Versatility of Y-family Sulfolobus solfataricus DNA Polymerase Dpo4 in Translesion Synthesis Past Bulky N²-Alkyguanine Adducts*†§

Huidong Zhang1,2, Robert L. Eoff1, Ivan D. Kozekov3, Carmelo J. Rizzo3, Martin Egli1, and F. Peter Guengerich1,2

From the 1Department of Biochemistry and Center in Molecular Toxicology, Vanderbilt University School of Medicine, Nashville, Tennessee 37232-0146 and the 3Department of Chemistry and Center in Molecular Toxicology, Vanderbilt University, Nashville, Tennessee 37235

In contrast to replicative DNA polymerases, Sulfolobus solfataricus Dpo4 showed a limited decrease in catalytic efficiency (Kcat/Km) for insertion of dCTP opposite a series of N²-alkyguanine templates of increasing size from (methyl (Me) to (9-anthracenyl)-Me (Anth)). Fidelity was maintained with increasing size up to (2-naphthyl)-Me (Naph). The catalytic efficiency increased slightly going from the N²-NaphG to the N²-AnthG substrate, at the cost of fidelity. Pre-steady-state kinetic bursts were observed for dCTP incorporation throughout the series (N²-MeG to N²-AnthG), with a decrease in the burst amplitude and kpol, the rate of single-turnover incorporation. The pre-steady-state kinetic courses with G and all of the six N²-alkyl G adducts could be fit to a general DNA polymerase scheme to which was added an inactive complex in equilibrium with the active ternary Dpo4-DNA-dNTP complex, and only the rates of equilibrium with the inactive complex and phosphodiester bond formation were altered. Two crystal structures of Dpo4 with a template N²-NaphG (in a post-insertion register opposite a 3′-terminal C in the primer) were solved. One showed N²-NaphG in a syn conformation, with the naphthyl group located between the template and the Dpo4 “little finger” domain. The Hoogsteen face was within hydrogen bonding distance of the N4 atoms of the cytosine opposite N²-NaphG and the cytosine at the −2 position. The second structure showed N²-Naph G in an anti conformation with the primer terminus largely disordered. Collectively these results explain the versatility of Dpo4 in bypassing bulky G lesions.

Accurate and efficient replication of DNA is crucial for the preservation of genomic integrity and survival of organisms (1). A major obstacle for DNA replication is caused by various DNA lesions, which are inevitably formed by both endogenous sources and exogenous mutagens in cells and are usually present in replicating DNA after avoiding repair (2). DNA polymerases often produce abnormal phenomena when encountering lesions during DNA replication, including misinsertion of incorrect nucleotides, slippage, and blockage of replication, all of which can result in mutation or cell death and can lead to detrimental effects, including aging and cancer in mammals (3).

Some DNA adducts are small (e.g. abasic sites and oxidative adducts), but others are very bulky, e.g. pyrimidine dimers, photoproducts, large carcinogen-bound adducts, and cross-links (2). The bulkiness of DNA adducts may be a differentiating factor in the misincorporation and blockage of DNA polymerases. The N2 atom of guanine is susceptible to modification by various potential carcinogens, including formaldehyde (4), acetaldehyde (5–8), styrene oxide (9), N-nitrosopyrrolidine (10), oxidation products of heterocyclic aromatic amines (e.g. N-hydroxy-2-amino-3-methylimidazo[4,5-f]quinoline and N-hydroxy-2-amino-3,8-dimethylimidazo[4,5-f]quinoline (11)), and the oxidation products of various polycyclic aromatic hydrocarbons (e.g. benzo[a]pyrene (BPDE)3 (12, 13), N²-Et deoxyguanosine has been detected in granulocyte and lymphocyte DNA and urine of alcoholic patients (14–16). Misincorporation opposite N²-EtG by Escherichia coli DNA polymerase I Klenow fragment exonuclease (S) suggests possible relevance to mutation and subsequent carcinogenesis. Crotonaldehyde- and acetaldehyde-derived 1,N²-propanodeoxyguanosine adducts have been detected in DNA from human tissues (8). An N²-guanyl adduct has been detected in urine of rats treated with N-nitrosopyrrolidine (17), and pyridloxobutyl-derived N²-guanyl

* 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.). Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Basic Energy Sciences, Office of Science, under Contract W-31-109-Eng-38. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby deposited in the Protein Data Bank, Research Collaboratory for Structural

† The abbreviations used are: BPDE, benzo[a]pyrene diol epoxide; Anth, (9-anthracenyl)methyl; Bz, benzyl; CID, collision-induced dissociation; αTTPoS, 2′-deoxyctydylate 5′-O-({1-thiotriphosphate}; αTTPo5, 2′-deoxythymidine 5′-O-({1-thiotriphosphate); Dpo4, DNA polymerase IV; ESI, electrospray ionization; Et, ethyl; Ib, isobutyl; LC, liquid chromatography; Me, methyl; MS, mass spectrometry; MS/MS, tandem mass spectrometry; Naph, (2-naphthyl) methyl; 8-oxoG, 7,8-dihydro-8-oxodeoxyguanosine; pol, (DNA) polymerase; r.m.s.d., root mean square deviation; HIV, human immunodeficiency virus. For simplicity the N²-guanyl adducts are usually collectively referred to as “alkyl,” although some are aralkyl.

‡ 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.). Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Basic Energy Sciences, Office of Science, under Contract W-31-109-Eng-38. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby deposited in the Protein Data Bank, Research Collaboratory for Structural

§ The on-line version of this article (available at http://www.jbc.org) contains supplementary Figs. S1–S6 and Tables S1–S5.

1 Both authors contributed equally to this work.

2 To whom correspondence should be addressed: Dept. of Biochemistry and Center in Molecular Toxicology, Vanderbilt University School of Medicine, 638 Robinson Research Bldg., Nashville, TN 37232-0146. Tel.: 615-322-2261; Fax: 615-322-3141; E-mail: fguergerich@vanderbilt.edu.
Function of Dpo4 with N²-Guanyl Adducts

Adducts have been found in rats treated with the tobacco-specific nitrosamine N'-nitrosornicotine (18).

The nature of the mutations produced during replication is not only dependent on the specific damage in the DNA but also on the DNA polymerases involved (19, 20). Replicative DNA polymerases replicate unmodified DNA bases with high fidelity and efficiency but are generally intolerant of the DNA distortions caused by many DNA lesions; thus, these polymerases are blocked and/or inefficiently misinsert bases opposite DNA adducts during replication. The “translesion synthesis” DNA polymerases have more open and larger active sites and can synthesize DNA across and beyond various replication-blocking DNA lesions (21, 22). However, synthesis can still be inhibited (or blocked) by certain DNA lesions and bypass can be error-prone or error-free, depending on the adduct (23).

We previously studied incorporation and extension past the adducts N²-MeG, N²-EtG, N²-IbG, N²-BzG, N²-AnthG, N²-(2,3,4-trihydroxy-1-butyl)G, 8-hydroxy-1,N²-propanoG, 8-hydroxy-6-methyl-1,N²-propanoG, and the styrene oxide and polycyclic hydrocarbon products formed at the guanyl-N2 atom with the replicative DNA polymerases bacteriophage pol T7 and HIV-1 reverse transcriptase (24–26). Even the relatively small N²-MeG and N²-EtG adducts were miscoded with Klenow fragment (4, 5) and pol T7 and HIV-1 reverse transcriptase (24). N²-EtG and larger N²-guanyl adducts cause strong blockage of polymerization by pol T7 and HIV-1 reverse transcriptase (24). The eukaryotic DNA polymerases pol λ, κ, and η and Rev1 still preferred to incorporate dCTP opposite the N²-guanyl adducts, but the incorporation efficiency and the fidelity with bulkier N²-guanyl adducts were decreased compared with unmodified G (27–29). Although kinetic data have been accumulated for translesion synthesis by various DNA polymerases, the structures of the N²-guanyl adducts in the enzyme active sites and mechanistic details are still not clear.

