THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 286, NO. 20, pp. 17503–17511, May 20, 2011

The Roles of DNA Polymerases κ and ι in the Error-free Bypass of N^2-Carboxyalkyl-2′-deoxyguanosine Lesions in Mammalian Cells*[S]

Received for publication, February 18, 2011, and in revised form, March 18, 2011 Published, JBC Papers in Press, March 28, 2011, DOI 10.1074/jbc.M111.232835

Bifeng Yuan[‡§], Changjun You[‡], Nisana Andersen[‡], Yong Jiang[¶], Masaaki Moriya[‖], Timothy R. O'Connor[**], and Yinsheng Wang[‡¶1]

From the [‡]Department of Chemistry and [¶]Environmental Toxicology Graduate Program, University of California, Riverside, California 92521, the [§]College of Chemistry and Molecular Sciences, Wuhan University, Wuhan 430072, China, the [‖]Laboratory of Chemical Biology, Department of Pharmacological Sciences, State University of New York, Stony Brook, New York 11794, and the [**]Biology Division, Beckman Research Institute, City of Hope National Medical Center, Duarte, California 91010

To counteract the deleterious effects of DNA damage, cells are equipped with specialized polymerases to bypass DNA lesions. Previous biochemical studies revealed that DinB family DNA polymerases, including *Escherichia coli* DNA polymerase IV and human DNA polymerase κ, efficiently incorporate the correct nucleotide opposite some N^2-modified 2′-deoxyguanosine derivatives. Herein, we used shuttle vector technology and demonstrated that deficiency in *Polk* or *Poli* in mouse embryonic fibroblast (MEF) cells resulted in elevated frequencies of G→T and G→A mutations at N^2-(1-carboxyethyl)-2′-deoxyguanosine (N^2-CEdG) and N^2-carboxymethyl-2′-deoxyguanosine (N^2-CMdG) sites. Steady-state kinetic measurements revealed that human DNA polymerase ι preferentially inserts the correct nucleotide, dCMP, opposite N^2-CEdG lesions. In contrast, no mutation was found after the N^2-CEdG- and N^2-CMdG-bearing plasmids were replicated in *POLH*-deficient human cells or *Rev3*-deficient MEF cells. Together, our results revealed that, in mammalian cells, both polymerases κ and ι are necessary for the error-free bypass of N^2-CEdG and N^2-CMdG. However, in the absence of polymerase κ or ι, other translesion synthesis polymerase(s) could incorporate nucleotide(s) opposite these lesions but would do so inaccurately.

DNA is susceptible to damage by endogenous and exogenous agents (1). To minimize cell death arising from replication blockage and mutations emanating from nucleotide misincorporation opposite the damage site, cells are equipped with several pathways to repair DNA lesions (2). Additionally, cells have evolved two major mechanisms to tolerate unrepaired DNA lesions (3), including homologous recombination and translesion synthesis (TLS)[2] (3). In this regard, synthesis past many DNA lesions requires the replacement of replicative DNA polymerases with one or a few specialized polymerases, most of which belong to the Y-family (4). Along this line, phylogenetic analysis revealed five branches of Y-family polymerases, which include UmuC, DinB, polymerase η, polymerase ι, and REV1 (4, 5). All of these, except for UmuC, are found in eukaryotic cells (4, 5). Apart from the Y-family DNA polymerases, some B-family DNA polymerases also participate in TLS. These include DNA polymerase II in *Escherichia coli* and polymerase ζ, which is composed of the catalytic subunit REV3 and the regulatory subunit REV7, in eukaryotic cells (4, 5).

These specialized DNA polymerases are characterized by their relatively low fidelity in replicating undamaged DNA but have the capability to bypass those DNA lesions that normally block DNA synthesis by replicative DNA polymerases (4). It was advocated that specialized DNA polymerases are evolved by nature to copy DNA base damage accurately, but they no longer bear the ability to replicate unmodified DNA with high fidelity (6). The best known example of this is the efficient and accurate bypass of *cis*,*syn*-cyclobutane pyrimidine dimers by polymerase η (7, 8), which is encoded by the xeroderma pigmentosum-variant gene (*i.e. POLH*) in humans (9). Xeroderma pigmentosum-variant patients exhibit elevated susceptibility to developing skin cancer (10). Several recent studies also demonstrated that DinB DNA polymerase, a Y-family polymerase conserved in all three kingdoms of life (11), can insert the correct dCMP opposite several N^2-substituted guanine lesions at an efficiency that is similar to or better than opposite an unmodified guanine (12–14). The x-ray crystal structure of the catalytic core of human DNA polymerase κ, along with primer/template DNA and an incoming nucleotide, reveals the lack of steric hindrance in the minor groove at the primer-template junction (15), which may account for the tolerance of polymerase κ toward the minor groove N^2-dG lesions. By using shuttle vector technology, we also showed that DinB (*i.e.* polymerase IV) is the major polymerase involved in the accurate bypass of the two diastereomers of N^2-(1-carboxyethyl)-2′-deoxyguanosine (N^2-CEdG) in *E. coli* cells (14). However, it remains unclear whether this finding can be extended to mammalian cells and whether other mammalian TLS polymerase(s) are also required for the error-free bypass of N^2-CEdG.

* This work was supported by National Institutes of Health Grant R01 DK082779 (to Y. W.). This work was also supported by the California Tobacco-related Disease Research Program Grants 17RT-0109 (to T. R. O.), 18DT-0009 (to N. A.), and 18XT-0073 (to Y. W.).

[S] The on-line version of this article (available at http://www.jbc.org) contains supplemental Figs. S1–S5.

[1] To whom correspondence should be addressed: Dept. of Chemistry, University of California, Riverside, CA 92521-0403. Tel.: 951-827-2700; Fax: 951-827-4713; E-mail: yinsheng.wang@ucr.edu.

[2] The abbreviations used are: TLS, translesion synthesis; MEF, mouse embryonic fibroblast; N^2-CEdG, N^2-(1-carboxyethyl)-2′-deoxyguanosine; N^2-CMdG, N^2-carboxymethyl-2′-deoxyguanosine; ODN, oligodeoxyribonucleotide; TGFP, Turbo GFP.

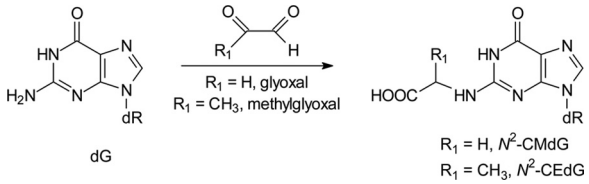

FIGURE 1. **The formation of N^2-CEdG and N^2-CMdG.**

N^2-CEdG is the major stable DNA adduct formed from methylglyoxal, a glycolytic metabolite (Fig. 1) (16). This modified nucleoside was found to be present in urine samples from healthy human subjects (17), and the lesion was observed more frequently in kidney and aortic cells of diabetic and uremic patients than in the corresponding cells of healthy human subjects (18). N^2-CEdG can also be detected in cultured human cells, and the exposure of these cells to methylglyoxal or glucose further enhanced the formation of this lesion (14). Moreover, recent studies illustrated that the exposure of isolated DNA to the structurely related glyoxal can give rise to the formation of N^2-carboxymethyl-2′-deoxyguanosine (N^2-CMdG) (19, 20). The major objective of the present study is to define the roles of polymerase κ and other mammalian TLS DNA polymerases in bypassing the N^2-CEdG and N^2-CMdG lesions in cells.

