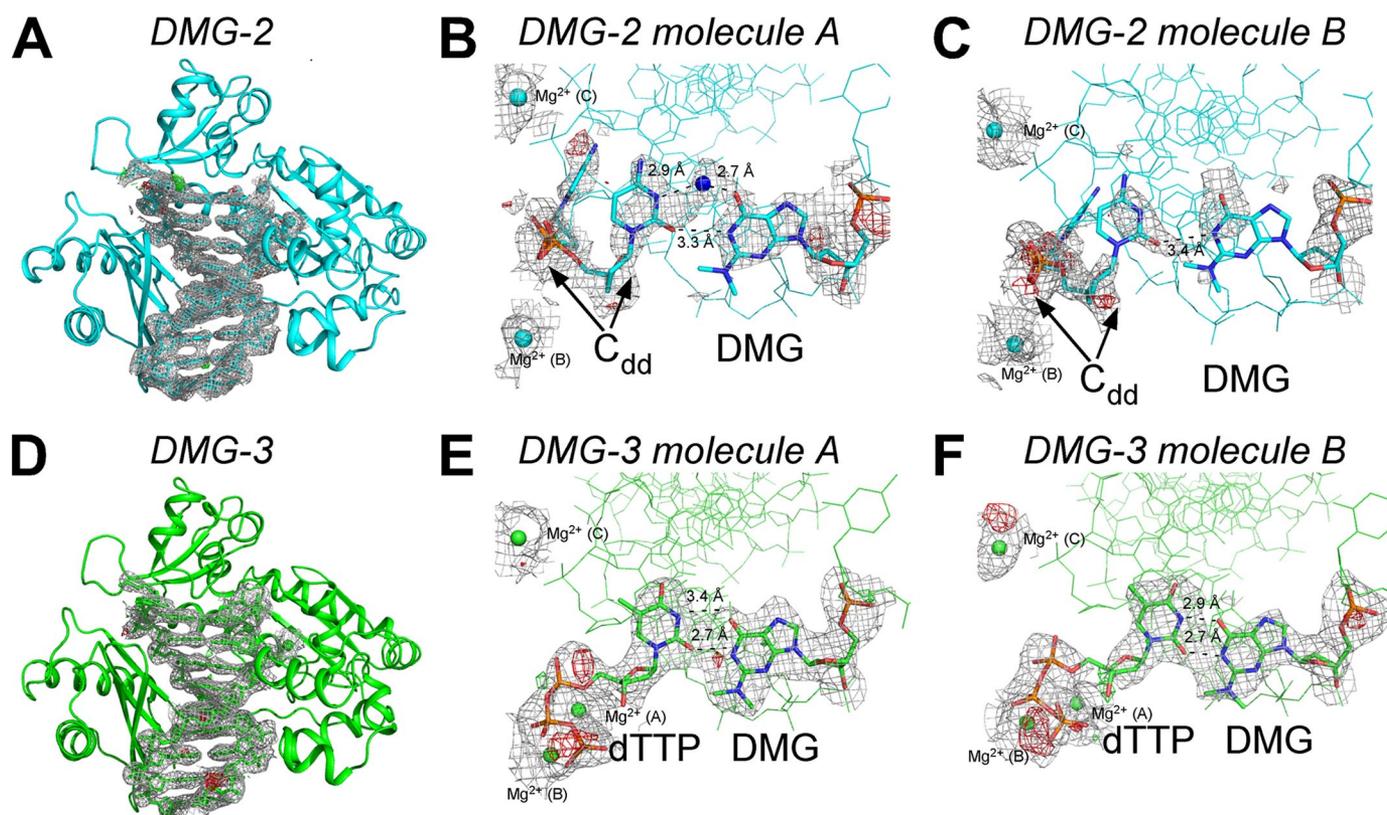


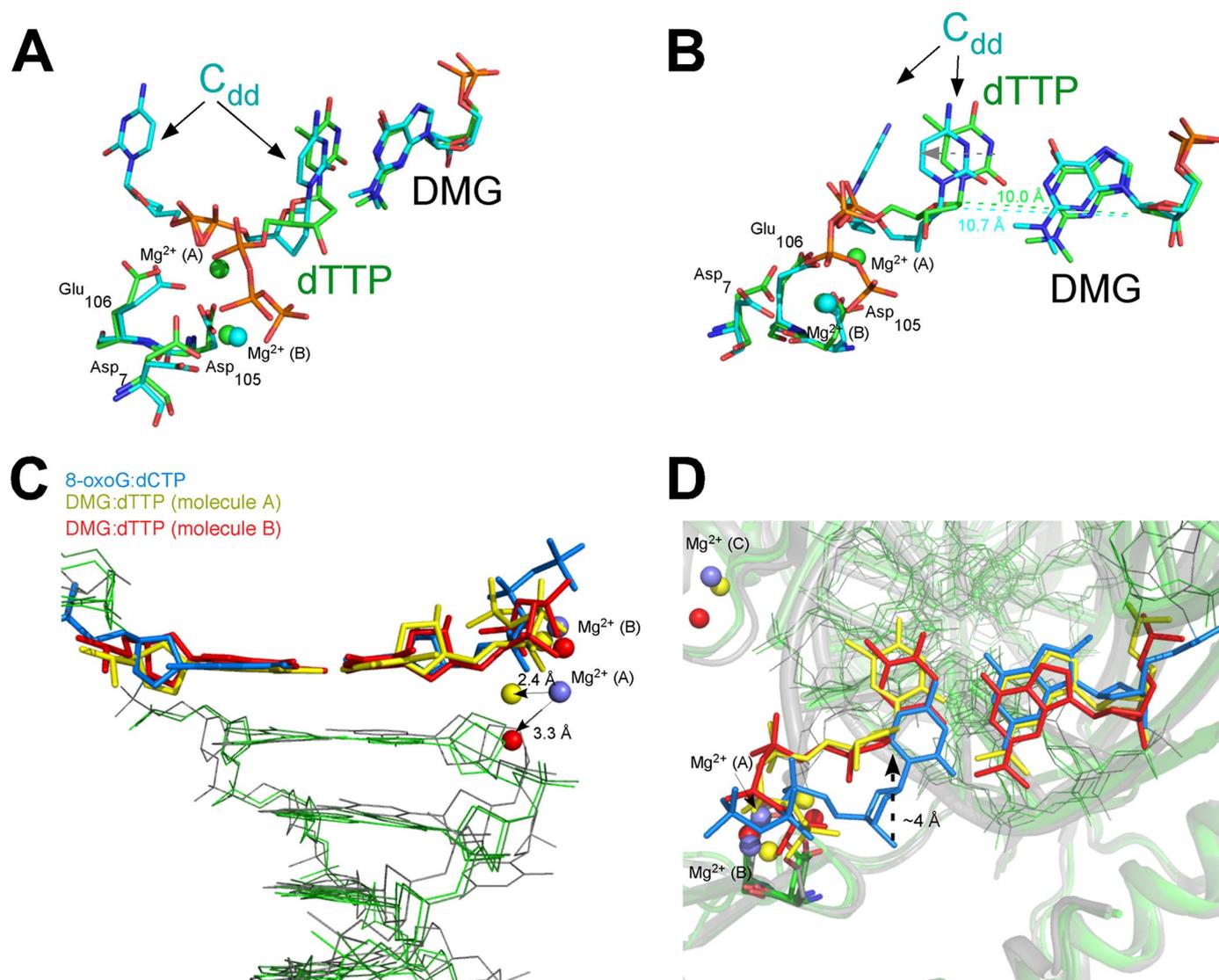
FIGURE 4. **Structural examination of wobble pairing with  $N^2,N^2$ -Me<sub>2</sub>G-modified DNA in Dpo4.** A, overall structure of DNA and corresponding electron density observed in the active site of DMG-2 are shown. The orientation of bases observed in molecule A (B) and molecule B (C) of the DMG-2 structure is shown. D, overall structure of DNA and corresponding electron density observed in the active site of DMG-3 are shown. The orientation of bases observed in molecule A (E) and molecule B (F) of the DMG-3 structure is shown. In all panels, the  $3F_o - 2F_c$  map (gray mesh) for DNA bound in the DMG-2 structure is shown contoured to  $1\sigma$  level (gray mesh) with the  $F_o - F_c$  difference maps shown contoured to  $3\sigma$  and  $-3\sigma$  for positive (red mesh) and negative (green mesh) density, respectively.