The DNA polymerase Dpo4 from Sulfolobus solfataricus P2 (polymerase IV) (30) has been investigated in considerable detail both in terms of its function (31–35) and structure (36–38). Dpo4 can bypass various DNA adducts, including 8-oxoG (39, 40), O²-MeG (41), O⁶-BzG (42), UV-cross-linked products (43), BPDE adducts (44, 45), and abasic lesions (38). Several crystal structures of ternary complexes (with Dpo4, DNA template–primer construct containing an adducted base, and incoming dNTP) have been determined (40–44, 46, 47), and this structural data has revealed details of misincorporation in replication bypass.

In the present work, Dpo4 was studied in detail regarding both kinetics and structure in the processing of a series of N²-guanyl adducts ranging in size, including N²-MeG, N²-EtG, N²-IbG, N²-BzG, N²-NaphG, and N²-AnthG. Two crystal structures of Dpo4 with an oligonucleotide containing an N²-NaphG adduct were determined. The complexes are identical in their composition but differ in the arrangement of the template N²-NaphG and in the pairing schemes. A kinetic model involving active and inactive DNA polymerase complexes was developed to fit the available data and is discussed in relation to the available structural data.

EXPERIMENTAL PROCEDURES

Materials—Dpo4 was expressed in E. coli and purified to electrophoretic homogeneity as described previously (46). Unlabeled dNTPs were obtained from Amersham Biosciences, (S)p-dCTPαS and (S)p-dTTPαS were purchased from Biolog Life Science Institute (Bremen, Germany), and [γ-³²P]ATP (specific activity 3 × 10⁸ Ci mmol⁻¹) was from PerkinElmer Life Sciences. T4 polynucleotide kinase and restriction endonucleases were purchased from New England Biolabs (Beverly, MA). Bio-spin columns were obtained from Bio-Rad (Hercules, CA).

Oligonucleotides—Unmodified 24-mer, 25-mer, and 36-mer (Table 1) were purchased from Midland Certified Reagent Co. (Midland, TX). Six 36-mers, each containing an N²-alkyl G adduct (N²-MeG, N²-EtG, N²-IbG, N²-BzG, N²-NaphG, or N²-AnthG, Fig. 1) were prepared as previously described and characterized by capillary gel electrophoresis and matrix-assisted laser desorption ionization/time-of-flight MS (24, 27–29, 48). The extinction coefficients for the oligonucleotides, estimated by the Borer method (49), were: 24-mer, ε₂₆₀ = 224 mm⁻¹ cm⁻¹; 25-mer, ε₂₆₀ = 232 mm⁻¹ cm⁻¹; and 36-mer, ε₂₆₀ = 310 mm⁻¹ cm⁻¹ (Table 1) (24, 27).

Reaction Conditions for Enzyme Assays and Analysis—Standard DNA polymerase reactions were carried out in 50 mM Tris-HCl buffer (pH 7.5 at 25 °C) containing 50 mM NaCl, 5 mM dithiothreitol, 100 μg of bovine serum albumin ml⁻¹, and 5% (v/v) glycerol at 37 °C (24, 27–29, 48). Primer (³²P-labeled at the 5'-end using T4 polynucleotide kinase/[γ-³²P]ATP) was annealed to either an unmodified or modified (N²-alkyl-G) template by heating the primer and template (equimolar mixture) to 95 °C for 5 min and then slowly cooling to room temperature. All reactions were initiated by the addition of dNTP solutions containing MgCl₂ (5 mM final concentration) to the preincubated enzyme/DNA mixtures. After reaction, 5-μl aliquots of the reaction mixture were quenched with EDTA-formamide solution (50 μl of 20 mM EDTA in 95% formamide (v/v) with 0.5% bromphenol blue (w/v) and 0.05% xylene cyanol (w/v)). Products were resolved using a 20% polyacrylamide (w/v) denaturing gel electrophoresis system containing 8 M urea and then visualized and quantitated by PhosphorImager analysis using a Bio-Rad Molecular Imager FX instrument and Quantity One software (Bio-Rad).

Primer Extension Assay with All Four dNTPs—Unmodified or modified DNA (100 nM) was incubated with Dpo4 (0, 0.5, 2,
or 10 nM) in 50 mM Tris-HCl buffer (pH 7.5 at 25 °C) containing 50 mM NaCl, 5 mM dithiothreitol, 100 ng of bovine serum albumin ml⁻¹, and 5% glycerol (v/v)). Reactions were initiated by adding dNTP-Mg²⁺ (100 μM each dNTP and 5 mM MgCl₂) solution at 37 °C. At the indicated time points, reactions were quenched by the addition of EDTA-formamide solution and analyzed by gel electrophoresis.

**Steady-state Kinetic Analyses**—A 32P-labeled primer/template complex was extended in the presence of varying concentrations of a single dNTP. The molar ratio of Dpo4 to DNA was <0.10. Polymerase concentrations and reaction time were adjusted to keep the extent of product formation <20% (50). All determinations were usually done at twelve dNTP concentrations (with the Km values in the center of the points). Reactions were quenched by the addition of EDTA-formamide solution, analyzed using gel electrophoresis, and then quantitated (phosphorimaging). Graphs of product formation versus dNTP concentration were fit using nonlinear regression (hyperbolic fits) using GraphPad Prism Version 3.0 (San Diego, CA) for the determination of kcat and Km values.

**Pre-steady-state Reactions**—Rapid quench experiments were performed using a model RQF-3 KinTek Quench Flow Apparatus (KinTek Corp., Austin, TX), with 50 mM Tris-HCl (pH 7.4, 25 °C) aqueous solution in the drive syringes. Reactions were initiated by rapid mixing of 32P-primer/template/Dpo4 mixtures (12.5 μM) with the dNTP-Mg²⁺ complex (10.9 μM) and were then quenched with 0.6 M EDTA after varying times. 20-μl aliquots of the reaction product solution were mixed with 100 μl of EDTA-formamide solution, and the products were separated using denaturing gel electrophoresis (see above) and quantitated. Pre-steady-state data, with excess DNA or Dpo4, were fit to Equations 1 or 2, respectively, using nonlinear regression analysis in GraphPad Prism Version 3.0, (San Diego, CA) for the determination of kₚₑₒₙ and Km values.

**Pre-steady-state Trap Experiments**—Pre-steady-state incorporation of dCTP opposite N²-AnthG was also analyzed in a RQF-3 KinTek Quench Flow Apparatus, with excess unlabeled trap DNA added along with dCTP. Reactions were initiated by rapid mixing of a 70 nM Dpo4 plus 120 nM 32P-labeled 24-mer primer/36-mer template (containing N²-AnthG) complex in 50 mM Tris-HCl buffer (pH 7.4) with a second solution containing 1 mM dCTP, 5 mM Mg²⁺, and 1.2 μM unlabeled trap 24-mer/36-mer in the same buffer. Reactions were quenched with 0.6 M EDTA after varying times. The reaction products were quantitated by denaturing gel electrophoresis and the data points were fit to Equation 1.

**Phosphorothioate Analysis**—Phosphorothioate analysis was done to compare the burst rates for the incorporation of dCTP and (S)p-dCTPαS opposite the template bases. Reactions were initiated by rapid mixing of 70 nM 32P-primer/template and 120 nM Dpo4 in 50 mM Tris-HCl buffer (pH 7.4, 12.5 μl) with a mixture of 1 mM dNTP or (S)p-dNTPαS and 5 mM MgCl₂ solution (10.9 μl) in the quench flow apparatus and were then quenched with 0.6 M EDTA after varying reaction times. The reaction products were quantified by denaturing gel electrophoresis and the data points were fit to Equation 1.

**Determination of kpol and Kₚₑₒₙ**—kpol and Kd,dCTP were estimated by performing pre-steady-state reactions in the quenched-flow apparatus with varying reaction times at different dNTP concentrations. Graphs of the burst rates (kₚₑₒₙ) versus dNTP concentration were fit to a hyperbolic equation,

\[
kₚₑₒₙ = kₚₑₒₙ[dNTP]/[dNTP] + Kₚₑₒₙ \quad \text{(Eq. 3)}
\]

where kₚₑₒₙ is the maximal rate of nucleotide incorporation and Kd,dCTP is the equilibrium dissociation constant for dCTP (51, 52).