EXPERIMENTAL PROCEDURES

Materials—Unmodified ODNs used in this study were purchased from Integrated DNA Technologies (Coralville, IA). [γ-^{32}P]ATP was obtained from Perkin Elmer Life Sciences, and chemicals unless otherwise noted were from Sigma-Aldrich. Human polymerases η and ι were purchased from Enzymax (Lexington, KY). Shrimp alkaline phosphatase was obtained from USB Corp. (Cleveland, OH), and all other enzymes unless otherwise specified were from New England Biolabs.

The *Polk*-, *Rev3*-, and *Poli*-deficient mouse embryonic fibroblast (MEF) cells were kindly provided by Professors Haruo Ohmori (21), Richard D. Wood (22), and Roger Woodgate (23), respectively. The C-Tag cells (*POLH*-deficient cells), which are SV40-immortalized cells derived from xeroderma pigmentosum-variant patients, were provided generously by Professor William K. Kaufmann (24).

Preparation and LC-MS/MS Characterizations of ODN Substrates Containing an (S)-N^2-CEdG, (R)-N^2-CEdG, or N^2-CMdG—The 17-mer ODN 5′-GCGCAAA*X*CTAGAGCTC-3′ (*X* = N^2-CEdG) was synthesized following previously reported procedures (25). The corresponding ODN bearing an N^2-CMdG was prepared by using the same procedures except that glycine was used in lieu of alanine to displace the fluorine atom in 2-fluoro-2′-deoxyinosine after ODN assembly.

The identities of N^2-CEdG- and N^2-CMdG-containing 17-mer ODNs were confirmed by online capillary HPLC electrospray ionization-MS/MS (supplemental Figs. S1 and S2) using an Agilent 1100 capillary HPLC pump (Agilent Technologies) interfaced with an LTQ linear ion-trap mass spectrometer (Thermo Fisher Scientific, San Jose, CA), which was configured for monitoring the fragmentation of the $[M - 4H]^{4-}$ ions of the 17-mer ODNs. A 5-min linear gradient of 0–30% methanol followed by a 30 min of 30–50% methanol in 400 mM 1,1,1,3,3,3-hexafluoro-2-propanol buffer (pH adjusted to 7.0 by addition of triethylamine) was employed.

Construction of pTGFP-Hha10 and pMTEX4 Genomes Harboring a Site-specifically Inserted (S)-N^2-CEdG, (R)-N^2-CEdG, N^2-CMdG, or dG—A previously described method (26) was used to insert a structurally defined lesion into a unique site in the pTGFP-Hha10 vector. Briefly, we first nicked the original vector with N.BstNBI to produce a gapped vector and a 33-mer single-stranded DNA (see Fig. 2, *A* and *B*). In this respect, the pTGFP-Hha10 vector contains two N.BstNBI recognition sites upstream of the coding region of the Turbo GFP (TGFP) gene, and N.BstNBI nicks only one strand of duplex DNA at the fourth nucleotide 3′ to the GAGTC restriction recognition site. The 33-mer single-stranded ODN was removed subsequently from the nicked plasmid by annealing the cleavage mixture with the complementary 33-mer ODN in 50-fold molar excess. The gapped plasmid was isolated from the mixture by using 100-kDa cut-off ultracentrifugation filter units (Centricon 100 from Millipore). The purified gapped construct (see Fig. 2*C*, *lane 2*) was subsequently annealed with a 5′-phosphorylated 17-mer lesion-carrying insert and a 16-mer unmodified ODN and ligated with T4 DNA ligase (see Fig. 2*B*). In this regard, we used a 17-mer rather than 33-mer lesion-containing insert because of the relative ease in the synthesis and purification of the former substrate. The ligation mixture was incubated with ethidium bromide, and the resulting supercoiled, lesion-carrying double-stranded plasmid was isolated from the mixture by agarose gel electrophoresis (see Fig. 2*C*, *lanes 4* and *5*). The pTGFP-Hha10 vector contains an SV40 origin, which allows the vectors to be replicated in the SV40-transformed human 293T and *POLH*-deficient cells.

The replicating vector pMTEX4 was used for experiments conducted in the MEF cell lines. pMTEX4 contains a polyomavirus T antigen gene and a polyomavirus origin of replication, which allow pMTEX4 to be replicated in MEF cells. A 54-bp duplex DNA with Mlu I and NheI cohesive ends was ligated to the Mul I-NheI digested pMTEX4. The resulting pMTEX4 vector was nicked with Nb.BbvCI to produce a gapped vector and a 41-mer single-stranded DNA (supplemental Fig. S3). The subsequent procedures were similar as those for the construction of lesion-containing pTGFP-Hha10 vector as described above (supplemental Fig. S3B).

Cell Culture, Transfection, and in Vivo Replication—The 293T, *POLH*-deficient human fibroblast cells and all MEF cells were cultured at 37 °C in a 5% CO_2 atmosphere and in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum (Invitrogen), 100 units/ml penicillin, and 100 μg/ml streptomycin (ATCC, Manassas, VA). Cells (5×10^5) were seeded in six-well plates and cultured overnight, after which they were transfected with 300 ng of the lesion-containing or lesion-free pTGFP-Hha10 or pMTEX4 vector using Lipofectamine 2000 (Invitrogen). After *in vivo* replication for 24 h, DNase I was added to digest the residual DNA adsorbed to the exterior of cells (27). Cells were subsequently detached by treating with trypsin-EDTA (ATCC), the progenies of the plasmid were isolated by using an alkali lysis method (28), and residual unreplicated plasmid was further removed by DpnI digestion.

Determination of Bypass Efficiency and Mutation Frequency—The bypass efficiencies and mutation frequencies were deter-