polymerase active site of the DMG-1 structure. A third Mg<sup>2+</sup> ion (Fig. 3B, labeled C) is coordinated by water molecules and the carbonyl moieties in the amide backbone at residues Ala-181 and Ile-186, which are located in the thumb domain of Dpo4 (Fig. 3). A fourth Mg<sup>2+</sup> ion (Fig. 3B, labeled D) is coordinated between the  $\alpha$ -phosphate of the incoming dGTP and the

phosphate connecting the 13C and <sup>14</sup>C<sub>dd</sub> residues in the primer strand.

The second crystal structure (DMG-2) was found to be a binary complex with no incoming nucleoside triphosphate (Fig. 4). A single Mg<sup>2+</sup> ion was observed in both molecules of the asymmetric unit near the enzyme active site, in a position that is

## DNA Polymerase Dpo4 Structure and Function



**FIGURE 5. Comparison between catalytically inhibited Dpo4  $N^2,N^2$ -Me<sub>2</sub>G wobble pairs and the highly efficient 8-oxoG:dCTP structure (68, 69).** A, orientation of wobble base pairs observed in DMG-2 (cyan carbons) and DMG-3 (green carbons) is shown with the active-site residues and Mg<sup>2+</sup> ions labeled accordingly. B, rotated view of A illustrating the widened C-1'-C-1' distance observed for the  $N^2,N^2$ -Me<sub>2</sub>G:C<sub>dd</sub> wobble relative to  $N^2,N^2$ -Me<sub>2</sub>G:dTTP. C, superimposing molecules A (base pair and metal ions in yellow, everything else in green) and B (base pair and metal ions in red, everything else in green) from the DMG-3 structure with a ternary structure of Dpo4 inserting dCTP opposite 8-oxoG (base pair and metal ions in blue, everything else in gray; PDB code 2c2e) shows that the catalytic metal ion is shifted in  $N^2,N^2$ -Me<sub>2</sub>G:dTTP pairing events. D, rotated view of C shows how the wobble pairing mode shifts the incoming dNTP away from the active site of Dpo4.

normally associated with coordination of the  $\beta$ - and  $\gamma$ -phosphate groups on the incoming dNTP (Fig. 4). A second Mg<sup>2+</sup> ion was coordinated by residues in the thumb domain, similar to those described for DMG-1 (Fig. 4). The primer used in the DMG-2 complex was designed to pair the terminal C<sub>dd</sub> with the  $N^2,N^2$ -Me<sub>2</sub>G template residue. Clear positive difference Fourier electron density was observed for two alternate conformations for the primer terminus (Fig. 4, B and C). One orientation was similar to that observed in DMG-1, where the primer terminus folds back into the growing minor groove and is clearly nonproductive. The second conformation paired C<sub>dd</sub> with  $N^2,N^2$ -Me<sub>2</sub>G. The pairing was loose in the sense that the acceptor and donor atoms for one hydrogen bonding pair (O-2 of C<sub>dd</sub> and N-1 of  $N^2,N^2$ -Me<sub>2</sub>G) are located 3.3 Å from one another, and a second hydrogen bond between N-3 of C<sub>dd</sub> and O-6 of  $N^2,N^2$ -Me<sub>2</sub>G is mediated by a water molecule (Fig. 4B). Of the

two molecules in the asymmetric unit, molecule A was better ordered and showed better density near the  $N^2,N^2$ -Me<sub>2</sub>G site (Fig. 4, B and C).

In the DMG-3 insertion complex, dTTP forms a wobble pair with the  $N^2,N^2$ -Me<sub>2</sub>G template in the Dpo4 active site (Fig. 4, D–F). Two hydrogen bonds were formed between the N-1 atom of  $N^2,N^2$ -Me<sub>2</sub>G and the O-2 atom of the incoming dTTP (2.7 Å for molecules A and B in the asymmetric unit, respectively) and between the O-6 atom of  $N^2,N^2$ -Me<sub>2</sub>G and the N-1 atom of dTTP (3.4 and 2.9 Å for molecules A and B in the asymmetric unit, respectively) (Fig. 4, E and F). The presence of the two Me groups at the 2-position of G pushes the dTTP toward the major groove side to form a “wobble” pairing mode. The wobble observed here is different from the wobble pair formed between the N-1 atom and the exocyclic amino group of a template O<sup>6</sup>-MeG (70) or O<sup>6</sup>-BzG (57) and the O-4 and N-3 atoms of

