9256–9265 *Nucleic Acids Research, 2016, Vol. 44, No. 19* *Published online 27 July 2016*
doi: 10.1093/nar/gkw662

Translesion synthesis of O^4-alkylthymidine lesions in human cells

Jun Wu[1,†], Lin Li[1,†], Pengcheng Wang[2], Changjun You[1], Nicole L. Williams[2] and Yinsheng Wang[1,2,*]

[1]Department of Chemistry, University of California, Riverside, CA 92521-0403, USA and [2]Environmental Toxicology Graduate Program, University of California, Riverside, CA 92521-0403, USA

Received June 2, 2016; Revised July 9, 2016; Accepted July 13, 2016

ABSTRACT

Environmental exposure, endogenous metabolism and cancer chemotherapy can give rise to alkylation of DNA, and the resulting alkylated thymidine (alkyldT) lesions were found to be poorly repaired and persistent in mammalian tissues. Unrepaired DNA lesions may compromise genomic integrity by inhibiting DNA replication and inducing mutations in these processes. In this study, we explored how eight O^4-alkyldT lesions, with the alkyl group being a Me, Et, *n*Pr, *i*Pr, *n*Bu, *i*Bu, (*R*)-*s*Bu and (*S*)-*s*Bu, are recognized by DNA replication machinery in HEK293T human embryonic kidney cells. We found that the O^4-alkyldT lesions are moderately blocking to DNA replication, with the bypass efficiencies ranging from 20 to 33% in HEK293T cells, and these lesions induced substantial frequencies T→C transition mutation. We also conducted the replication experiments in the isogenic cells where individual translesion synthesis (TLS) DNA polymerases were depleted by the CRISPR/Cas9 genome editing method. Our results showed that deficiency in Pol η or Pol ζ, but not Pol κ or Pol ι, led to pronounced drops in bypass efficiencies for all the O^4-alkyldT lesions except O^4-MedT. In addition, depletion of Pol ζ resulted in significant decreases in T→C mutation frequencies for all the O^4-alkyldT lesions except O^4-MedT and O^4-*n*BudT. Thus, our study provided important new knowledge about the cytotoxic and mutagenic properties of the O^4-alkyldT lesions and defined the roles of TLS polymerases in bypassing these lesions in human cells.

INTRODUCTION

DNA alkylation constitutes a major type of DNA damage, where humans are exposed to alkylating agents present in the environment and from endogenous metabolism (1). Alkylating agents are also widely used as cancer chemotherapeutic drugs (2). These agents can, directly or after metabolic activation, alkylate DNA at each of the four nucleobases and the phosphate backbone (1). Among the various DNA alkylation products, those with an alkyl group attached to the *N*7 position of adenine and guanine, the *N*3 position of adenine and the O^2 position of cytosine are known to possess a labile *N*-glycosidic bond, which may give rise to abasic sites after spontaneous loss of the alkylated nucleobase (1,3). On the other hand, alkylation occurring at the O^6 position of guanine and the O^4 position of thymine alters the hydrogen bonding properties of the two nucleobases, which direct DNA polymerases to preferentially incorporate the incorrect dTMP and dGMP, respectively (1,4).

Alkylated DNA lesions, if not repaired timely, may perturb the flow of genetic information by diminishing the fidelity and efficiency of DNA replication and transcription (5–7). Previous studies have shown that, relative to the O^6-alkylated 2′-deoxyguanosines, the O^4-alkylated thymidines (O^4-alkyldT) are more resistant to repair (8–10). In addition, higher levels of O^4-EtdT could be detected in lung and leukocyte DNA of smokers than non-smokers (11,12). Thus, it is important to assess how the O^4-alkyldT lesions perturb the transmission of genetic information by inhibiting DNA replication and inducing mutations in these processes. In addition, depending on the identities of alkylating agents involved, the alkyl group adducted to DNA can range from a simple methyl group to a large pyridyloxobutyl or pyridylhydroxybutyl functionality (1,13). Therefore, it is also important to examine how the efficiency and fidelity of DNA replication across the O^4-alkylthymidine (O^4-alkyldT) lesions are influenced by the size and structure of the alkyl group. In this respect, our recently published study showed that, in *Escherichia coli* cells, all the O^4-alkyldT lesions direct high frequencies of dGMP misinsertion (14–16). A few studies have also been conducted on assessing the effects of the O^4-alkyldT lesions on DNA repli-