mined by a method adapted from the restriction endonuclease and post-labeling assay (29, 30), which was developed initially by Essigmann and co-workers (31–33) for assessing the cytotoxic and mutagenic properties of DNA lesions in *E. coli* cells using the single-stranded M13 genome. The progeny genomes were amplified subsequently by PCR using Phusion high fidelity DNA polymerase. The primers were 5′-GCAGAGCTGGTTTAGTGAACCGTCAG-3′ and 5′-CTCTGCTGAAGCCAGTTACCTTCGG-3′ for pTGFP-Hha10 vector, 5′-GCTATCGAATTAATACGACTCATTATAGG-3′ and 5′-CCTTCGGAAAAAGAGTTGGTAGC-3′ for pMTEX4 vector. Target DNA was amplified for 35 cycles, each consisting of 10 s at 98 °C, 30 s at 62 °C, 2 min at 72 °C, with a final extension at 72 °C for 5 min. The resulting 3950-bp PCR products (from pTGFP-Hha10 progenies) and 3412-bp PCR products (from pMTEX4 progenies) were purified by using a QIAquick PCR purification kit (Qiagen, Valencia, CA). For the bypass efficiency assay, a portion of the above PCR fragments was treated with 10 units of SacI and 1 unit of shrimp alkaline phosphatase (USB Corp.) in 10-μl New England Biolabs buffer 4 at 37 °C for 1 h, followed by heating at 65 °C for 20 min to deactivate the phosphatase. The above mixture was then treated in a 15-μl NEB buffer 4 with 5 mM DTT, ATP (50 pmol (cold), premixed with 1.66 pmol [γ-^{32}P]ATP), and 10 units of polynucleotide kinase. The reaction was continued at 37 °C for 1 h, followed by heating at 65 °C for 20 min to deactivate the polynucleotide kinase. 10 units of FspI was added to the reaction mixture (see recognition sites highlighted in *boldface* in Fig. 3*A*), and the solution was incubated at 37 °C for 1 h, followed by quenching with 15 μl of formamide gel loading buffer containing xylene cyanol FF and bromphenol blue dyes. The mixture was loaded onto a 30% native polyacrylamide gel (acrylamide:bis-acrylamide = 19:1), and products were quantified by a phosphorimaging analysis (14). After these cleavages, the original lesion site was housed in a 14-mer/10-mer duplex, d(pGCAAAMCTAGAGCT)/d(p*CTAGNTTTGC) (bottom strand product), where M represents the nucleobase incorporated at the initial damage site during *in vivo* DNA replication, N is the paired nucleobase of M in the complementary strand, and p* designates the 5′-radiolabeled phosphate (see Fig. 3*A*). The mutation frequencies were determined from the relative amounts of different 10-mer products from the gel band intensities. On the other hand, the restriction cleavage of the product arising from the replication of the lesion-free top strand gave d(pGCAAAGCTTGAGCT)/d(p*CAAGCTTTGC). The 10-mer products were monitored instead of the 14-mer products because the former products could be resolved by native polyacrylamide gel. The bypass efficiency was calculated using the following formula, % bypass = (lesion vector bottom strand signal/lesion vector top strand signal)/(control vector bottom strand signal/control vector top strand signal).

Identification of Replication Products by Using LC-MS/MS—To identify the replication products using LC-MS, a nested PCR was performed using primers 5′-ACGGTGCCGACAGGTGCTTC-3′ and 5′-CCCTTCCGGAGACGACTAGTGC-3′. The 238-bp PCR products were treated with 50 units of SacI, 50 units of FspI, and 20 units of shrimp alkaline phosphatase in a 250-μl NEB buffer 4 at 37 °C for 2 h, followed by heating at 65 °C for 20 min. The resulting solution was extracted with phenol/chloroform/isoamyl alcohol (25:24:1, v/v), and the aqueous portion was dried with Speed-vac and dissolved in 12 μl of water. The ODN mixture was subjected to LC-MS/MS analysis. A 0.5 × 150 mm Zorbax SB-C18 column (5 μm in particle size, Agilent Technologies) was used for the separation, and the flow rate was 8.0 μl/min, which was delivered by using an Agilent 1100 capillary HPLC pump. A 5-min gradient of 0–20% methanol followed by a 35-min of 20–50% methanol in 400 mM 1,1,1,3,3,3-hexafluoro-2-propanol (pH was adjusted to 7.0 by the addition of triethylamine) was employed for the separation. The effluent from the LC column was coupled directly to an LTQ linear ion trap mass spectrometer (Thermo Electron, San Jose, CA), which was configured for monitoring the fragmentation of the $[M - 3H]^{3-}$ ions of the 14-mer ODNs (*i.e.* d(GCAAAMCTAGAGCT), where M designates A, T, C, or G).

siRNA Knockdown of POLK Gene in 293T Cells and Replication Studies in siRNA-treated Cells—POLK ON-TARGET*plus* SMARTpool (L-021038) and siGENOME Nontargeting siRNA (D-001210) were purchased from Dharmacon (Lafayette, CO). 293T cells were grown in DMEM medium (ATCC) containing 10% fetal bovine serum (Invitrogen) and penicillin/streptomycin (ATCC). Cells were seeded in six-well plates at 70% confluence and transfected with 1 μg of synthetic duplex siRNAs using Lipofectamine 2000 following the manufacturer's instructions. After a 48-h incubation, cells were co-transfected with another aliquot of siRNA and pTGFP-Hha10 lesion-containing vector followed by isolation of progeny vectors after a 24-h incubation with the same methods used previously. The isolated progeny vectors were processed using a protocol identical to that described above.

Real-time Quantitative RT-PCR and Immunoblot Analysis—Total RNA was extracted from the cells 48 and 72 h after transfection with siRNA using the RNeasy Mini Kit (Qiagen). cDNA was synthesized by using iScript™ cDNA synthesis kit (Bio-Rad) according to the manufacturer's recommended procedures. Briefly, 1 μg of total RNA was reverse-transcribed with 1 μl of iScript reverse transcriptase in a 20-μl volume reaction. The reaction was carried out at 25 °C for 5 min and at 42 °C for 30 min. The reverse transcriptase was then deactivated by heating at 85 °C for 5 min. Analyses of the transcripts for *POLK* were performed by real-time quantitative RT-PCR using iQ™ SYBR Green Supermix kit (Bio-Rad) according to the manufacturer's recommendations on a Bio-Rad iCycler system (Bio-Rad). The following primers were used for the real-time PCR: 5′-TGCTGATTTTCCACATCCCTTGAG-3′ and 5′-CTCCTTTGTTGGTGTTTCCTGTCC-3′ for *POLK*, 5′-TTTGTCAAGCTCATTTCCTGGTATG-3′ and 5′-TCTCTTCCTCTTGTGCTCTTGCTG-3′ for GAPDH. The comparative cycle threshold (Ct) method ($\Delta\Delta Ct$) was used for the relative quantification of gene expression (34). Immunoblot analysis was preformed with a total of 40 μg of whole cell lysate. Antibody that specifically recognizes human DNA polymerase κ (Santa Cruz Biotechnology) was used at a 1:2000 dilution. Human β-actin antibody (Abcam, Cambridge, MA) was used at a 1:5000 dilution. Horseradish peroxidase-conjugated secondary goat anti-mouse antibody (Santa Cruz Biotechnology) and goat anti-rabbit antibody (Abcam) were used at a 1:10,000 dilution.

Steady-state Kinetic Measurements—A 24-mer ODN, d(5′-GCGCAAAXCTAGAGCTCGAGATCT-3′) (X = S/(*R*)-N^2-CEdG) and a 5′-^{32}P-labeled 16-mer primer, d(AGATCTCGAGCTCTAG), were employed for steady-state kinetic measurements.