**LC-MS/MS Analysis of Oligonucleotide Products from Dpo4 Reactions** (46, 53)—Dpo4 (5 μM) was preincubated with primer/template (10 μM), and the reaction was initiated by addition of dNTP (1 mM each) and MgCl₂ (5 mM) in 50 mM Tris-HCl buffer (pH 7.4) at 37 °C for 4 h. Reactions were terminated by extraction of the remaining dNTPs using a size-exclusion chromatography column (Bio-Spin 6 chromatography column, Bio-Rad). Concentrated stocks of Tris-HCl, dithiothreitol, and EDTA were added to restore the concentrations to 50 mM, 5 mM, and 1 mM, respectively. 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. 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 samples were resuspended in 100 μl of H₂O for MS analysis.

LC-MS/MS analysis was performed on an Acquity UPLC system (Waters, Milford, MA) connected to a Finnigan LTQ mass spectrometer (Thermo Fisher Scientific, Waltham, MA), operating in the ESI-negative ion mode. An Acquity UPLC BEH octadecylsilane (C₁₈) column (1.7 μm, 1.0 mm × 100 mm) was used with the following LC conditions: Buffer A contained 10 mM CH₃COONH₄ plus 2% CH₃CN (v/v), and buffer B contained 10 mM CH₃COONH₄ plus 95% CH₃CN (v/v). The following gradient program was used, with a flow rate of 150 μl min⁻¹: 0–2.5 min, linear gradient from 100% A to 95%/5% B (v/v); 2.5–6.0 min, linear gradient to 75%/25% B (v/v); 6–6.5 min, linear gradient to 100% B; 6.5–8.0 min, hold at 100% B; 8.0–9.0 min, linear gradient to 100% A; 9.0–12.0 min, hold at 100% A. The temperature of the column was maintained at 50 °C. Samples were injected with an autosampler system. ESI conditions were as follows: source voltage, 4 kV; source current, 100 μA; auxiliary 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. The doubly charged (negative ion) species were generally used for CID analysis. Calculations for 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.
Function of Dpo4 with N²-Guanyl Adducts

**TABLE 1**

| Oligodeoxynucleotides used in this study | 
|----------------------------------------|
| G¢·G, N²-MeG, N²-EtG, N²-IBG, N²-BzG, N²-NaphG, or N²-AnthG, Cₛᵦ terminal dioxyacetide. |

Kinetic Simulations—Kinetic simulations were performed using two mechanisms: a minimal mechanism, which has been previously defined for DNA polymerization (54, 55), and a mechanism with the presence of a non-productive ternary complex (56, 57). The fits to the experimentally determined data were developed utilizing DynaFit (BioKin, Watertown, MA) (58), run on an iMac computer using a Macintosh OS 8.5.1 operating system (Apple Computer, Cupertino, CA).

Crystallographic and X-ray Diffraction Data Collection—The 18-mer containing N²-NaphG and 14-mer primer containing a 3′-terminal dioxyacetide (Table 1) were annealed to form a duplex. Dpo4 was mixed with DNA (1:1.2 molar ratio) in 20 mM Tris-HCl buffer (pH 8.0, 25 °C) containing 60 mM NaCl, 4% glycerol (v/v), and 5 mM MgCl₂ and 1 mM dGTP. The final Dpo4 concentration was 18-mer template containing 60 mM NaCl, 4% glycerol (v/v), and 5 mM mercaptoethanol and then placed on ice for 1 h prior to incubation with 5 mM MgCl₂ and 1 mM dGTP. The final Dpo4 concentration was ~10 mg ml⁻¹. Crystals were grown using the sitting drop/vapor-diffusion method with the reservoir solution containing 20 mM Tris-HCl (pH 8.0 at 25 °C), 15% polyethylene glycol 3350 (w/v), 60 mM NaCl, 5 mM MgCl₂, and 4% glycerol (v/v). Droplets consisted of a 1:1 (v/v) mixture of the Dpo4-DNA:Mg²⁺·dGTP complex and the reservoir solutions and were equilibrated against the reservoir solutions. 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-N²-NaphG-1 (Npg-1) and Dpo4-N²-NaphG-2 (Npg-2) complex crystals were collected at the Advanced Photon Source (Argonne National Laboratory, Argonne, IL) on the 21-ID (LS-CAT) and 22-ID (SER-CAT) beam lines, respectively. Both data sets were recorded from 25-mer template residue 5, metal ions, and dGTP served as the starting model for the Npg-2 structure, and the locations of the individual models were optimized by several rounds of rigid body refinement while gradually increasing the resolution of the diffraction data.

Manual model rebuilding was done with the program TURBO-FRODO. The maps were computed using the sA-modified coefficients (62). Clear positive density for the Mg²⁺ ions and the dGTP was observed in the initial difference Fourier electron density maps of both complexes, although the positive density was sparser for dGTP in the Npg-1 structure. Positive density was observed for the N²-NaphG-modified template residue, but several rounds of modeling and refinement were required to ascertain the final orientation of the base and the naphthalene moiety in both structures. The CNS package (63) 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 2). The crystallographic figures were prepared using PyMOL. Calculation of the DNA helical parameter descriptions was performed using CURVES (64).

RESULTS

Primer Extension by Dpo4 Using All Four dNTPs—Full-length extension beyond various N²-alkyl G adducts in the presence of all four dNTPs by Dpo4 is shown in Fig. 2. An increase in the Dpo4 concentration enhanced extension for all DNA adducts examined. Dpo4 readily extended beyond the unmodified G and N²-MeG, N²-EtG, N²-IBG, and N²-BzG, and N²-NaphG templates to yield 35-, 36-, and 37-mer products, although the extensions were partially blocked opposite these adducts to yield (unelongated) 25-mer product. Other products (28-, 29-, and 31-mers) were also produced. With N²-AnthG, only one-base extension product was observed, and subsequent extension was severely blocked under these conditions.

Steady-state kinetics of dNTP Incorporation by Dpo4 Opposite G and N²-Alkyl G Adducts—Steady-state kinetic parameters (kₘᵦ and Kₘ) were measured for the single incorporation of each dNTP opposite G and N²-alkyl G adducts, with varying concentrations of dNTP (Table 3). Dpo4 preferentially incorporated dCTP opposite all of the N²-alkyl G adducts. However, the efficiencies (kₘⁱ⁄Kₘ) for correct incorporation opposite N²-alkyl G adducts were decreased 3- to 125-fold compared with unmodified G, due to both the decreased catalytic rates (kₘⁱ) and increased Kₘ values. The efficiencies for the misincorporation of dATP, dGTP, and dTTP opposite all the N²-alkyl G adducts were similar to those measured with unmodified G.

Correct incorporation of dCTP opposite unmodified G was preferred over misincorporation for Dpo4, yielding a misincorporation frequency (defined as f = (kₘᵦ/Kₘ)ᵈC₅P/(kₘⁱ/Kₘ)ᵈC₅P, where dNTP ≠ dCTP) of 10⁻⁵. The misincorporation frequency was 10⁻⁴ to 10⁻³ for all the N²-guanylated adducts (10⁻⁴ to 10⁻³-fold increase), mainly due to decreased correct incorporation efficiency (as opposed to changes in the misincorporation efficiency).