*To whom correspondence should be addressed. Tel: +1 951 827 2700; Fax: +1 951 827 4713; Email: Yinsheng.Wang@ucr.edu
†These two authors contributed equally to the paper as first authors.

cation in mammalian cells, and the results from these studies illustrated that O^4-MedT, O^4-EtdT and O^4-*n*PrdT (Figure 1A) could induce exclusively the T→C transition mutation (17–19).

To cope with unrepaired DNA lesions, cells are equipped with DNA damage tolerance pathways to promote the replication across damaged nucleosides in template DNA. In this vein, mammalian cells employ translesion synthesis (TLS) DNA polymerases to bypass those modified nucleosides that normally block DNA replication catalyzed by high-fidelity replicative DNA polymerases (20,21). These include polymerases η, ι and κ in the Y-family and polymerase ζ in the B-family (20,21). A recent study showed that, while depletion of any of the SOS-induced DNA polymerases alone did not compromise the bypass of the O^4-alkyldT lesions in *E. coli* cells, simultaneous deletion of all three SOS-induced DNA polymerases led to markedly reduced bypass of these lesions, suggesting that the SOS-induced DNA polymerases assume somewhat redundant roles in replication across these lesions (14). It remains unexploited how the O^4-alkyldT lesions are recognized by TLS polymerases in mammalian cells.

Unambiguous delineation of the roles of TLS polymerases in bypassing specific DNA lesions in human cells necessitates the availability of polymerase-deficient cells for replication studies. In this respect, patient-derived human skin fibroblasts deficient in DNA polymerase η and the corresponding cells complemented with wild-type human polymerase η have facilitated the assessment of the role of this polymerase in bypassing various DNA lesions (22). The investigations about the roles of other TLS polymerases in bypassing structurally defined DNA lesions have so far relied on the use of siRNA for knocking down the expression of these polymerases. However, the siRNA method frequently results in incomplete depletion of TLS polymerases, which sometimes renders it difficult to draw unequivocal conclusions about the involvement of specific polymerases in lesion bypass (22,23).

In the present study, we set out to fill in the above knowledge gaps by constructing double-stranded plasmids containing eight O^4-alkyldT lesions with varying sizes and structures of the alkyl group (Figure 1A) and by investigating how these lesions impair the efficiency and accuracy of DNA replication in HEK293T cells. We also employed the CRISPR/Cas9 genome editing method (24,25) to knock out individually the TLS polymerases in HEK293T cells and, by using these cells, we were able to unveil unambiguously the roles of these polymerases in modulating the cytotoxic and mutagenic properties of the O^4-alkyldT lesions.

MATERIALS AND METHODS

Materials

All chemicals, if not specifically described, were from Sigma-Aldrich (St. Louis, MO, USA) or EMD Millipore, and all enzymes, unless otherwise noted, were obtained from New England Biolabs (Ipswich, MA, USA). 1,1,1,3,3,3-Hexafluoro-2-propanol (HFIP) was obtained from TCI America (Portland, OR, USA) and [γ-^{32}P]ATP was purchased from Perkin Elmer (Piscataway, NJ, USA). All unmodified oligodeoxyribonucleotides (ODNs) were from Integrated DNA Technologies (IDT, Coralville, IA, USA). The 12-mer ODNs harboring a site-specifically incorporated O^4-alkyldT were synthesized using conventional phosphoramidite chemistry, as described previously (14). The identified and purities of all the lesion-harboring ODNs were confirmed by liquid chromatography-mass spectrometry (LC-MS) and tandem MS (MS/MS) analyses prior to their insertion into double-stranded plasmids.