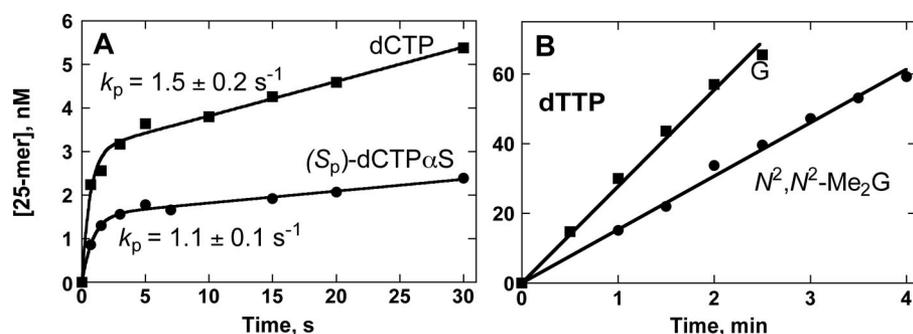


FIGURE 6. Incorporation of dCTP, ( $S_p$ )-dCTP $\alpha$ S, and dTTP opposite G and  $N^2,N^2$ -Me $_2$ G by Dpo4. A, pre-steady-state incorporation of dCTP (1 mM,  $\blacksquare$ ) or ( $S_p$ )-dCTP $\alpha$ S (1 mM,  $\bullet$ ) into a 24-mer primer opposite a 36-mer template containing  $N^2,N^2$ -Me $_2$ G (120 nM) by Dpo4 (70 nM). The burst rates ( $k_p$ ) were estimated to be  $1.5 (\pm 0.2) \text{ s}^{-1}$  for dCTP and  $1.1 (\pm 0.1) \text{ s}^{-1}$  for dCTP $\alpha$ S. B, incorporation of dTTP (1 mM) opposite 36-mer at the G ( $\blacksquare$ ) and  $N^2,N^2$ -Me $_2$ G ( $\bullet$ ) sites by Dpo4 (70 nM). No bursts were observed, and the data points were fit to a linear equation, with rates estimated as  $0.46 (\pm 0.01) \text{ s}^{-1}$  for the G oligonucleotide and  $0.026 (\pm 0.004) \text{ s}^{-1}$  for the  $N^2,N^2$ -Me $_2$ G oligonucleotide.

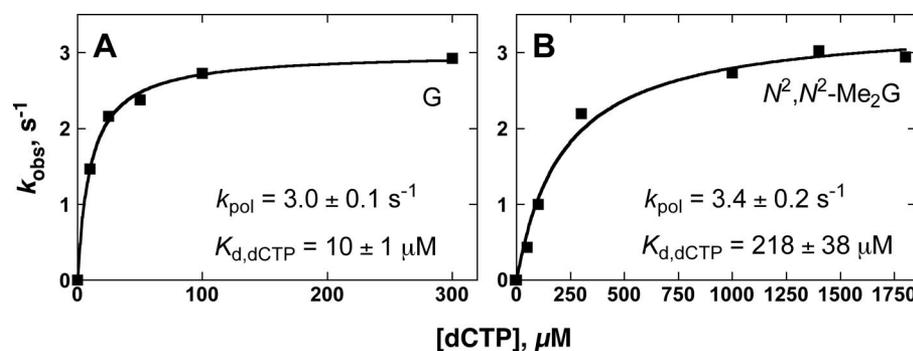


FIGURE 7. Estimation of  $k_{\text{pol}}$  and  $K_{d,\text{dCTP}}$  for by pre-steady-state burst kinetic analysis. Single turnover experiments were done with Dpo4 (200 nM) and 100 nM 24-mer primer/36-mer template complex containing G (A) or  $N^2,N^2$ -Me $_2$ G (B) and varying dCTP concentrations (2–1800  $\mu\text{M}$ ). Burst rates ( $k_{\text{obs}}$ , fit to Equation 2) versus [dCTP] were fit to a hyperbolic equation (Equation 3). (Data from the experiments with  $N^2$ -MeG are not shown.) The burst amplitudes for G,  $N^2$ -MeG, and  $N^2,N^2$ -Me $_2$ G were 42, 71, and 3 nM, respectively. Estimated values of  $k_{\text{pol}}$  for G,  $N^2$ -MeG, and  $N^2,N^2$ -Me $_2$ G were 3.0, 2.8, and  $3.4 \text{ s}^{-1}$ , respectively. Estimated  $K_{d,\text{dCTP}}$  values for G,  $N^2$ -MeG, and  $N^2,N^2$ -Me $_2$ G were 10, 9, and 220  $\mu\text{M}$ , respectively.

dCTP, respectively. The  $^{13}\text{C}_{\text{dd}}$  at the primer termini of DMG-3 was paired with the 3'-template G next to  $N^2,N^2$ -Me $_2$ G in a standard Watson-Crick base pairing mode. The  $\alpha$ -phosphate of the incoming dTTP is positioned  $\sim 4 \text{ \AA}$  from the primer terminus, and two of the  $\text{Mg}^{2+}$  ions were coordinated in the active site of the enzyme by residues Asp-7, Asp-105, and Glu-106. However, the catalytic  $\text{Mg}^{2+}$  ion ( $\text{Mg}^{2+}$  A) was shifted away from the aspartate/glutamate side chains by 2.4 and 3.3  $\text{\AA}$  in molecules 1 and 2 of the asymmetric unit, respectively, relative to what was observed in the DMG-1 structure. The relevance of this change in metal ion coordination toward Dpo4 catalysis is examined further under "Discussion."

**Pre-steady-state Incorporation Opposite G,  $N^2$ -MeG, or  $N^2,N^2$ -Me $_2$ G**—Pre-steady-state kinetics of dCTP or dTTP incorporation opposite  $N^2,N^2$ -Me $_2$ G were analyzed to discriminate correct incorporation from other events at the active site of Dpo4. Dpo4 clearly showed the expected bursts for dCTP incorporation opposite G and  $N^2$ -MeG (see below) and also a small burst phase for incorporation of dCTP opposite  $N^2,N^2$ -Me $_2$ G (Fig. 6A), indicating that the initial reaction at the active site of Dpo4 was still faster than the following steady-state rate, *i.e.* a step following product formation is rate-limiting in steady-state kinetics. The 36-mer template containing  $N^2,N^2$ -Me $_2$ G had been prepared according to a post-oligomerization meth-

odology, in which  $(\text{CH}_3)_2\text{NH}$  is reacted with 2-fluoro- $O^6$ -(trimethylsilylethyl)-2'-deoxyinosine in a 36-mer template oligonucleotide, and the product is purified by high pressure liquid chromatography and denaturing gel electrophoresis (43). To verify that the small burst was real, the purity of the 36-mer template containing  $N^2,N^2$ -Me $_2$ G was reanalyzed by matrix-assisted laser desorption MS. Only one peak was detected, corresponding to the 36-mer with  $N^2,N^2$ -Me $_2$ G (supplemental Fig. S6), as in the original analysis (43), supporting the view that the small burst originated from the fast incorporation of dCTP opposite  $N^2,N^2$ -Me $_2$ G.

If the phosphodiester bond formation steps were rate-limiting, then the incorporation rate of ( $S_p$ )-dCTP $\alpha$ S should be reduced compared with dCTP because the positive charge on the  $\alpha$ -P atom is decreased by the substitution of an  $\alpha$ -oxygen by sulfur (71). The ratio of the burst rates for dCTP or ( $S_p$ )-dCTP $\alpha$ S incorporation was 1.4 (Fig. 6A), *i.e.* without an obvious thio substitution effect, suggesting that phosphodiester bond formation is not the rate-limiting event in the burst phase reaction. No burst was

detected for the incorporation of dTTP opposite  $N^2,N^2$ -MeG (Fig. 6B), and product formation fit to a linear equation.