4 C. Cambillau and A. Roussel (1997) Turbo Frodo, Version Open GL.1, Université Aix-Marseille II, Marseille, France.

5 W. L. DeLano (2002) The PyMOL Molecular Graphics System, DeLano Scientific LLC, San Carlos, CA.
Steady-state Next-base Extension following C or T Paired with G or N²-Alkyl G Adducts—Steady-state next-base extension (i.e. incorporation of dGTP opposite the next template base, C

TABLE 2

| Parameter | Npg-1 | Npg-2 |
|-----------|-------|-------|
| X-ray source | APS (LS-CAT) | APS (SER-CAT) |
| Beam line | ID-21 | ID-22 |
| Detector | MAR CCD | MAR CCD |
| Wavelength (Å) | 0.98 | 0.99 |
| Temperature (K) | 110 | 110 |
| Number of crystals | 1 | 1 |
| Space group | P2₁,2,2 | P2₁,2,2 |
| Unit cell (Å) | 95.04, 103.91, 52.32 | 95.92, 103.83, 52.49 |
| Resolution range (Å) | 19.97–3.10 | 29.25–3.0 |
| Highest resolution shell | (3.19–3.10) | (3.19–3.0) |
| Number of measurements | 63,806 (6,312) | 78,594 (1,746) |
| Number of unique reflections | 9,733 (912) | 11,751 (1,607) |
| Redundancy | 6.5 (6.9) | 6.6 (7.3) |
| Completeness (%) | 98.4 (95.1) | 98.5 (97.1) |
| Rmerge | 9.4 (64.9) | 7.8 (60.7) |
| Bond angles (°) | 1.6 | 1.6 |
| Dihedral angles (°) | 22.8 | 22.2 |
| Improper angles (°) | 1.96 | 1.77 |
| Rfree (%) | 23.3 | 22.6 |
| Rfree (%) | 28.2 | 28.2 |
| Estimated coordinate error (Å) | 0.45 | 0.40 |
| Temperature factors | 62.4 | 57.3 |
| Mean isotropic (Å²) | 88.9 | 77.4 |
| r.m.s.d. in temperature factors | 1.25 | 1.50 |
| Bonded main-chain atoms (Å²) | 1.53 | 2.25 |
| r.m.s.d. from ideal values | 0.009 | 0.009 |
| Bond lengths (Å) | 1.6 | 1.6 |
| Dihedral angles (°) | 22.8 | 22.2 |
| Improper angles (°) | 1.96 | 1.77 |

a Values in parentheses correspond to the highest resolution shells.

b Rmerge = ΣhkIhk − ΣhkIhk|hk|, where the outer sum (hk) is taken over the unique reflections.

c Rfree = ΣhkIhk − ΣhkIhk|hk|, where |hk| and |hk| are the observed and calculated structure factor amplitudes, respectively.

d Rfree, idem, for the set of reflections (5% of the total) omitted from the refinement process.

Value differences were decreased 4- to 32-fold compared with the unmodified G.

Pre-steady-state Burst Kinetics of dCTP Incorporation Opposite G or N²-Alkyl G Adducts by Dpo4—Dpo4 clearly showed burst phases for the correct incorporation of dCTP opposite unmodified G or N²-al kyl G adducts, even with the bulky adduct N²-AnthG, and burst rates were also attenuated compared with the unmodified G.

Steady-state and Pre-steady-state Burst Kinetics of dTTP Incorporation Opposite G and N²-Alkyl G Adducts by Dpo4—No bursts were detected for the incorporation of dTTP opposite G or any N²-al kyl G adducts, except for N²-MeEG (Fig. 4A). The incorrect incorporation of dTTP opposite N²-AnthG yielded a higher misincorporation rate (Fig. 4A). Misincorporation of both dTTP and (S₉)₉-dTTPαS opposite N²-MeEG showed small bursts (Fig. 4B), with much lower burst amplitudes (~3% of the enzyme concentration), burst rates, and steady-state rates.

Determination of kpol and Kd,dCTP Values for dCTP Incorporation by Dpo4—Analysis of the changes in the pre-steady-state burst rates as a function of increasing dTTP concentration (supplemental Fig. S1) yielded maximum rates of nucleotide incorporation (kpol) and apparent binding affinity of the dCTP to the Dpo4 DNA binary complex to form a ternary complex poised for catalysis (Kd,dCTP) (51, 52). The incorporation of dCTP opposite unmodified G or N²-MeEG showed similar kpol (~3.0 s⁻¹) and Kd,dCTP (~10 μM) values (Fig. 5, A and B). For the incorporation opposite bulky N²-al kyl G adducts (N²-EtG to N²-NaphG, Fig. 5, C–F) the kpol values were decreased 4- to 32-fold compared with G, but the Kd,dCTP value for N²-AnthG (Fig. 5G) increased unexpectedly (compared with N²-NaphG). The Kd,dCTP values for...
**Function of Dpo4 with N²-Guanyl Adducts**

### Table 3

Steady-state kinetic parameters for one-base incorporation opposite G and N²-alkyl G adducts by Dpo4

| Template  | dNTP | $k_{m\text{dNTP}}$ | $k_{\text{cat}} \times 10^3$ | $k_{\text{cat}}/K_m \times 10^5$ | Efficiency relative to dCTP:G | Misincorporation frequency |
|-----------|------|------------------|-----------------|-----------------|-----------------------------|---------------------------|
| G         | C    | $1.2 \geq 0.3$  | $960 \geq 50$  | $8.2 \times 10^4$ | 1                           | 1.4 $\times 10^{-5}$     |
|           | A    | $490 \geq 60$   | $5.7 \geq 0.2$ | $1.2 \times 10^{-5}$ |                            |                           |
|           | G    | $190 \geq 20$   | $6.3 \geq 0.1$ | $3.3 \times 10^{-5}$ |                            |                           |
|           | T    | $780 \geq 10$   | $34 \geq 2$    | $4.4 \times 10^{-5}$ |                            |                           |
| N²-MeG    | C    | $1.9 \geq 0.5$  | $500 \geq 21$  | $2.7 \times 10^4$   | 3-fold less                 |                           |
|           | A    | $650 \geq 70$   | $7.6 \geq 0.3$ | $1.2 \times 10^{-5}$ |                            |                           |
|           | G    | $380 \geq 80$   | $23 \geq 2$    | $6.1 \times 10^{-5}$ |                            |                           |
|           | T    | $320 \geq 90$   | $40 \geq 5$    | $12 \times 10^{-5}$  |                            |                           |
| N²-EtG    | C    | $7.4 \geq 1.0$  | $83 \geq 3$    | $1100 \times 10^{-5}$ | 75-fold less                |                           |
|           | A    | $360 \geq 60$   | $3.4 \geq 0.2$ | $0.93 \times 10^{-5}$ |                            |                           |
|           | G    | $210 \geq 30$   | $5.5 \geq 0.3$ | $2.6 \times 10^{-5}$  |                            |                           |
|           | T    | $78 \geq 18$    | $2.6 \geq 0.1$ | $3.4 \times 10^{-5}$  |                            |                           |
| N²-IbG    | C    | $16 \geq 3$     | $100 \geq 5$   | $650 \times 10^{-5}$ | 126-fold less               |                           |
|           | A    | $280 \geq 50$   | $3.0 \geq 0.2$ | $1.1 \times 10^{-5}$  |                            |                           |
|           | G    | $170 \geq 10$   | $8.8 \geq 0.2$ | $5.2 \times 10^{-5}$  |                            |                           |
|           | T    | $140 \geq 10$   | $1.6 \geq 0.0$ | $1.2 \times 10^{-5}$  |                            |                           |
| N²-BzG    | C    | $21 \geq 4$     | $270 \geq 20$  | $1300 \times 10^{-5}$ | 63-fold less                |                           |
|           | A    | $190 \geq 30$   | $2.1 \geq 0.1$ | $1.1 \times 10^{-5}$  |                            |                           |
|           | G    | $120 \geq 20$   | $2.2 \geq 0.1$ | $1.8 \times 10^{-5}$  |                            |                           |
|           | T    | $320 \geq 53$   | $4.4 \geq 0.2$ | $1.4 \times 10^{-5}$  |                            |                           |
| N²-NaphG  | C    | $5.2 \geq 1.0$  | $110 \geq 3$   | $2200 \times 10^{-5}$ | 37-fold less                |                           |
|           | A    | $320 \geq 60$   | $1.0 \geq 0.1$ | $0.33 \times 10^{-5}$ |                            |                           |
|           | G    | $180 \geq 20$   | $3.8 \geq 0.1$ | $2.2 \times 10^{-5}$  |                            |                           |
|           | T    | $120 \geq 20$   | $1.5 \geq 0.1$ | $1.3 \times 10^{-5}$  |                            |                           |
| N²-AnthG  | C    | $1.3 \geq 0.15$ | $43 \geq 1$    | $3200 \times 10^{-5}$ | 26-fold less                |                           |
|           | A    | $360 \geq 450$  | $0.78 \geq 0.03$ | $0.21 \times 10^{-5}$ |                            |                           |
|           | G    | $870 \geq 140$  | $6.5 \geq 0.5$ | $0.75 \times 10^{-5}$ |                            |                           |
|           | T    | $100 \geq 10$   | $2.3 \geq 0.05$ | $2.3 \times 10^{-5}$  |                            |                           |

all the N²-alkyl G adducts (8 to 29 μM) were similar to that of unmodified G (10 μM) (3-fold variation compared with G).