The 24-mer N^2-CEdG-containing ODNs, 5′-GCGCAAAX-CTAGAGCTCGAGATCT-3′, *X* = (*S/R*)-N^2-CEdG), were prepared, and steady-state kinetic analyses were performed following procedures described previously (14). The 24-mer lesion-containing template or nondamaged template (20 nM) with dG in place of the N^2-CEdG was annealed with a 5′-^{32}P-labeled 16-mer primer d(AGATCTCGAGCTCTAG) to give the primer-template complex. The primer-template complex (10 nM) was incubated at 37 °C with human polymerase η (5 ng) for 10 min or polymerase ι (5 ng) for 30 min in the presence of various concentrations of an individual dNTP and in a buffer containing 20 mM Tris-HCl (pH 7.5), 2 mM dithioreitol, 100 μg/ml bovine seine albumin, 10% glycerol, and 5 mM (for polymerase η) or 2 mM (for polymerase ι) $MgCl_2$. The individual dNTP concentration was optimized for different insertion reactions to allow for a <20% primer extension (35). The products were resolved by denaturing PAGE analysis, and 20% (29:1, acrylamide/bisacrylamide) polyacrylamide gels containing 8 M urea were used. Gel band intensities for the substrates and products were quantified by using a Typhoon 9410 Variable Mode Imager (Amersham Biosciences) and ImageQuant (version 5.2; Amersham Biosciences). The observed rate of dNTP incorporation (V_{obs}) was plotted as a function of dNTP concentration, and the apparent K_m and V_{max} steady-state kinetic parameters for the incorporation of both the correct and incorrect nucleotides were determined by fitting the rate data with the Michaelis-Menten equation.

$$V_{obs} = \frac{V_{max} \times [dNTP]}{K_m + [dNTP]} \quad \text{(Eq. 1)}$$

The k_{cat} values were then calculated by dividing the V_{max} values with the concentration of the polymerase used. The efficiency of nucleotide incorporation was determined by the ratio of k_{cat}/K_m, and the fidelity of nucleotide incorporation was calculated by the frequency of misincorporation (f_{inc}) with Equation 2.

$$f_{inc} = \frac{(k_{cat}/K_m)_{incorrect}}{(k_{cat}/K_m)_{correct}} \quad \text{(Eq. 2)}$$

RESULTS

Construction and Characterization of Plasmids Housing a Site-specifically Inserted N^2-CEdG or N^2-CMdG—Our previous studies revealed that DinB is the major DNA polymerase involved in the accurate and efficient bypass of N^2-CEdG in *E. coli* cells. The goal of the present investigation is to understand how N^2-CEdG and its structurally related N^2-CMdG are replicated in mammalian cells. To this end, we first constructed double-stranded pTGFP-Hha10 and pMTEX4 vectors housing these lesions (Fig. 2 and supplemental Fig. S3).

Constructing the lesion-bearing plasmid is among the most challenging steps when a double-stranded shuttle vector is employed for *in vivo* replication experiments. However, the use of the pTGFP-Hha10 and pMTEX4 vectors allows for the relatively easy preparation of the lesion-containing plasmid because the ligation only necessitates the insertion of a single-stranded damage-containing ODN into a gapped double-stranded vector. As depicted in Fig. 2*C* (*lane 4*), the strategy facilitated the efficient generation of lesion-containing double-stranded vector, and agarose gel electrophoresis afforded convenient purification of the resulting double-stranded vector from the ligation mixture (Fig. 2*C*); we could routinely produce the N^2-CMdG- and N^2-CEdG-containing pTGFP-Hha10 vectors at an overall yield of ~30%.

We further confirmed the incorporation of the lesion-containing insert by employing a restriction digestion/post-labeling assay. In this experiment, the aforementioned double-stranded genomes were digested with EcoRI (the unique EcoRI site is shown in Fig. 2*B*.) The nascent terminal phosphate groups were removed by using alkaline phosphatase, and the 5′-termini were subsequently rephosphorylated with [γ-^{32}P]ATP. The linearized vector was further cleaved with NheI (unique NheI site shown in Fig. 2*B*), which affords a 52-mer lesion-containing ODN if the ligation is successful (Fig. 2*B*). Indeed, denaturing PAGE analysis revealed a distinct 52-mer ^{32}P-labeled fragment, and no shorter fragments were detected (Fig. 2*D*), supporting the successful incorporation of the 17-mer lesion-containing insert and the 16-mer unmodified ODN into the gapped construct.

Effect of N^2-CEdG and N^2-CMdG on Efficiency and Fidelity of DNA Replication in Mammalian Cells—When the lesion-containing, double-stranded shuttle vector is replicated in mammalian cells, the undamaged strand may be replicated preferentially over the lesion-carrying strand (36, 37), rendering it difficult to determine accurately the mutation frequencies. To overcome this difficulty, we employed a similar strategy as reported previously (26, 36, 37) and incorporated an A:A mismatch three nucleotides away from the lesion site (Figs. 2*B* and 3*A*). This method facilitated the independent assessment of the products arising from the replication of the lesion-containing strand and its opposing unmodified strand (Fig. 3*A*).

The lesion-containing and the control lesion-free vectors were transfected separately into several mammalian cell lines, which include the 293T cells and wild-type MEF cells. To assess the roles of various TLS DNA polymerases in bypassing the above-mentioned N^2-dG lesions, we also evaluated the replication of the lesion-carrying and control lesion-free vectors in immortalized MEF cells deficient in *Polk* (21), *Rev3* (22), or *Poli* (23), as well as C-Tag cells, which are deficient in human *POLH* (24).

Our results revealed that none of the N^2-dG lesions are mutagenic in human 293T cells or wild-type MEF cells. However, all of them are mutagenic in MEF cells that are deficient in *Polk*. As depicted in Fig. 3*B*, replication of A:A mismatch-containing, lesion-free genome (labeled as "control") in wild-type MEF gives two 10-mer products, *i.e.* d(CAAGCTTTGC) (10-mer-AC) and d(CTAGCTTTGC) (10-mer-TC), from the respective replication of the top and bottom strands of the plasmid. The observation of similar amounts of products originating from the replication of the top and bottom strands indicates the lack of repair of the A:A mismatch in the plasmid. The replication of