Analysis of the pre-steady-state burst rate as a function of increasing dCTP concentration yielded a maximum rate of nucleotide incorporation opposite  $N^2,N^2$ -Me $_2$ G ( $k_{\text{pol}}$ ) of  $3.4 (\pm 0.2) \text{ s}^{-1}$  and a  $K_{d,\text{dCTP}}$  of  $220 (\pm 40) \mu\text{M}$  (Fig. 7 and supplemental Fig. S5), the latter reflects the apparent binding affinity of the dCTP to the enzyme-DNA binary complex to form a ternary complex poised for catalysis (54, 55). The presence of one Me group at G (*i.e.*  $N^2$ -MeG) increased the burst amplitude but had little effect on the burst rate or dCTP binding affinity (46). The presence of the second Me group ( $N^2,N^2$ -Me $_2$ G) notably decreased the burst amplitude by 24-fold and dCTP binding affinity by 24-fold. However, the measured burst rate ( $k_{\text{pol}}$ ) was similar for the incorporation of dCTP opposite G,  $N^2$ -MeG, or  $N^2,N^2$ -Me $_2$ G.

**Kinetic Simulation of a Polymerase Mechanism for Dpo4**—The minimal mechanism (Fig. 8, without step 8) is generally applicable for DNA polymerases when a stoichiometric burst amplitude is observed (64, 72, 73). As expected from kinetic modeling work with replicative DNA polymerases (64, 65) and Dpo4 with  $N^2$ -alkyl G adducts (46), a modified mechanism with an additional nonproductive E-DNA-dNTP ternary complex (Fig. 8A, step 8) provides a reasonable fit to the experimental

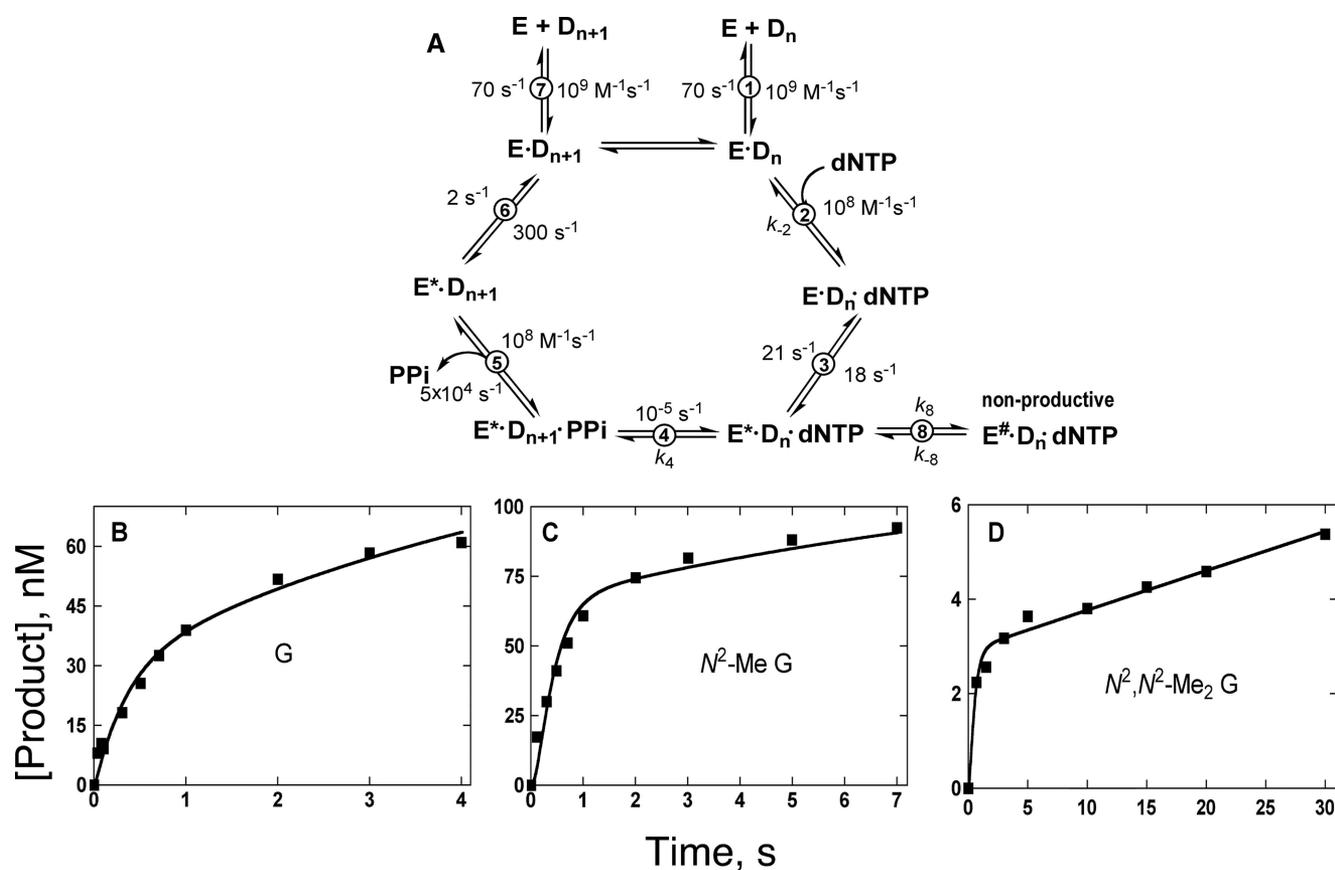


FIGURE 8. **Kinetic simulations of bursts.** *A*, general DNA polymerase mechanism. Individual steps are numbered. *E*, polymerase;  $D_n$ , DNA substrate;  $E^*$ , conformationally modified polymerase;  $E^\#$ , inactive polymerase conformation;  $D_{n+1}$ , DNA extended by one base, and  $\text{PP}_i$ , pyrophosphate. Forward and reverse rate constants for each step are presented from simulations of pre-steady-state nucleotide incorporation reactions with unmodified substrate. See Table 6 for adjustments of  $k_{-2}$ ,  $k_4$ ,  $k_8$ , and  $k_{-8}$ . All other rate constants remain the same in *B–D*. *B*, simulated fit to G data points. *C*, simulated fit to  $N^2$ -MeG data points. *D*, simulated fit to  $N^2,N^2$ -Me<sub>2</sub>G data points.

**TABLE 6**  
Kinetic modeling of pre-steady-state rate constants for Dpo4 reaction steps

In Fig. 8A, ED is  $E \cdot D_n$ , N is dNTP, EDN is  $E^* \cdot D_n \cdot \text{dNTP}$ , XDN is  $E^* \cdot D_{n+1} \cdot \text{PP}_i$ , ODN is  $E^* \cdot D_n \cdot \text{dNTP}$ , and  $k_{-2}$ ,  $k_4$ ,  $k_8$ , and  $k_{-8}$  are as indicated.