**Phosphorothioate Analysis of dCTP Incorporation Opposite G and N²-Alkyl G Adducts by Dpo4**—Incorporation of dCTP or (S<sub>p</sub>)-dCTP<sub>10</sub> opposite G, N²-MeG, and N²-EtG (Fig. 3, A–C) showed no obvious thio effects (burst rate ratios of 1.0, 1.1, and 1.4, respectively). In contrast, the incorporation of dCTP or (S<sub>p</sub>)-dCTP<sub>10</sub> opposite N²-IbG, N²-BzG, or N²-NaphG (Fig. 3, D–F) resulted in different rates and thio effects (ratios higher than 13, 25, and 15, respectively). Unexpectedly, incorporation of (S<sub>p</sub>)-dCTP<sub>10</sub> opposite N²-AnthG (Fig. 3G) showed a similar burst rate as in the incorporation of dCTP, but without an obvious thio effect.

**Determination of $k_{\text{pol}}$ and $K_{d\text{dCTP}}$ for dTTP Incorporation Opposite N²-MeG by Dpo4**—The incorporation of dTTP opposite N²-MeG also showed a burst phase, and the corresponding $k_{\text{pol}}$ and $K_{d\text{dCTP}}$ were determined to be 0.95 (± 0.06) s<sup>-1</sup> and 132 (± 25) μM, respectively (supplemental Fig. S2). Compared with correct dCTP incorporation (Fig. 4B), dTTP incorporation had a decreased amplitude (23-fold), decreased burst rate (3-fold), and weakened dNTP binding affinity (15-fold).

**Pre-steady-state Incorporation of dCTP Opposite N²-AnthG with Excess Trap DNA**—A 10-fold molar excess of unlabeled unmodified DNA was used to trap free Dpo4 after Dpo4 dissociated from the 32P-labeled DNA containing N²-AnthG; this experiment was done to establish whether a burst was indeed observed, because only Dpo4 that was still bound to the labeled DNA could extend the (radiolabeled) primer during the single binding event (24). Incorporation of dCTP opposite N²-AnthG with or without trap DNA showed a burst phase, and the burst amplitudes were similar (Fig. 6). A slow linear phase is observed in the trap experiment, most likely due to the relatively low burst in product formation (~17% of [Dpo4]) and the fact that the protein trap is not 100% efficient. The major point of the trapping experiment was to be certain that the small burst amplitude observed with the N²-AnthG-containing substrate is representative of true pre-steady-state conditions.

**Analysis of Primer Incorporation/Extension Products using LC-MS/MS**—Full-length extension assays showed that all of the adducts, except for N²-AnthG, can be readily extended (Fig. 2). LC-MS/MS was utilized for the sequence analysis of the oligonucleotide extension products (46, 67). A key element for the success of this method is the placement of uracil residues in the primer strand and the use of uracil DNA glycosylase to hydrolyze at the introduced uracil sites, to reduce the product length for MS sequence analysis (46). Each product in the oligonucleotide product mixtures could be readily identified and quantified by LC-MS/MS (Table 5, supplemental Figs. S3, S4, and S5, and Tables S1–S4). From unmodified G to the
values (9.4% and 7.8%, for Npg-1 and Npg-2, respectively). Both crystals belong to the space group P2_1_2_1, with a single complex per asymmetric unit.

The two structures are very similar overall (Fig. 7A). Analysis of the helical parameters of the DNA substrates reveals perturbations that are observed in both complexes. Similar to results obtained with O^6^-BzG-modified templates (42), each of the two refined structures containing N^2^-NaphG possesses a distinct arrangement of the DNA in the active site. The naphthyl moiety exhibits well defined electron density for both orientations (Fig. 7A). A simulated annealing omit map (excluding template N^2^-NaphG, primer terminal ddC, incoming dGTP, and active site Mg2^+ ion) further supports the correct positioning of the residues bound in the active site of both complexes (supplemental Fig. S6). The N^2^-NaphG residue is found in the syn orientation in the Npg-1 complex, which allows hydrogen bonding to occur between the exocyclic amino groups of the last two cytosines of the primer with the O6 atom of N^2^-NaphG (Fig. 7C). The 3'-phosphate of the incoming dGTP is positioned ~5 Å from the primer terminus, and two of the Mg2^+ ions are coordinated in the active site of the enzyme by residues Asp-7, Asp-105, Glu-106, and Asp-156. Overall, Npg-1 would appear to be an “active” complex.

The N^2^-NaphG residue is found in the anti orientation in the Npg-2 complex but the terminal C of the primer is largely disordered. Based on the density surrounding the terminal phosphate and sparse density in the region of the base, the deoxycytosine was flipped into the minor groove, as has been observed in other non-productive Dpo4 structures (38, 42).

The N^2^-NaphG residue is found in the anti orientation in the Npg-2 complex. The terminal deoxycytosine of the primer is flipped into the minor groove, based on the density surrounding the terminal phosphate and sparse density in the region of the pyrimidine ring system. Such a conformation has been observed in other non-productive Dpo4 structures (38, 42).
**Function of Dpo4 with $N^2$-Guanyl Adducts**

There is considerable distortion of the DNA duplex in both Npg-1 and Npg-2 structures when compared with the structure with a DNA lesion that replicated with high efficiency (dCTP opposite 8-oxoG by Dpo4; pdb ID code 2c2e; Fig. 7B). In both instances the base pairs to the 3’-side of the $N^2$-NaphG residue are severely buckled (35° and 24° for Npg-1 and Npg-2, respectively), with large tilts for the cytosine at the −2 position ($t = 27°$ and $t = 19°$ for Npg-1 and Npg-2, respectively). Furthermore, the base pairing distance between hydrogen bond donors/acceptors is increased several residues upstream from $N^2$-NaphG, indicating that the lesion can distort the helix several base pairs away from the actual site of damage.

**Kinetic Simulation to a Mechanism**—A general minimal mechanism is shown in Fig. 8A (51, 54, 55, 68–72), modified with the addition of step 8 (56, 66). The binding affinity and a burst rate for the incorporation of dCTP opposite G and $N^2$-alkyl G adducts by Dpo4 were initially set at 15 μM and 3 s$^{-1}$, respectively, based on the pre-steady-state kinetic determinations (Fig. 5). Recently, our laboratory utilized mutant Dpo4 containing Trp as a fluorescent probe to rigorously examine the kinetics describing Dpo4-catalyzed polymerization (66). It was concluded that a slow conformational relaxation following pyrophosphate release is the major rate-limiting determinant in the Dpo4 catalytic cycle. The kinetic mechanism used in our modeling studies reflects these new insights into Dpo4 catalysis (66).