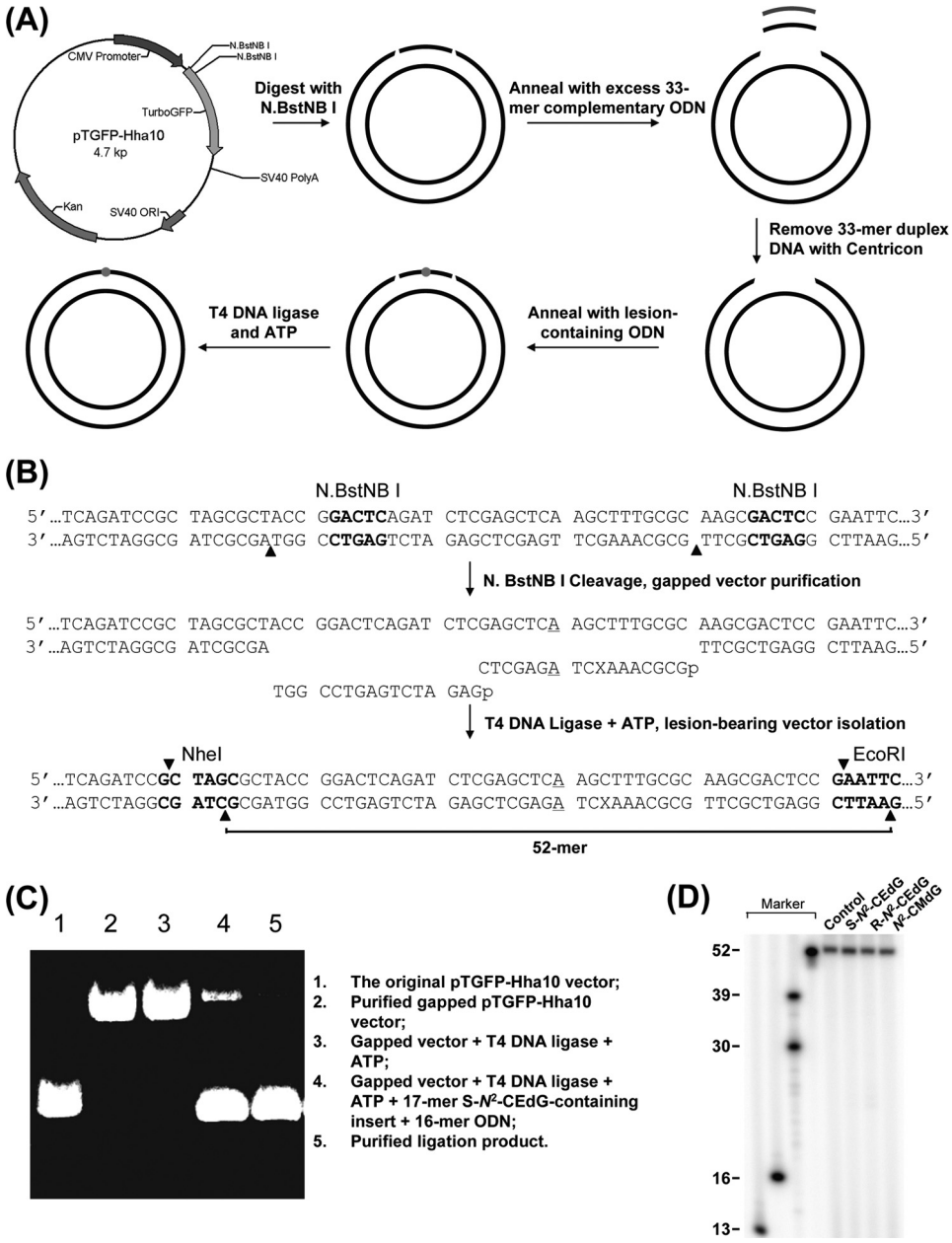

FIGURE 2. **Construction and characterization of the lesion-bearing pTGFP-Hha10 shuttle vector.** *A*, a schematic diagram showing the procedures for the preparation of the lesion-containing plasmid. SV40 ORI represents SV40 replication origin. See "Experimental Procedures" for details. *B*, enzymatic digestion and ligation for the insertion of damage-carrying ODN into the gapped pTGFP-Hha10 vector. *X* represents a dG, (*S*)-N^2-CEdG, (*R*)-N^2-CEdG, or N^2-CMdG, and the A:A mismatch site is *underlined*. The restriction recognition sites are highlighted in *boldface*, and cleavage sites are indicated by *filled triangles*. *C*, agarose (1%) gel electrophoresis for monitoring the processes for the construction of the (*S*)-N^2-CEdG-bearing pTGFP-Hha10 vector. *D*, restriction digestion and post-labeling for assessing the integrity of the control and lesion-containing vectors (see text).

the N^2-CEdG- and N^2-CMdG-harboring genomes in *Polk*-deficient background results in the appearance of two new bands in the gel (Fig. 3*B*), which are attributed to d(CTAGTTTGC) (10-mer-TT) and d(CTAGATTTGC) (10-mer-TA) based on LC-MS/MS monitoring of the complementary 14-mer fragments from the same restriction digestion mixtures (supplemental Fig. S4 and Fig. 3*A*). In this context, the latter analysis revealed the formation of only the two mutated products as labeled to the right of Fig. 3*B*; other mutagenic products are below the detection limit of the LC-MS/MS method. Quantification revealed that the replication of (*S*)-N^2-CEdG-, (*R*)-N^2-CEdG-, and N^2-CMdG-containing genomes in *Polk*-deficient background can give rise to G→A transition at frequencies of 23.3, 23.6, and 23.9%, respectively, and G→T transversion at frequencies of 15.1, 15.8, and 15.0%, respectively (Fig. 4*B*). The bypass efficiencies for the three lesions are similar in 293T cells and in wild-type MEF cells, and the deficiency in *Polk* did not compromise the bypass efficiencies. In this regard, the bypass efficiencies for (*S*)-N^2-CEdG, (*R*)-N^2-CEdG, and N^2-CMdG in wild-type MEF are 99.4, 100.6, and 99.0%, respectively, whereas the corresponding bypass efficiencies in the *Polk*-deficient MEF are 97.8, 95.4, and 98.3% (Fig. 4*A*).

The above results support that the polymerase κ-mediated bypass of N^2-substituted dG lesions in mammalian cells is accu-