	ED + N $\rightleftharpoons$ EDN, $k_{-2}$	EDN $\rightleftharpoons$ XDN, $k_4$	XDN $\rightleftharpoons$ ODN		$k_8/k_4$
			$k_8$	$k_{-8}$	
	$\text{s}^{-1}$	$\text{s}^{-1}$	$\text{s}^{-1}$	$\text{s}^{-1}$	
G	1500	$3.2 \pm 0.7$	$2.1 \pm 1.2$	$0.07 \pm 0.19$	0.66
$N^2$ -MeG	1500	$6.1 \pm 0.6$	$1.4 \pm 0.2$	$10^{-5}$	0.23
$N^2,N^2$ -Me <sub>2</sub> G	21,800	$0.23 \pm 0.04$	$5.0 \pm 0.9$	$0.025 \pm 0.003$	22

Indicated rates and errors were estimated using the program Dynafit. See Fig. 7 for fits.

results for a system with partial bursts. Initial rate constants used in fitting were based upon a recent kinetic analysis of some Dpo4 mutants (66), and  $K_{d,\text{dCTP}}$  values of 15, 15, and 220  $\mu\text{M}$  were used for oligonucleotides containing G,  $N^2$ -MeG, and  $N^2,N^2$ -Me<sub>2</sub>G, respectively, based on the pre-steady-state kinetic determinations (Fig. 6). The alternate mechanism (Fig. 8A) can fit the experimental results well, using the enzyme concentrations estimated by UV measurements, and yields  $k_4$ ,  $k_8$ , and  $k_{-8}$  for the oligonucleotides containing G,  $N^2$ -MeG or  $N^2,N^2$ -Me<sub>2</sub>G (Fig. 8, B–D, and Table 6). The ratio of rate constants ( $k_8/k_4$ ) for the conversion of the intermediate  $E^* \cdot D_n \cdot \text{dNTP}$  to the nonproductive complex *versus* product ( $E^* \cdot D_{n+1} \cdot \text{PP}_i$ ) was increased from 0.23 for  $N^2$ -MeG to 22 for  $N^2,N^2$ -Me<sub>2</sub>G. When the mechanism was simulated with the same dCTP binding affinity ( $k_{-2}/k_2$ ) for the G,  $N^2$ -MeG, and  $N^2,N^2$ -Me<sub>2</sub>G substrates, the mechanism with the nonproduc-

tive complex still fit the experimental data well and yielded similar rate constants for  $k_4$ ,  $k_8$ , and  $k_{-8}$  (results not shown).

## DISCUSSION

Chemical modification at the N-2 atom of guanine has been studied extensively in the context of understanding DNA damage and mechanisms of miscoding by DNA polymerases. Replicative DNA polymerases are challenged by N-2 bulk on G even as small as Me or Et. Y-family DNA polymerases efficiently replicate past these forms of damage (42–46). However, the addition of a second Me group to  $N^2$ -MeG poses a strong block to Y-family DNA polymerases unless they use alternate pairing modes (*e.g.* Hoogsteen pairing or an amino acid in the protein (42, 44)). We utilized this block with the Y-family DNA polymerase of *S. solfataricus*, Dpo4, in a series of kinetic and structural studies to better understand how such a small chemical

change results in such a large functional difference for this important class of enzymes.

At the outset, two major hypotheses could be proposed to account for the dramatic loss of catalytic activity upon addition of the second Me group at the N-2 atom. The first hypothesis would predict that the presence of one Me group does not influence the normal hydrogen bonding pattern between C and G but the addition of a second Me will and that this loss of a hydrogen bond disrupts catalysis. There is some evidence to support that idea that Y-family polymerases are more sensitive to loss of hydrogen bonding capability than replicative pols (74–77). However, in recent work we have found that hypoxanthine is nearly as good in directing dCTP incorporation by Dpo4 as a template G.<sup>4</sup> Hypoxanthine does not have an exocyclic amino group at the C-2 position and can only form two hydrogen bonds with an incoming dCTP. Thus, the results with hypoxanthine would seem to minimize the importance of hydrogen bonding in explaining the reduced efficiency of Dpo4 during  $N^2,N^2$ -Me<sub>2</sub>G bypass. Moreover, Watson-Crick A:T pairs have only two hydrogen bonds. An alternate hypothesis is that the presence of two Me groups on the N-2 atom induces an orientation in the enzyme·DNA complex that is distinct from normal Watson-Crick geometry and inhibitory to polymerization. In this study we provide structural evidence that clearly supports the latter hypothesis, but there is also reason to consider the importance of hydrogen bonds during nucleotide selectivity by Dpo4 when Watson-Crick geometry is distorted.

Major changes in the protein itself were not seen in the crystal structures but multiple orientations of the DNA/dNTP substrates were observed, including one structure with C<sub>dd</sub> positioned opposite  $N^2,N^2$ -Me<sub>2</sub>G (Fig. 4, A–C) and one with dTTP opposite  $N^2,N^2$ -Me<sub>2</sub>G (Fig. 4, D–F). The unique arrangement observed in *DMG-2* illustrates why the efficiency of dCTP incorporation opposite  $N^2,N^2$ -Me<sub>2</sub>G is so severely perturbed. If we assume that the pairing mode observed between C<sub>dd</sub> and  $N^2,N^2$ -Me<sub>2</sub>G in *DMG-2* is the orientation observed during dCTP incorporation, then the contacts between Dpo4 and the template near the active site would need to shift the incoming dCTP residue down into the polymerase active site for catalysis to occur (Fig. 4B). Even though there is no incoming dNTP in the *DMG-2* structure, it is interesting to note that the binding site for the catalytic metal ion is empty (Fig. 4B). Only the Mg<sup>2+</sup> ion that would coordinate the β- and γ-phosphates is present in the context of the  $N^2,N^2$ -Me<sub>2</sub>G:C pair. The conformation of C<sub>dd</sub> positioned in the active site of the *DMG-2* structure is in a wobble orientation relative to the template  $N^2,N^2$ -Me<sub>2</sub>G. There is only one pair of hydrogen bond acceptor and donor atoms possible in the wobble orientation, namely the O-2 atom of C<sub>dd</sub> and the N-1 atom of  $N^2,N^2$ -Me<sub>2</sub>G. A second hydrogen bond between the N-3 atom of C<sub>dd</sub> and the O-6 atom of  $N^2,N^2$ -Me<sub>2</sub>G is mediated by a water molecule. It should be noted that electron density for the mediating water molecule is present in only one of the molecules of the asymmetric unit. However, the distances between C<sub>dd</sub> and  $N^2,N^2$ -Me<sub>2</sub>G are similar in both molecules.