The minimal mechanism fit generally well for the DNA polymerization with a stoichiometric burst amplitude. In this study, the incorporation of dCTP opposite G and $N^2$-alkyl G adducts (except $N^2$-MeG) by Dpo4 resulted in reduced burst amplitudes (Fig. 5). Simulation of experimental results for partial bursts cannot yield a reasonable fit for G itself or the other

---

**TABLE 5**

Products of extension of $N^2$-alkylG template-primer complexes by Dpo4

| Compound         | AGG AGG AGG AGG | AGG AGG AGG AGG | AGG AGG AGG AGG |
|------------------|------------------|------------------|------------------|
| $N^2$-MeG        | AGG AGG AGG AGG | AGG AGG AGG AGG | AGG AGG AGG AGG |
| $N^2$-EtG        | AGG AGG AGG AGG | AGG AGG AGG AGG | AGG AGG AGG AGG |
| $N^2$-BzG        | AGG AGG AGG AGG | AGG AGG AGG AGG | AGG AGG AGG AGG |
| $N^2$-NaphG      | AGG AGG AGG AGG | AGG AGG AGG AGG | AGG AGG AGG AGG |
| $N^2$-AnthG      | AGG AGG AGG AGG | AGG AGG AGG AGG | AGG AGG AGG AGG |

---

**FIGURE 5.** Estimation of $k_{pol}$ and $K_{uCTP}$ for Dpo4 by pre-steady-state burst rate dependence on dCTP concentration. Dpo4 (200 nM) was incubated with 100 nM 24-mer/36-mer primer-template complexes ($^{32}$P-labeled) in a rapid quench-flow instrument and mixed with varying dCTP concentrations (2–300 μM) to initiate reactions. Plots of burst rates ($k_{obs}$, fit by Equation 2) versus [dCTP] were fit to a hyperbolic equation (Equation 3). A, G (unmodified); B, $N^2$-MeG; C, $N^2$-EtG; D, $N^2$-BzG; E, $N^2$-NaphG; and G, $N^2$-AnthG. The calculated $k_{pol}$ and $K_{uCTP}$ values are shown in each panel of the figure.

**FIGURE 6.** Pre-steady-state kinetics of incorporation of dCTP opposite $N^2$-AnthG by Dpo4 in the absence and presence of excess unlabeled DNA trap. Dpo4 (70 nM) was incubated with 24-mer/36-mer primer-template complex containing $N^2$-AnthG (120 nM) in a rapid quench-flow instrument and mixed with a mixture of either (i) 5 mM MgCl$_2$/1 mM dCTP or (ii) 5 mM MgCl$_2$/1 mM dCTP/1200 nM unlabeled trap 24-mer/36-mer DNA to initiate reactions. All polymerization reactions were quenched with 0.6 M EDTA at the indicated time intervals. The data were fit to the burst equation, $y = A(1 - \exp(-k_{pol}t)) + k_{obs}t$ (Equation 1). The burst rates were estimated to be 0.46 (± 0.04) s$^{-1}$ for the case with dCTP only and 1.7 (± 0.2) s$^{-1}$ for the case with the DNA trap.

---

There is considerable distortion of the DNA duplex in both Npg-1 and Npg-2 structures when compared with the structure with a DNA lesion that replicated with high efficiency (dCTP opposite 8-oxoG by Dpo4; pdb ID code 2c2e; Fig. 7B). In both instances the base pairs to the 3’-side of the $N^2$-NaphG residue are severely buckled (35° and 24° for Npg-1 and Npg-2, respectively), with large tilts for the cytosine at the −2 position ($t = 27°$ and $t = 19°$ for Npg-1 and Npg-2, respectively). Furthermore, the base pairing distance between hydrogen bond donors/acceptors is increased several residues upstream from $N^2$-NaphG, indicating that the lesion can distort the helix several base pairs away from the actual site of damage.

**Kinetic Simulation to a Mechanism**—A general minimal mechanism is shown in Fig. 8A (51, 54, 55, 68–72), modified with the addition of step 8 (56, 66). The binding affinity and a burst rate for the incorporation of dCTP opposite G and $N^2$-alkyl G adducts by Dpo4 were initially set at 15 μM and 3 s$^{-1}$, respectively, based on the pre-steady-state kinetic determinations (Fig. 5). Recently, our laboratory utilized mutant Dpo4 containing Trp as a fluorescent probe to rigorously examine the kinetics describing Dpo4-catalyzed polymerization (66). It was concluded that a slow conformational relaxation following pyrophosphate release is the major rate-limiting determinant in the Dpo4 catalytic cycle. The kinetic mechanism used in our modeling studies reflects these new insights into Dpo4 catalysis (66).

The minimal mechanism fit generally well for the DNA polymerization with a stoichiometric burst amplitude. In this study, the incorporation of dCTP opposite G and $N^2$-alkyl G adducts (except $N^2$-MeG) by Dpo4 resulted in reduced burst amplitudes (Fig. 5). Simulation of experimental results for partial bursts cannot yield a reasonable fit for G itself or the other
Function of Dpo4 with N²-Guanyl Adducts

**DISCUSSION**

Although the Y-family DNA polymerase Dpo4 has been used extensively as a model for translesion polymerization, systematic studies on its ability to incorporate past lesions of different sizes have not been done previously. We had a set of oligonucleotides containing guanines in the N2-position of N2-guanyl adducts differing only in their N2-substitution at a single site available and used these in this study with Dpo4. As in the cases of the human Y-family polymerases we have analyzed (24, 27–29, 48), the catalytic efficiency (steady-state $k_{cat}/K_m$) showed relatively limited attenuation with increasing size of the lesion at the G N2 atom, in contrast to replicative DNA polymerases (e.g. human pol δ, HIV-1 reverse transcriptase, and bacteriophage pol T7) (24). Very recently RNA polymerases have been found to vary in their abilities to insert and copy past N²-guanylation (16). With Dpo4 and this series of N²-guanyl derivatives, the fidelity showed only a limited decrease until the size increased to that of N²-anthracene guanines (Table 5). Of interest was the maintenance of a pre-steady-state kinetic burst throughout the entire series of N²-guanine substitutions, with a decrease in the burst amplitude. The $K_{dCTP}$ did not change very much with increasing bulk at the guanyl N2 atom.

Two crystal structures of Dpo4 with template DNA containing N²-guanyl were solved. In both cases, the N²-guanyl adduct was positioned opposite deoxythymidine at the 3'-end of a primer (at the −1 position, i.e. post-insertion), but two different active site orientations were observed. The N²-guanyl moiety adopts the syn orientation in the Npg-1 structure and the anti conformation in the active site of Npg-2. The DNA duplex suffers some perturbation in both structures, including a slight widening of the major groove near the primer-template junction, as are evident. C, the active site of the Npg-1 crystal structure is shown in schematic form. The $3F_{o} - 2F_{c}$ electron density map (cyan mesh) is contoured at the 1σ level around the N²-guanyl residue. The incoming dGTP (red) is shown along with the magnesium ions (green spheres). D, base pairing and stacking interactions observed in the Npg-1 structure. E, the active site of the Npg-1 crystal structure is shown in schematic form. The $3F_{o} - 2F_{c}$ electron density map (purple mesh) is contoured at the 1σ level around the N²-guanyl residue. The incoming dGTP (red) is shown along with the magnesium ions (green spheres). F, base pairing and stacking interactions observed in the Npg-2 structure.

N²-alkyl G adducts using the minimal mechanism (except for N²-MeG), as expected (56), even with adjustment of the rates of steps after phosphodiester bond formation (step 4 in Fig. 8A). An alternate mechanism with an additional non-productive E-DNA-dNTP ternary complex (56, 66) (Fig. 8A) yielded good fits to the experimental results for the incorporation of dCTP opposite G (Fig. 8B) and all of the N²-alkyl G adducts (except for N²-MeG, Fig. 8C), using the experimentally measured enzyme concentration (Fig. 8, C–H). The same rate constants shown in the scheme of Fig. 8A were used in the simulations to generate values of $k_4$, $k_{\text{on}}$, and $k_{\text{off}}$ (Table 6), and the ratios for the conver-

deasing burst amplitude. The $K_{dCTP}$ did not change very much with increasing bulk at the guanyl N2 atom.