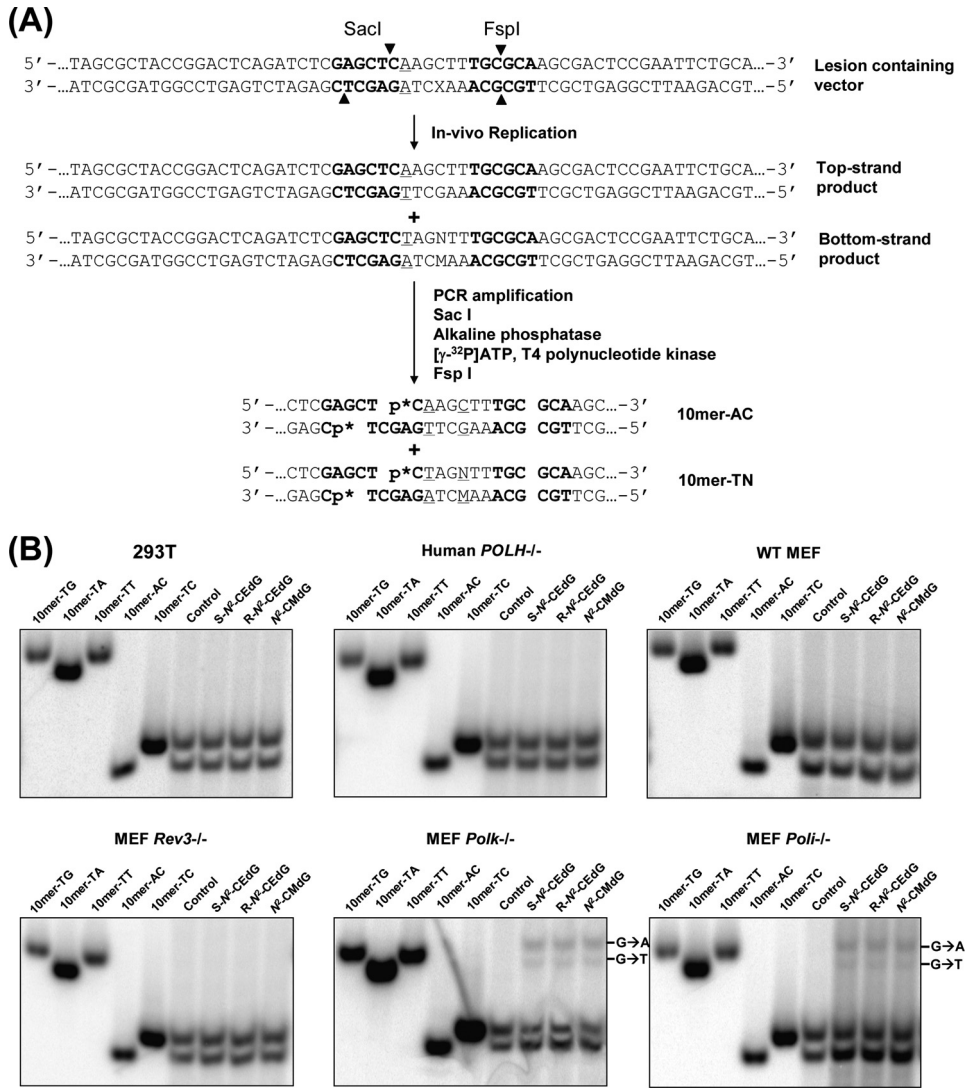

FIGURE 3. **Determination of bypass efficiencies and mutation frequencies of (*S*)- or (*R*)-N^2-CEdG and N^2-CMdG.** *A*, restriction digestion of PCR products of the progeny genome arising from *in vivo* replication of N^2-CEdG- and N^2-CMdG-bearing genome in mammalian cells. The PCR fragments were digested with SacI, and the nascent terminal phosphate groups in the resulting DNA fragments were removed by shrimp alkaline phosphatase. The 5′-termini were labeled with [γ-^{32}P]ATP, and the resulting DNA fragments were further digested with FspI. DNA fragments subsequently were resolved on a 30% native PAGE gel. *B*, nondenaturing PAGE for assessing the bypass efficiencies and mutation frequencies of N^2-CEdG and N^2-CMdG in mammalian cells that are proficient in TLS DNA polymerase (human 293T cells and wild-type MEF cells) or deficient in polymerase η (*POLH*$^{-/-}$), polymerase ζ (*Rev3*$^{-/-}$), polymerase κ (*Polk*$^{-/-}$), and polymerase ι (*Poli*$^{-/-}$).

rate. This result is in keeping with our recently published results showing that polymerase IV, the *E. coli* ortholog of mammalian polymerase κ, is the major polymerase involved in the accurate bypass of N^2-CEdG in *E. coli* cells and that the human polymerase κ-mediated *in vitro* nucleotide incorporation opposite N^2-CEdG is both accurate and efficient (14).

We next assessed whether other TLS DNA polymerase(s) are also involved in the bypass of N^2-CEdG and N^2-CMdG in mammalian cells. Our results revealed that *Poli* deficiency also led to elevated frequencies of G→A (18.0, 19.3, and 20.4% for (*S*)-N^2-CEdG, (*R*)-N^2-CEdG, and N^2-CMdG, respectively) and G→T mutations (12.6, 13.1, and 13.6%, respectively, Fig. 4*B*). In addition, the bypass efficiencies for the three lesions in *Poli*-deficient cells were 105.3, 101.2, and 99.5%, respectively (Fig. 4*A*), which were similar to those found in the wild-type MEF cells.

The observation of substantial frequencies of G→A and G→T mutations in *Poli*-deficient MEF suggests that the polymerase ι-mediated nucleotide insertion opposite the above N^2-dG lesions is also accurate. To examine whether this is the case, we performed steady-state kinetic measurements for human polymerase ι-induced nucleotide incorporation opposite the two diastereomers of N^2-CEdG as well as an unmodified dG. The results indeed demonstrated that human polymerase ι preferentially inserts the correct dCMP across the N^2-CEdG lesions (Table 1). Importantly, the efficiencies for dCMP incorporation opposite the two N^2-CEdG derivatives are only slightly lower than that for dCMP insertion opposite an unmodified dG (Table 1). This is consistent with the previous observations showing that polymerase ι-mediated nucleotide insertion opposite other minor groove N^2-dG adducts is both accurate and efficient (38, 39).

We next assessed the role of polymerases η and ζ on the bypass of N^2-CEdG and N^2-CMdG in mammalian cells. It turned out that the replication of N^2-CEdG and N^2-CMdG-

bearing plasmids in *POLH*- or *Rev3*-deficient cells did not give rise to detectable mutations. These data underscored the lack of involvement of polymerase η or ζ in bypassing those lesions in a background that is proficient in *Polk* or *Poli*.

DISCUSSION

We developed a shuttle vector-based method for assessing quantitatively how (*S*)-N^2-CEdG, (*R*)-N^2-CEdG, and N^2-CMdG perturb the efficiency and fidelity of DNA replication in mammalian cells. With the use of this method and mammalian cells that are deficient in TLS DNA polymerases, we investigated the roles of various mammalian TLS DNA polymerases in bypassing these lesions. The most important finding made from the present study is that, in mammalian cells, the polymerase κ- and polymerase ι-mediated nucleotide incorporation opposite N^2-CEdG and N^2-CMdG is error-free. This result reinforced the notion that N^2-substituted dG lesions are cognate substrates for DinB DNA polymerases (12, 14), and it also ascertained that polymerase ι, a Y-family polymerase present only in higher eukaryotes (4), accurately bypasses these lesions.

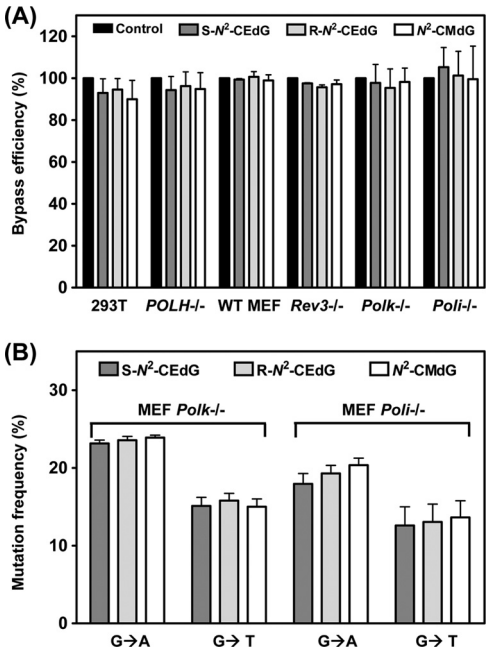

FIGURE 4. **Bypass efficiencies (*A*) and mutation frequencies (*B*) of dG, (*S*)-N^2-CEdG, (*R*)-N^2-CEdG, and N^2-CMdG lesions in wild-type, *POLH*-, *Rev3*-, *Polk*-, and *Poli*-deficient mammalian cells.** *Black*, *dark gray*, *light gray*, and *white columns* represent the results for substrates carrying dG, (*S*)-N^2-CEdG, (*R*)-N^2-CEdG and N^2-CMdG, respectively. The data represent the means and S.D. of results from three independent experiments.