In contrast, the incoming dTTP found in the *DMG-3* structure forms two stable hydrogen bond pairs with  $N^2,N^2$ -Me<sub>2</sub>G (Fig. 4, E and F). One hydrogen bond occurs between the O-2 atom of dTTP and the N-1 atom of  $N^2,N^2$ -Me<sub>2</sub>G with a second pairing between the N-3 atom of dTTP and the O-6 atom of  $N^2,N^2$ -Me<sub>2</sub>G. In essence, the  $N^2,N^2$ -Me<sub>2</sub>G:C<sub>dd</sub> and  $N^2,N^2$ -Me<sub>2</sub>G:T pairs exhibit similar wobble geometry (Fig. 5A). The C-1'–C-1' distance for the  $N^2,N^2$ -Me<sub>2</sub>G:C<sub>dd</sub> pair is 0.7 Å wider than for the  $N^2,N^2$ -Me<sub>2</sub>G:T (10.7 and 10.0 Å, respectively) because of the mediating water molecule (Fig. 5B). The kinetic preference for dTTP insertion (Table 2) may represent the fact that one of the two hydrogen bonds occurring between dCTP and  $N^2,N^2$ -Me<sub>2</sub>G relies upon a water-mediated association, which is transient and displays a wider C-1'–C-1' distance compared with the  $N^2,N^2$ -Me<sub>2</sub>G:T pairing. Superimposition of *DMG-3* with a structure of Dpo4 inserting dCTP opposite 8-oxoG (PDB code 2C2E) reveals how loosely bound the incoming dTTP is relative to the 8-oxoG:dCTP pair, which is a highly efficient catalytic event (Fig. 5, C and D). The  $N^2,N^2$ -Me<sub>2</sub>G and 8-oxoG residues superimpose well, but the 3'-hydroxyl group of the incoming dTTP is shifted ~4 Å away from the positions of the corresponding moiety in the 8-oxoG:dCTP pair (Fig. 5). The looseness of the binding event is further evidenced by the catalytic magnesium ion in *DMG-3*, which is shifted away from the Asp-7, Asp-105, and Glu-106 residues in the palm domain (2.4 Å for molecule A and 3.3 Å for molecule B) relative to the catalytic metal ion position in the 8-oxoG:dCTP structure (Fig. 5C). The metal ion that stabilizes the triphosphate moiety of the incoming dNTP is coordinated in a similar fashion for both the *DMG-3* and 8-oxoG:dCTP structures (Fig. 5, C and D). Previous reports from Aggarwal and co-workers (78) and Yang and co-workers (79) showed that the catalytic metal ion shifts ~3 Å away from the active site when Dpo4 binds a T:dGTP wobble mispair, similar to the structure containing a  $N^2,N^2$ -Me<sub>2</sub>G:dTTP pair. However, the steady-state catalytic efficiency of T:dGTP (incorporation) is only reduced ~27-fold relative to dC:dGTP, indicating asymmetry in the efficiencies of the T:G wobble pairs that is dependent upon the identity of the incoming dNTP. Superimposition of the T:dGTP complex (PDB code 2agp) with the *DMG-3* complex (*i.e.*  $N^2,N^2$ -Me<sub>2</sub>G:dTTP pair) shows that both the primer terminus and the incoming dTTP in the *DMG-3* complex fail to shift down into the active site in a manner similar to what is observed in the active site of the T:dGTP wobble pair, which apparently serves to dramatically reduce catalytic efficiency (from 27- to 6,700-fold) when combined with a loosely coordinated metal ion. The catalytic Mg<sup>2+</sup> ion is important for proton abstraction at the 3'-hydroxyl group and subsequent nucleophilic attack upon the α-phosphate of the incoming dNTP. The distance between the dideoxy-primer terminus and the α-phosphate of the incoming dTTP in our structure is near optimal (~4 Å), but one would predict that Dpo4 has a decreased ability to catalyze phosphoryl transfer based on the shifted position of the catalytic magnesium ion observed in *DMG-3*.

A rapid kinetic burst of product formation was observed for the incorporation of dCTP (but not dTTP) opposite  $N^2,N^2$ -Me<sub>2</sub>G (Fig. 6A). The observed burst amplitude only accounts

<sup>4</sup> H. Zhang, U. Bren, I. D. Kozekov, C. J. Rizzo, D. F. Stec, and F. P. Guengerich, manuscript in preparation.

## DNA Polymerase Dpo4 Structure and Function

for ~5% of the enzyme, consistent with the view that Dpo4 adopts nonproductive complex structures regardless of which dNTP is selected for insertion opposite  $N^2,N^2$ -Me<sub>2</sub>G. A kinetic simulation in which the rate constants of steps are adjusted only in the basic catalytic cycle, without multiple complexes, cannot fit the partial burst results, as we first demonstrated in a study with HIV-1 reverse transcriptase (64, 65). As an additional comparison, we checked the ability of bacteriophage DNA polymerase T7 (exonuclease<sup>-</sup>) to form productive complexes with  $N^2,N^2$ -Me<sub>2</sub>G, but no burst in product formation was observed for polymerase T7<sup>-</sup>-catalyzed insertion of dCTP opposite  $N^2,N^2$ -Me<sub>2</sub>G (results not presented). For Dpo4-catalyzed insertion of dCTP opposite  $N^2,N^2$ -Me<sub>2</sub>G the burst rate constant,  $k_{\text{pol}}$ , was very similar to that measured for G (Fig. 7), although the apparent  $K_{d,\text{dCTP}}$  was 20-fold higher. The increased  $K_{d,\text{dCTP}}$  is supportive of the view that Dpo4 uses the wobble pairing observed in DMG-2 to insert dCTP opposite  $N^2,N^2$ -Me<sub>2</sub>G. In most instances the wobble pairing will result in nonproductive complex formation, as judged by the reduced burst amplitude. However, we conclude that a small fraction of the ternary (enzyme·DNA·dNTP) complex is able to form product rapidly (Fig. 6A), presumably by trapping the template  $N^2,N^2$ -Me<sub>2</sub>G base in a conformation that forms a wobble pair with the incoming dCTP but one that is also conducive to rapid polymerization (*i.e.* tighter coordination of the catalytic metal ion). Such a conclusion is supported by our structural data, and the existence of a burst phase in the transient state that is defined by a higher  $K_{d,\text{dCTP}}$  for polymerization. Why Dpo4 does not exhibit burst kinetics during dTTP insertion opposite  $N^2,N^2$ -Me<sub>2</sub>G is unclear, given the fact that both dCTP and dTTP are likely to form wobble base pairs with the adduct. The basic kinetic scheme can be readily adapted to model systems that equilibrate between productive and nonproductive ternary systems (Fig. 8A), although the point should be made that this is still a minimal kinetic scheme and multiple E·DNA·dNTP complexes might be represented by the collective entity  $E^{\#}\cdot D_n\cdot d\text{NTP}$ , each with its own rate constants.