Two crystal structures of Dpo4 with template DNA containing N²-guanyl were solved. In both cases, the N²-guanyl adduct was positioned opposite deoxythymidine at the 3'-end of a primer (at the −1 position, i.e. post-insertion), but two different active site orientations were observed. The N²-guanyl moiety adopts the syn orientation in the Npg-1 structure and the anti conformation in the active site of Npg-2. The DNA duplex suffers some perturbation in both structures, including a slight widening of the major groove near the primer-template juncti...
junction and buckling of the base pairs upstream from the adduct site. Of the two structures, the Npg-1 complex appears to be productive, whereas the Npg-2 complex is considered non-productive.

In the structure Npg-2, the anti form of N2-NaphG has the primer 14-C flipped out of the active site and distorted (Fig. 7F), indicating that incorporation of dCTP opposite the bulkier N2-guanyl adducts will be very unfavorable, with a tendency to form less productive complexes. In comparison to the Watson-Crick base pairing presumably used for dCTP with template G and small N2-alkyl G adducts, the fraction of non-productive complex(es) is presumably increased with N2-guanyl adduct bulkiness, and incorporation burst amplitudes are decreased (Fig. 3). Kinetic modeling also yielded an increase in the residence time in the non-productive complex with N2-adduct bulkiness (Table 6 and Fig. 8). The bulkier N2-NaphG can also be in a syn configuration (Npg-1), which we conclude can incorporate dCTP opposite the Hoogsteen face of a modified G (Fig. 7). Thus the N2-alkyl G adducts may be incorporating dCTP via either of two (or more) mechanisms: with the small N2-alkyl G adducts, mainly via the usual anti configuration/Watson-Crick pairing mode; for the bulkier N2-aralkyl G adducts, via a syn/Hoogsteen mode. Changes in the use of these mechanisms with increasing bulkiness of the adducts may be reflected in the variations of the observed kinetic parameters (Fig. 3).

Of the two structures solved here, the complex that contains syn N2-NaphG in a Hoogsteen-type pairing with ddC appears to be the more productive. However, the use of a syn-oriented template base would not be expected to yield the level of fidelity observed for Dpo4 (Table 3). The Y-family member pol ϊ has been shown to rely upon Hoogsteen base pairing even with unmodified DNA (73, 74). In biochemical terms, pol ϊ has an
unusual incorporation spectrum in that it incorporates opposite template dA with much greater efficiency and fidelity than other bases (75). When incorporating opposite template dG, pol α does not possess the level of fidelity observed for any of the N²-alkyl G adducts tested with Dpo4. Crystal structures of pol α have shown that the Hoogsteen pairing mode may arise, because the template sugar is embedded in a “cavity” formed primarily by residues from the finger domain (specifically Leu-62, Lys-60, and Gln-59), which shortens the C1’–C1’ distance from 10.65 Å (anti) to 8.6 Å (syn) (74). Such a shortening of the C1’–C1’ distance is not observed in the Npg-1 structure, but a direct comparison cannot be made because the syn-oriented G in the pol α structure is in the insertion context, while syn N²-NaphG is in the post-insertion context of Dpo4. Nevertheless, some differences in this region are worth noting. Of the three residues postulated to influence the C1’–C1’ distance in pol α, only Gln-59 is distinctive from the Val-32 of Dpo4, because the amide side chain of Gln-59 protrudes into the region of the active site that contains the template base. The Lys-60 of pol α contacts the phosphate backbone of the template DNA, whereas the analogous residue in Dpo4 (Phe-33) marks the beginning of a loop between β-sheets 2 and 3 in the finger domain. The loop is comprised of eight amino acids residues in Dpo4, which contrast to the short three-residue β-hairpin observed in pol α. Previous work has shown that pol α is not particularly tolerant of bulk at the N2 atom of guanine and suffers a greater loss of fidelity during bypass of N²-dG adducts compared with either Dpo4 or pol κ (28, 48). The fidelity observed for pol α bypass of N²-alkyl G adducts is ~100-fold more error-prone than the results obtained with Dpo4. Based on the crystal structures reported here and previous kinetic analysis it appears that the molecular features resulting in Dpo4-catalyzed bypass of N²-alkyl G adducts are not homologous to the ones that influence the ability of pol α to preferentially utilize Hoogsteen pairs. Comparison to pol α suggests that perhaps we have not captured the true active complex representing Dpo4-catalyzed bypass of N²-NaphG.

The presence of two substrate orientations in the Dpo4 active site is reminiscent of a study in which the trans N²-(105) BPDE-G adduct was observed in two different orientations when bound to Dpo4 (45). Modeling studies of Dpo4-catalyzed bypass of a bulky N²-BPDE-G adduct had predicted that the adduct could be accommodated in both syn and anti orientations during insertion opposite the lesion (76). The resulting orientations observed in the N²-BPDE-G crystal structures were distinct from the models, but it should be noted here that the modeling study focused on insertion, whereas in the crystal structures the N²-BPDE-G adduct is in the post-insertion context. In the “productive” crystal structure (BPG-1B) (45), the N²-BPDE-G moiety is flipped into a gap between the finger and little finger domains, which was not predicted by molecular modeling. The BPG-1B complex of Bauer et al. (45) appears to be associated with −1 frameshift events, as judged by the available gel data (i.e. the predominant dNTP incorporation is the one corresponding to the residue 5’ of the BPDE adduct in the template). These events would argue that the “Type II” crystal structure is most relevant to the observed function during bypass of N²-BPDE-G. No quantitative kinetic results were reported, so comparisons to the N²-guanyl adducts used here are not possible. The syn orientation of the N²-BPDE-G adduct in the Dpo4 active site results in a complex that is non-productive, because the BPDE ring system is stacked between the last two bases in the primer and this distortion prevents optimal binding of the incoming dATP. By way of comparison, Dpo4 places the naphthyl group between the template DNA and the little finger domain when the N²-NaphG moiety in the syn orientation, not into the “gap” observed for the BPDE adduct (supplemental Fig. S6). With the placement of the N²-NaphG moiety in the syn orientation, Dpo4 can achieve what appears to be a productive post-insertion complex. However, there is significant distortion to the base pairs adjacent to the lesion and the major groove is widened by ~3–4 Å starting three base pairs upstream of the adduct site. Superimposition of the Npg-1 complex with BPG-1B (45) reveals that, in contrast to BP-dG, the purine ring of N²-NaphG remains stacked in the template strand, perhaps due to the added flexibility of a methylene group between the exocyclic amino group and the naphthalene moiety (supplemental Fig. S6). The major similarity between the N²-NaphG structures solved here and the N²-BPDE-G structures is the ability of Dpo4 to allow several substrate conformations without greatly altering the polymerase structure.

It is of interest to compare the Npg-1 and Npg-2 structures solved here with the reported structure of pol κ bound to unmodified DNA (77), because Dpo4 and pol κ are both members of the DinB subfamily and can be considered homologues. Superimposition of the complexes reveals that, not surprisingly, the enzymes are quite similar in domain arrangement and orientation of the DNA substrate (Fig. 9A). Closer inspection of the polymerase active site shows that pol κ can clearly accommodate the N²-NaphG adduct in the syn orientation (Fig. 9B). Furthermore, the N-terminal portion of pol κ, which forms a “clasp” that encircles the DNA, has several aromatic side chains that could in theory promote the syn orientation of aromatic N²-alkyl G adducts during translesion synthesis by way of π–π interactions. Such a hypothesis seems somewhat counter-intuitive. One might initially predict (as we did) that minor groove adducts could be housed in the larger active sites of Y-family DNA polymerases, thereby allowing normal Watson-Crick geometry to occur during bypass. A computational study of pol κ insertion opposite N²-BPDE-dG and N²-BPDE-dA adducts predicted that the enzyme can effectively bypass N²-BPDE-dG in the anti orientation (78). The same study postulates that the N-clasp is important for “near error-free” bypass of N²-BPDE-dG and blockage of pol κ by the N²-BPDE-dA adduct. The authors propose that the N-clasp promotes the anti orientation during accurate bypass of N²-BPDE-dG. Even though a productive configuration is not observed in the Npg-2 structure, we cannot rule out the possibility that accurate bypass of the larger N²-alkyl G adducts proceeds through some orientation that is closer to normal Watson-Crick geometry than the pairing mode observed in the Npg-1 complex. Because the N-clasp is unique to pol κ, it may play a role in promoting accurate bypass of minor groove lesions. Indeed, several lines of evidence are consistent with the idea that human pol κ is an important enzyme during transle-
Direction of Dpo4 with N²-Guanyl Adducts

Dpo4 appears to favor accurate bypass of the bulky adducts over −1 frameshift deletions. Analysis of the crystal structures suggests the possibility that, as the bulk at the N2 position increases, accurate bypass of the N²-NaphG adduct occurs when the lesion adopts the syn orientation. Still, it is difficult to explain the limited decrease in Dpo4 fidelity observed with the N²-NaphG adduct (decreased ~20-fold relative to unmodified DNA) if the Hoogsteen face is indeed being presented to the incoming dNTP.