Our current study also underscored that, in a genetic background lacking polymerase κ or ι, some other polymerase(s) can also bypass the above two lesions, though with compromised fidelity. In this context, our mutagenesis data revealed substantial frequencies of dAMP and dTMP misincorporation opposite the N^2-dG lesions when the replication was performed in *Polk*- or *Poli*-deficient cells. The *in vitro* steady-state kinetics experiments with purified human polymerase η illustrated that, aside from inserting the correct nucleotide (dCMP), the polymerase can incorporate dAMP and dTMP at efficiencies that are similar or slightly lower than for dCMP insertion (Table 1), suggesting that polymerase η might be responsible for the mutagenic bypass of the N^2-dG lesions in a *Polk*- or *Poli*-deficient background. This finding is reminiscent of what was recently observed for the bypass of *cis*,*syn*-cyclobutane thymine dimer in mammalian cells. In this regard, polymerase η can perform the accurate and efficient nucleotide insertion opposite the lesion (40); in the absence of polymerase η, the lesion also can be bypassed by other TLS polymerases, but with poor accuracy (40, 41). In addition, a previous report showed that another N^2-modified 2′-deoxyguanosine derivatives, N^2-benzo[*a*]pyrene-guanine, can be bypassed efficiently and accurately

TABLE 1

Fidelity of nucleotide incorporation by human polymerases η and ι on N^2-CEdG-containing substrates and the undamaged substrate as determined by steady-state kinetic measurements. k_{cat} and K_m are average values based on three independent measurements

dNTP	k_{cat}	K_m	k_{cat}/K_m	f_{inc}
	min^{-1}	nM	$nM^{-1}\ min^{-1}$	
Human polymerase η				
(*S*)-N^2-CEdG-containing substrate				
dATP	$(5.06 \pm 0.01) \times 10^{-2}$	$(3.53 \pm 0.29) \times 10^{3}$	1.43×10^{-5}	0.63
dGTP	$(5.73 \pm 0.37) \times 10^{-2}$	$(2.05 \pm 0.42) \times 10^{4}$	2.79×10^{-6}	0.12
dCTP	$(3.01 \pm 0.02) \times 10^{-2}$	$(1.32 \pm 0.04) \times 10^{3}$	2.28×10^{-5}	1
dTTP	$(1.50 \pm 0.02) \times 10^{-2}$	$(6.90 \pm 0.64) \times 10^{2}$	2.17×10^{-5}	0.95
(*R*)-N^2-CEdG-containing substrate				
dATP	$(3.53 \pm 0.01) \times 10^{-2}$	$(2.85 \pm 0.25) \times 10^{3}$	1.24×10^{-5}	0.52
dGTP	$(4.86 \pm 0.01) \times 10^{-2}$	$(5.60 \pm 0.62) \times 10^{4}$	8.69×10^{-7}	0.069
dCTP	$(2.61 \pm 0.02) \times 10^{-2}$	$(1.10 \pm 0.09) \times 10^{3}$	2.37×10^{-5}	1
dTTP	$(1.57 \pm 0.02) \times 10^{-2}$	$(5.62 \pm 0.36) \times 10^{2}$	2.79×10^{-5}	1.2
Control dG-containing substrate				
dATP	$(4.33 \pm 0.01) \times 10^{-2}$	$(9.27 \pm 0.52) \times 10^{3}$	4.67×10^{-6}	0.00013
dGTP	$(1.53 \pm 0.003) \times 10^{-2}$	$(2.52 \pm 0.44) \times 10^{3}$	6.07×10^{-6}	0.00017
dCTP	$(3.57 \pm 0.04) \times 10^{-2}$	$(9.92 \pm 0.72) \times 10^{-1}$	3.61×10^{-2}	1
dTTP	$(3.22 \pm 0.001) \times 10^{-2}$	$(2.56 \pm 0.21) \times 10^{4}$	1.26×10^{-6}	3.5×10^{-5}
Human polymerase ι[a]				
(*S*)-N^2-CEdG-containing substrate				
dCTP	$(1.16 \pm 0.01) \times 10^{-2}$	$(2.23 \pm 0.38) \times 10^{4}$	5.24×10^{-7}	1
(*R*)-N^2-CEdG-containing substrate				
dCTP	$(1.07 \pm 0.02) \times 10^{-2}$	$(1.97 \pm 0.12) \times 10^{4}$	5.43×10^{-7}	1
Control dG-containing substrate				
dCTP	$(1.32 \pm 0.04) \times 10^{-2}$	$(1.06 \pm 0.08) \times 10^{4}$	1.24×10^{-6}	1

[a] No apparent incorporation of dAMP, TMP, or dGMP was observed even when high concentrations of dNTP (2 mM) were used. Thus, steady-state kinetic parameters were determined only for dCMP incorporation.

by polymerase κ in wild-type MEF cells; however, the bypass of this adduct was error-prone in the absence of polymerase κ (42). Along this line, polymerase κ was found to be important to protect mammalian cells against the lethal and mutagenic effects of benzo[*a*]pyrene (43).

To further examine whether polymerase η was involved in mutagenic bypass of the N^2-dG lesions in cells, we employed siRNA to knock down the *POLK* gene in *POLH*-deficient cells to determine whether the genetic suppression of *POLK* in *POLH*-deficient background could abrogate the G→A and G→T mutations found in *Polk*$^{-/-}$ MEFs. As a control experiment, we also treated 293T cells with *POLK* siRNA while reasoning that siRNA-induced depletion of polymerase κ in 293T cells should recapitulate the findings made for *Polk*-deficient MEFs. However, no mutagenic products could be found from replication of the N^2-CEdG- or N^2-CMdG-containing plasmids in 293T cells treated with *POLK* siRNA. This result suggests that, although siRNA treatment results in an ~85% decrease in *POLK* expression as demonstrated by real-time quantitative RT-PCR and Western analysis (supplemental Fig. S5), the remaining polymerase κ might be sufficient for bypassing the N^2-carboxyalkyl-dG lesion in transfected plasmids. We also attempted but failed to effectively knock down the expression of *Polh* in the *Polκ*-deficient MEF cells using siRNA.

Based on the above findings, we put forward a plausible model for the bypass of the N^2-dG lesions in mammalian cells. In wild-type cells, polymerase κ and/or ι insert the correct dCMP opposite the lesion and the resulting N^2-dG:dC base pair at the primer/template junction can be extended by polymerase ζ, a general mismatch extender (44), or by another TLS polymerase. In cells deficient in *Polk* or *Poli*, another polymerase (*e.g.* polymerase η) can insert a nucleotide opposite the N^2-dG lesion, but the nucleotide incorporation is error-prone, with substantial frequencies of dAMP and dTMP misincorporation. The nascent base pair at the primer/template junction can again be extended by polymerase ζ or another TLS polymerase.

N^2-dG adducts can arise from byproducts of an array of cellular processes, including glycolysis (14, 18) and lipid peroxidation (45). In addition, there is supporting evidence that these minor groove adducts may be resistant toward excision repair (46). In this vein, the N^2-dG adduct is the least abundant dG lesion produced from acetylaminofluorine, but it persists in tissues of animals treated with this carcinogen (47). The ubiquitous formation of N^2-dG adducts from byproducts of endogenous metabolism and the poor repair of these lesions may justify the need of multiple polymerases (polymerase κ and ι) for their accurate bypass in mammalian cells. In this context, little is known about how N^2-CEdG and N^2-CMdG lesions are repaired, and such studies are currently ongoing in our laboratory.

Acknowledgments—We thank Professors Haruo Ohmori, Richard D. Wood, Roger Woodgate, and William K. Kaufmann for generously providing the TLS polymerase-deficient mammalian cells. David Baker is thanked for technical assistance.