In conclusion, our structures show that there are at least three possible active site orientations for the DNA polymerase Dpo4 and the small but very blocking adduct,  $N^2,N^2$ -Me<sub>2</sub>G. The modified  $N^2,N^2$ -Me<sub>2</sub>G residue remains in the *anti* orientation for all of the structures solved. Hence, the reduced efficiency and fidelity of Y-family pols that utilize normal Watson-Crick base pairing as the preferred mode of nucleotide selection result from the stable formation of wobble pairs, which more loosely associates the incoming dNTP and the catalytic metal ion with the polymerase active site. Y-family DNA polymerases that utilize non-Watson-Crick modes of pairing during nucleotide selection, *e.g.* pol  $\iota$  and REV1, do not exhibit such dramatic inhibition of polymerase activity (43, 45). In the case of Dpo4, a small fraction of the polymerase·oligonucleotide complex remains productive during dCTP incorporation and is as inherently fast as with the normal G. The rest of the polymerase·oligonucleotide population is in slow equilibrium with the productive complex, and this model can be fit with rate constants for individual reaction steps that can explain all of the kinetic data, which in turn has support in the structural work.

*Acknowledgments*—We thank K. C. Angel for technical assistance and K. Trisler for assistance in preparation of the manuscript. We are grateful to Z. Wawrzak, Northwestern University, for assistance with x-ray diffraction data collection. Vanderbilt University is a member institution of the Life Sciences Collaborative Access Team at Sector 21 of the Advanced Photon Source, Argonne National Laboratory, Argonne, IL. Use of the Advanced Photon Source was supported by the United States Department of Energy, Basic Energy Sciences, Office of Science, under Contract W-31-109-Eng-38.

## REFERENCES

1. Schaaper, R. M., Glickman, B. W., and Loeb, L. A. (1982) *Cancer Res.* **42**, 3480–3485
2. Nakamura, J., Walker, V. E., Upton, P. B., Chiang, S. Y., Kow, Y. W., and Swenberg, J. A. (1998) *Cancer Res.* **58**, 222–225
3. Wink, D. A., Kasprzak, K. S., Maragos, C. M., Elespuru, R. K., Misra, M., Dunams, T. M., Cebula, T. A., Koch, W. H., Andrews, A. W., Allen, J. S., and Keefer, L. K. (1991) *Science* **254**, 1001–1003
4. Nguyen, T., Brunson, D., Crespi, C. L., Penman, B. W., Wishnok, J. S., and Tannenbaum, S. R. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89**, 3030–3034
5. Kuchino, Y., Mori, F., Kasai, H., Inoue, H., Iwai, S., Miura, K., Ohtsuka, E., and Nishimura, S. (1987) *Nature* **327**, 77–79
6. Aruoma, O. I., Halliwell, B., and Dizdaroglu, M. (1989) *J. Biol. Chem.* **264**, 13024–13028
7. Cheng, K. C., Cahill, D. S., Kasai, H., Nishimura, S., and Loeb, L. A. (1992) *J. Biol. Chem.* **267**, 166–172
8. Searle, C. E. (ed) (1984) *Chemical Carcinogens*, 2nd Ed, American Chemical Society, Washington, DC
9. Friedberg, E. C., Walker, G. C., Siede, W., Wood, R. D., Schultz, R. A., and Ellenberger, T. (2006) *DNA Repair and Mutagenesis*, 2nd Ed., American Society for Microbiology, Washington, DC
10. Mouret, S., Baudouin, C., Charveron, M., Favier, A., Cadet, J., and Douki, T. (2006) *Proc. Natl. Acad. Sci. U.S.A.* **103**, 13765–13770
11. Douki, T., and Cadet, J. (2001) *Biochemistry* **40**, 2495–2501
12. Mao, H., Schnetz-Boutaud, N. C., Weisenseel, J. P., Marnett, L. J., and Stone, M. P. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 6615–6620
13. Guengerich, F. P., Langouët, S., Mican, A. N., Akasaka, S., Müller, M., and Persmark, M. (1999) *IARC Sci. Publ.* **150**, 137–145
14. de los Santos, C., Zaliznyak, T., and Johnson, F. (2001) *J. Biol. Chem.* **276**, 9077–9082
15. Stone, M. P., Cho, Y. J., Huang, H., Kim, H. Y., Kozekov, I. D., Kozekova, A., Wang, H., Minko, I. G., Lloyd, R. S., Harris, T. M., and Rizzo, C. J. (2008) *Acc. Chem. Res.* **41**, 793–804
16. Petruska, J., Goodman, M. F., Boosalis, M. S., Sowers, L. C., Cheong, C., and Tinoco, I., Jr. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 6252–6256
17. Showalter, A. K., and Tsai, M. D. (2002) *Biochemistry* **41**, 10571–10576
18. Joyce, C. M., and Benkovic, S. J. (2004) *Biochemistry* **43**, 14317–14324
19. Guengerich, F. P. (2006) *Chem. Rev.* **106**, 420–452
20. Johnson, K. A. (2008) *J. Biol. Chem.* **283**, 26297–26301
21. Yasui, M., Matsui, S., Ihara, M., Laxmi, Y. R., Shibutani, S., and Matsuda, T. (2001) *Nucleic Acids Res.* **29**, 1994–2001
22. Terashima, I., Matsuda, T., Fang, T. W., Suzuki, N., Kobayashi, J., Kohda, K., and Shibutani, S. (2001) *Biochemistry* **40**, 4106–4114
23. Liu, X., Lao, Y., Yang, I. Y., Hecht, S. S., and Moriya, M. (2006) *Biochemistry* **45**, 12898–12905
24. Lao, Y., and Hecht, S. S. (2005) *Chem. Res. Toxicol.* **18**, 711–721
25. Zhang, S., Villalta, P. W., Wang, M., and Hecht, S. S. (2006) *Chem. Res. Toxicol.* **19**, 1386–1392
26. Forgacs, E., Latham, G., Beard, W. A., Prasad, R., Bebenek, K., Kunkel, T. A., Wilson, S. H., and Lloyd, R. S. (1997) *J. Biol. Chem.* **272**, 8525–8530
27. Wang, M., Lao, Y., Cheng, G., Shi, Y., Villalta, P. W., and Hecht, S. S. (2007) *Chem. Res. Toxicol.* **20**, 625–633
28. Turesky, R. J., Rossi, S. C., Welti, D. H., Lay, J. O., Jr., and Kadlubar, F. F. (1992) *Chem. Res. Toxicol.* **5**, 479–490
29. Meehan, T., and Straub, K. (1979) *Nature* **277**, 410–412