Our conclusion from the structural results with N²-NaphG is that this is a very distorted complex and the Npg-2 complex is catalytically inactive. We utilized a variation of the DNA polymerase catalytic scheme in which an inactive ternary complex is in equilibrium with the active complex (56). This model was used to fit the data for the burst plots from incorporation experiments with all of the N²-guanyl adducts (Fig. 8), with only the rate constants k₅, k₆, and k₇ allowed to vary. (The “trap” experiment results with N²-AnthG presented in Fig. 6 argue that the systems are operating under single-turnover conditions and that the course of each of these sets of assays occurs without dissociation of DNA from Dpo4.) The fits in Fig. 8 were good and support the model with the inactive complex. The partial-burst results cannot be accommodated with a minimal mechanism in which there is no alternate complex (56). In the model, the inactive complex is not necessarily a single one, e.g. the structure described in Fig. 7 (E and F) but could be a set of inactive complexes. The model shows an increased partitioning of an initial ternary Dpo4-DNA-dNTP complex into the inactive form, as reflected in the ratio k₅/k₆ (Table 6). Although we have discussed the Npg-2 structure (Fig. 7) in the context of being inactive, we cannot rule out the possibility that limited pairing with C has some productive nature and occurs at a slow rate. However, we do not believe that this is a dominant complex that explains the preferential incorporation of dCTP (Table 3).

Exactly why the catalytic efficiency with N²-AnthG is higher than with N²-NaphG is unclear. The trend is more obvious with the pre-steady-state results than those for the steady state (Fig. 10). We attempted to crystallize N²-AnthG complexes but have not been unsuccessful. The structures of the N²-NaphG complexes are suggestive that adding an even larger adduct to the N2 atom might drive the conformational equilibrium toward the syn orientation, simply because the added bulk cannot be accommodated easily in the Dpo4 active site when in the anti position. Such an occurrence was certainly true of the N²-BPDE-G structures in which the BPDE moiety was either in the syn configuration or flipped out of the active site completely (45). The efficiency of incorporation opposite N²-AnthG is still diminished relative to the smaller adducts, but “stabilizing” the syn conformation could allow for greater efficiency during nucleotide insertion (with an additional effect being downstream perturbation of the DNA helix). Alternatively, N²-AnthG could be accommodated in a manner similar to N²-BPDE-dG by moving into the gap between domains.

In the course of this work we examined the “thio” effects for incorporation of dCTP opposite the N²-alkyl G derivatives (Fig. 3). Little effect was seen with G, N²-MeG, and N²-EtG, but major attenuations of the burst phases were seen with N²-IbG, N²-BzG, and N²-NaphG. A simplistic view of the results is that
the phosphodiester bond-formation step is not rate-limiting
in the case of G and the two shorter alkyl derivatives
(though step 3 of Fig. 8A is rate-limiting), and the bond forma-
tion step (step 4, Fig. 8A) occurs in large derivatives.
The modeling provides some support for this view (Table 6).
However, the thio effect was not seen with N2-AnthG, which
does have a faster $k_{pol}$ value (Fig. 3). However, the modeling
does not dictate that step 4 (of Fig. 8A) is rate-limiting. The
difficulties in interpretation of the thio effect experiments
have been discussed elsewhere (82), and the work in Fig. 3
should not be interpreted in the context of a simple com-
petition between the rate constants $k_3$ and $k_4$ (Fig. 8A).

In conclusion, we have shown that the catalytic efficiency
of the Dpo4 is largely preserved with increasing adduct size, using
a homologous series of N2-guanyl derivatives (Figs. 1 and 10).
The results have general similarity to those described for
some mammalian Y-family DNA polymerases (27–29, 48), although
in this case burst kinetics persists throughout the entire series.
Also, the fidelity of incorporation was highly preserved through
the series. The structures of two bulky adduct complexes were
determined (N2-NaphG, Fig. 7), the first of which we believe
may represent an active complex (Npg-1). A second complex
(Npg-2) is not predicted to be active, but we cannot rule out the
possibility that some alternate form of the Npg-2 complex
allows Watson-Crick pairing between dCTP and N2-NaphG.
The kinetic courses of dCTP incorporation opposite all of
the N2-guanyl adducts could be described using a minimal model
with an inactive ternary complex in equilibrium with a catalyt-
ically competent form, although the caveat must be included
that this inactive complex may be more than one and the math-
ematical representation may be more complex. These struc-
ture-function relationships add to our general knowledge of
how Y-family DNA polymerases use their structural features to
process a large variety of DNA lesions.

Acknowledgments—We thank K. C. Angel for technical assistance
and K. Trisler for assistance in preparation of the manuscript. We are
grateful to Dr. 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 (APS), Argonne National
Laboratory, Argonne, IL. Additional x-ray diffraction data were col-
lected on the 22-ID beamline of the Southeast Regional Collaborative
Access Team (SER-CAT) at APS. SER-CAT supporting institutions
may be found on-line.

REFERENCES
1. Kornberg, A., and Baker, T. A. (1992) DNA Replication, 2nd ed., W. H.
Freeman, New York
2. Searle, C. E. (1984) Chemical Carcinogens, 2nd Ed., Am. Chem. Soc.,
Washington, DC
3. Friedberg, E. C., Walker, G. C., Siede, W., Wood, R. D., Schultz, R. A., and
Ellenberger, T. (2006) DNA Repair and Mutagenesis, 2nd Ed., ASM Press,
Washington, DC
4. Yasui, M., Matsui, S., Ihara, M., Laxmi, Y. R., Shibutani, S., and Matsuda, T.
(2001) Nucleic Acids Res. 29, 1994–2001
5. Terashima, I., Matsuda, T., Fang, T.-W., Suzuki, N., Kobayashi, J., Kohda,
K., and Shibutani, S. (2001) Biochemistry 40, 4106–4114
6. Liu, X., Lao, Y., Yang, Y., Hecht, S. S., and Moriya, M. (2006) Biochemistry
45, 12998–129905
7. Lao, Y., and Hecht, S. S. (2005) Chem. Res. Toxicol. 18, 711–721
8. Zhang, S., Villalta, P. W., Wang, M., and Hecht, S. S. (2006) Chem. Res.
Toxicol. 19, 1386–1392
9. 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
10. Wang, M., Lao, Y., Cheng, G., Shi, Y., Villalta, P. W., and Hecht, S. S. (2007)
Chem. Res. Toxicol. 20, 625–633
11. Turesky, R. J., Ross, S. C., Welti, D. H., Lay, J. O., Jr., and Kladubar, F. F.
(1992) Chem. Res. Toxicol. 5, 479–490
12. Meehan, T., and Straub, K. (1979) Nature 277, 410–412
13. Cheng, S. C., Hilton, B. D., Roman, J. M., and Dipple, A. (1989) Chem. Res.
Toxicol. 2, 334–340
14. Fang, J. L., and Vaca, C. E. (1997) Carcinogenesis 18, 627–632
15. Matsuda, T., Terashima, I., Matsumoto, Y., Yabushita, H., Matsu, S., and