REFERENCES

1. Lindahl, T. (1993) *Nature* **362,** 709–715
2. Friedberg, E. C., Walker, G. C., Siede, W., Wood, R. D., Schultz, R. A., and Ellenberger, T. (2006) *DNA Repair and Mutagenesis*, ASM Press, Washington, D.C.
3. Lehmann, A. R., Niimi, A., Ogi, T., Brown, S., Sabbioneda, S., Wing, J. F., Kannouche, P. L., and Green, C. M. (2007) *DNA Repair* **6,** 891–899
4. Ohmori, H., Friedberg, E. C., Fuchs, R. P., Goodman, M. F., Hanaoka, F., Hinkle, D., Kunkel, T. A., Lawrence, C. W., Livneh, Z., Nohmi, T., Prakash, L., Prakash, S., Todo, T., Walker, G. C., Wang, Z., and Woodgate, R. (2001) *Mol. Cell.* **8,** 7–8
5. Prakash, S., Johnson, R. E., and Prakash, L. (2005) *Annu. Rev. Biochem.* **74,** 317–353
6. Friedberg, E. C., Wagner, R., and Radman, M. (2002) *Science* **296,** 1627–1630
7. Johnson, R. E., Prakash, S., and Prakash, L. (1999) *Science* **283,** 1001–1004
8. Johnson, R. E., Washington, M. T., Prakash, S., and Prakash, L. (2000) *J. Biol. Chem.* **275,** 7447–7450
9. Masutani, C., Kusumoto, R., Yamada, A., Dohmae, N., Yokoi, M., Yuasa, M., Araki, M., Iwai, S., Takio, K., and Hanaoka, F. (1999) *Nature* **399,** 700–704
10. Cleaver, J. E. (2005) *Nat. Rev. Cancer* **5,** 564–573
11. Ohmori, H., Ohashi, E., and Ogi, T. (2004) *Adv. Protein Chem.* **69,** 265–278
12. Jarosz, D. F., Godoy, V. G., Delaney, J. C., Essigmann, J. M., and Walker, G. C. (2006) *Nature* **439,** 225–228
13. Choi, J. Y., Angel, K. C., and Guengerich, F. P. (2006) *J. Biol. Chem.* **281,** 21062–21072
14. Yuan, B., Cao, H., Jiang, Y., Hong, H., and Wang, Y. (2008) *Proc. Natl. Acad. Sci. U.S.A.* **105,** 8679–8684
15. Lone, S., Townson, S. A., Uljon, S. N., Johnson, R. E., Brahma, A., Nair, D. T., Prakash, S., Prakash, L., and Aggarwal, A. K. (2007) *Mol. Cell.* **25,** 601–614
16. Frischmann, M., Bidmon, C., Angerer, J., and Pischetsrieder, M. (2005) *Chem. Res. Toxicol.* **18,** 1586–1592
17. Schneider, M., Thoss, G., Hübner-Parajsz, C., Kientsch-Engel, R., Stahl, P., and Pischetsrieder, M. (2004) *Chem. Res. Toxicol.* **17,** 1385–1390
18. Li, H., Nakamura, S., Miyazaki, S., Morita, T., Suzuki, M., Pischetsrieder, M., and Niwa, T. (2006) *Kidney Int.* **69,** 388–392
19. Pluskota-Karwatka, D., Pawłowicz, A. J., Tomas, M., and Kronberg, L. (2008) *Bioorg. Chem.* **36,** 57–64
20. Wang, H., Cao, H., and Wang, Y. (2010) *Chem. Res. Toxicol.* **23,** 74–81
21. Bi, X., Slater, D. M., Ohmori, H., and Vaziri, C. (2005) *J. Biol. Chem.* **280,** 22343–22355
22. Wittschieben, J. P., Reshmi, S. C., Gollin, S. M., and Wood, R. D. (2006) *Cancer Res.* **66,** 134–142
23. McDonald, J. P., Frank, E. G., Plosky, B. S., Rogozin, I. B., Masutani, C., Hanaoka, F., Woodgate, R., and Gearhart, P. J. (2003) *J. Exp. Med.* **198,** 635–643
24. Bullock, S. K., Kaufmann, W. K., and Cordeiro-Stone, M. (2001) *Carcinogenesis* **22,** 233–241
25. Cao, H., Jiang, Y., and Wang, Y. (2007) *J. Am. Chem. Soc.* **129,** 12123–12130
26. Yuan, B., O'Connor, T. R., and Wang, Y. (2010) *ACS Chem. Biol.* **5,** 1021–1027
27. Baker, D. J., Wuenschell, G., Xia, L., Termini, J., Bates, S. E., Riggs, A. D., and O'Connor, T. R. (2007) *J. Biol. Chem.* **282,** 22592–22604
28. Hirt, B. (1967) *J. Mol. Biol.* **26,** 365–369
29. Hong, H., Cao, H., and Wang, Y. (2007) *Nucleic Acids Res.* **35,** 7118–7127
30. Yuan, B., and Wang, Y. (2008) *J. Biol. Chem.* **283,** 23665–23670
31. Delaney, J. C., and Essigmann, J. M. (2004) *Proc. Natl. Acad. Sci. U.S.A.* **101,** 14051–14056
32. Delaney, J. C., and Essigmann, J. M. (1999) *Chem. Biol.* **6,** 743–753
33. Delaney, J. C., and Essigmann, J. M. (2006) *Methods Enzymol.* **408,** 1–15
34. Livak, K. J., and Schmittgen, T. D. (2001) *Methods* **25,** 402–408
35. Creighton, S., Bloom, L. B., and Goodman, M. F. (1995) *Methods Enzymol.* **262,** 232–256

36. Levine, R. L., Yang, I. Y., Hossain, M., Pandya, G. A., Grollman, A. P., and Moriya, M. (2000) *Cancer Res.* **60,** 4098–4104
37. Yang, I. Y., Johnson, F., Grollman, A. P., and Moriya, M. (2002) *Chem. Res. Toxicol.* **15,** 160–164
38. Wolfle, W. T., Johnson, R. E., Minko, I. G., Lloyd, R. S., Prakash, S., and Prakash, L. (2006) *Mol. Cell. Biol.* **26,** 381–386
39. Pence, M. G., Blans, P., Zink, C. N., Hollis, T., Fishbein, J. C., and Perrino, F. W. (2009) *J. Biol. Chem.* **284,** 1732–1740
40. Yoon, J. H., Prakash, L., and Prakash, S. (2009) *Proc. Natl. Acad. Sci. U.S.A.* **106,** 18219–18224
41. Ziv, O., Geacintov, N., Nakajima, S., Yasui, A., and Livneh, Z. (2009) *Proc. Natl. Acad. Sci. U.S.A.* **106,** 11552–11557
42. Avkin, S., Goldsmith, M., Velasco-Miguel, S., Geacintov, N., Friedberg, E. C., and Livneh, Z. (2004) *J. Biol. Chem.* **279,** 53298–53305
43. Ogi, T., Shinkai, Y., Tanaka, K., and Ohmori, H. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99,** 15548–15553
44. Johnson, R. E., Washington, M. T., Haracska, L., Prakash, S., and Prakash, L. (2000) *Nature* **406,** 1015–1019
45. Marnett, L. J. (2000) *Carcinogenesis* **21,** 361–370
46. Zewail-Foote, M., Li, V. S., Kohn, H., Bearss, D., Guzman, M., and Hurley, L. H. (2001) *Chem. Biol.* **8,** 1033–1049
47. Yasui, M., Dong, H., Bonala, R. R., Suzuki, N., Ohmori, H., Hanaoka, F., Johnson, F., Grollman, A. P., and Shibutani, S. (2004) *Biochemistry* **43,** 15005–15013