30. Cheng, S. C., Hilton, B. D., Roman, J. M., and Dipple, A. (1989) *Chem. Res. Toxicol.* **2**, 334–340
31. Lao, Y., Yu, N., Kassie, F., Villalta, P. W., and Hecht, S. S. (2007) *Chem. Res. Toxicol.* **20**, 246–256
32. Fang, J. L., and Vaca, C. E. (1997) *Carcinogenesis* **18**, 627–632
33. Matsuda, T., Terashima, I., Matsumoto, Y., Yabushita, H., Matsui, S., and Shibutani, S. (1999) *Biochemistry* **38**, 929–935
34. Pattanayek, R., Sethaphong, L., Pan, C., Prhavc, M., Prakash, T. P., Manoharan, M., and Egli, M. (2004) *J. Am. Chem. Soc.* **126**, 15006–15007
35. Martin, J. L., and McMillan, F. M. (2002) *Curr. Opin. Struct. Biol.* **12**, 783–793
36. Limbach, P. A., Crain, P. F., and McCloskey, J. A. (1994) *Nucleic Acids Res.* **22**, 2183–2196
37. Edqvist, J., Stråby, K. B., and Grosjean, H. (1995) *Biochimie* **77**, 54–61
38. Steinberg, S., Misch, A., and Sprinzl, M. (1993) *Nucleic Acids Res.* **21**, 3011–3015
39. Armengaud, J., Urbonavicius, J., Fernandez, B., Chaussinand, G., Bujnicki, J. M., and Grosjean, H. (2004) *J. Biol. Chem.* **279**, 37142–37152
40. Marquet, R., Isel, C., Ehresmann, C., and Ehresmann, B. (1995) *Biochimie* **77**, 113–124
41. Choi, J. Y., and Guengerich, F. P. (2004) *J. Biol. Chem.* **279**, 19217–19229
42. Choi, J. Y., and Guengerich, F. P. (2005) *J. Mol. Biol.* **352**, 72–90
43. Choi, J. Y., and Guengerich, F. P. (2006) *J. Biol. Chem.* **281**, 12315–12324
44. Choi, J. Y., Angel, K. C., and Guengerich, F. P. (2006) *J. Biol. Chem.* **281**, 21062–21072
45. Choi, J. Y., and Guengerich, F. P. (2008) *J. Biol. Chem.* **283**, 23645–23655
46. Zhang, H., Eoff, R. L., Kozekov, I. D., Rizzo, C. J., Egli, M., and Guengerich, F. P. (2009) *J. Biol. Chem.* **284**, 3563–3576
47. Pence, M. G., Blans, P., Zink, C. N., Hollis, T., Fishbein, J. C., and Perrino, F. W. (2009) *J. Biol. Chem.* **284**, 1732–1740
48. Nair, D. T., Johnson, R. E., Prakash, L., Prakash, S., and Aggarwal, A. K. (2005) *Science* **309**, 2219–2222
49. Nair, D. T., Johnson, R. E., Prakash, S., Prakash, L., and Aggarwal, A. K. (2004) *Nature* **430**, 377–380
50. Nair, D. T., Johnson, R. E., Prakash, L., Prakash, S., and Aggarwal, A. K. (2008) *Structure* **16**, 239–245
51. 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
52. Borer, P. N. (1975) in *Handbook of Biochemistry and Molecular Biology* (Fasman, G. D., ed) 3rd Ed., pp. 589–590, CRC Press, Inc., Cleveland, OH
53. Goodman, M. F., Creighton, S., Bloom, L. B., and Petruska, J. (1993) *Crit. Rev. Biochem. Mol. Biol.* **28**, 83–126
54. Patel, S. S., Wong, I., and Johnson, K. A. (1991) *Biochemistry* **30**, 511–525
55. Johnson, K. A. (1995) *Methods Enzymol.* **249**, 38–61
56. Choi, J. Y., Chowdhury, G., Zang, H., Angel, K. C., Vu, C. C., Peterson, L. A., and Guengerich, F. P. (2006) *J. Biol. Chem.* **281**, 38244–38256
57. Eoff, R. L., Angel, K. C., Egli, M., and Guengerich, F. P. (2007) *J. Biol. Chem.* **282**, 13573–13584
58. Kabsch, W. (1988) *J. Appl. Crystallogr.* **21**, 916–924
59. Rechkooblit, O., Malinina, L., Cheng, Y., Kuryavii, V., Broyde, S., Geacintov, N. E., and Patel, D. J. (2006) *PLoS Biol.* **4**, e11
60. Vagin, A., and Teplyakov, A. (1997) *J. Appl. Crystallogr.* **30**, 1022–1025
61. Vellieux, F. M. D., and Dijkstra, B. W. (1997) *J. Appl. Crystallogr.* **30**, 396–399
62. 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. Sect. D Biol. Crystallogr.* **54**, 905–921
63. Kuzmic, P. (1996) *Anal. Biochem.* **237**, 260–273
64. Furge, L. L., and Guengerich, F. P. (1999) *Biochemistry* **38**, 4818–4825
65. Woodside, A. M., and Guengerich, F. P. (2002) *Biochemistry* **41**, 1039–1050
66. Beckman, J. W., Wang, Q., and Guengerich, F. P. (2008) *J. Biol. Chem.* **283**, 36711–36723
67. Mendelman, L. V., Petruska, J., and Goodman, M. F. (1990) *J. Biol. Chem.* **265**, 2338–2346
68. 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
69. Eoff, R. L., Irimia, A., Angel, K. C., Egli, M., and Guengerich, F. P. (2007) *J. Biol. Chem.* **282**, 19831–19843
70. Eoff, R. L., Irimia, A., Egli, M., and Guengerich, F. P. (2007) *J. Biol. Chem.* **282**, 1456–1467
71. Herschlag, D., Piccirilli, J. A., and Cech, T. R. (1991) *Biochemistry* **30**, 4844–4854
72. Suo, Z., and Johnson, K. A. (1997) *Biochemistry* **36**, 12459–12467
73. Suo, Z., and Johnson, K. A. (1998) *J. Biol. Chem.* **273**, 27259–27267
74. Irimia, A., Eoff, R. L., Pallan, P. S., Guengerich, F. P., and Egli, M. (2007) *J. Biol. Chem.* **282**, 36421–36433
75. Lee, H. R., Helquist, S. A., Kool, E. T., and Johnson, K. A. (2008) *J. Biol. Chem.* **283**, 14402–14410
76. Kim, T. W., Delaney, J. C., Essigmann, J. M., and Kool, E. T. (2005) *Proc. Natl. Acad. Sci. U.S.A.* **102**, 15803–15808
77. Mizukami, S., Kim, T. W., Helquist, S. A., and Kool, E. T. (2006) *Biochemistry* **45**, 2772–2778
78. Trincão, J., Johnson, R. E., Wölfle, W. T., Escalante, C. R., Prakash, S., Prakash, L., and Aggarwal, A. K. (2004) *Nat. Struct. Mol. Biol.* **11**, 457–462
79. Vaisman, A., Ling, H., Woodgate, R., and Yang, W. (2005) *EMBO J.* **24**, 2957–2967
80. Cambillau, C., and Roussel, A. (1997) *Turbo Frodo*, Version Open GL.1, Université Aix-Marseille II, Marseille, France
81. DeLano, W. L. (2002) *The PyMOL Molecular Graphics System*, DeLano Scientific LLC, San Carlos, CA