Construction of lesion-containing and lesion-free plasmids

The lesion-containing and the lesion-free control and competitor genomes were prepared according to the previously reported procedures (22). First, a parent vector was constructed by modifying the sequence of the original pTGFP-Hha10 plasmid, which carries an SV40 replication origin and is capable of being replicated in the SV40-transformed mammalian cells (22,26). The parent vector was subsequently digested with Nt.bstNBI to generate a gapped vector (Figure 1B), followed by removing the resulting 25-mer single-stranded ODN through annealing with a 25-mer complementary ODN in large excess. The gapped plasmid was then isolated from the mixture by using 100 kDa-cutoff ultracentrifugal filter units (Millipore). To ensure complete removal of the 25-mer restriction fragment, the annealing with its complementary strand and the centrifugation steps were conducted twice. The gapped vector was filled with a 5′-phosphorylated 13-mer lesion-free ODN (5′-AATTGAGTCGATG-3′) and a 5′-phosphorylated 12-mer lesion-carrying ODN (5′-ATGGCGXGCTAT-3′) (X = O^4-alkyldT or dT) by using T4 DNA ligase and adenosine triphosphate (ATP) in the ligation buffer (Figure 1). The ligation mixture was separated by using agarose gel electrophoresis in the presence of ethidium bromide to purify the successfully ligated supercoiled plasmid. The constructed lesion-containing double-stranded vector was normalized against the lesion-free competitor vector following published procedures (27).

CRISPR/Cas9-mediated genome editing of HEK293T cells

Genome editing by the CRISPR/Cas9 system was conducted following the previously reported protocols (28), where the single guide RNAs (sgRNAs) were designed by using the online sgRNA tool (http://www.broadinstitute.org/rnai/public/analysis-tools/sgrna-design). ODNs corresponding to target sequences were obtained from IDT and inserted into the hSpCas9 plasmid pX330 (Addgene, Cambridge, MA, USA). The constructed plasmids were then transfected into HEK293T cells using Lipofectamine 2000 (Invitrogen, Carlsbad, CA, USA) and individual cell was cultured for further analysis. Genomic DNA was extracted from individual clonal cell lines, and specific DNA regions flanking the targeted sites were screened by polymerase chain reaction (PCR), followed by agarose gel electrophoresis to assess the modification efficiency and by Sanger sequencing to identify the deletion loci (Supplementary Figure S1). A set of clones with both alleles being cleaved by Cas9 were isolated, and the successful deletion of the TLS polymerases were validated by western blot analysis, as described previously (22). The guide sequences were (underline indicates the PAM motif):

Human *POLH* gene (Pol η): CTGCTCCCACGGTGAG CTGCAGG
Human *POLI* gene (Pol ι): CTACAGAGAAATGTCT TATAAGG
Human *POLK* gene (Pol κ): ATCCATGTCAATGTGC ACTATGG
Human *REV3L* gene (Pol ζ): AATGAGCCAACCTG AGTCACAAG

Cellular DNA replication and plasmid isolation

The lesion-bearing and the corresponding non-lesion control plasmids were premixed individually with the competitor genome for *in vivo* transfection; the molar ratio of competitor over control genome was 1:1, and those for the competitor over the lesion-bearing genome varied from 1:1.5 to 1:10. The HEK293T cells were cultured in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum (Invitrogen, Carlsbad, CA, USA), 100 U/ml penicillin and incubated at 37°C in 5% CO_2 atmosphere. The HEK293T cells and the CRISPR/Cas9 genome-engineered cells (1×10^5) were seeded in 24-well plate and cultured for 24 h before they were transfected with 300 ng of mixed vector by using Lipofectamine 2000 following the manufacturer's recommended procedures. The cells were harvested at 24 h following the transfection, and the progeny genomes were isolated using Qiagen Spin kit (Qiagen, Valencia, CA, USA) (29). The residual unreplicated plasmids were further digested by restriction endonuclease DpnI, followed by removing the resulting linear DNA with exonuclease III, as reported elsewhere (22). In this vein, the parent plasmid carried 25 DpnI recognition sites and cleavage at any of these sites would result in the degradation of the entire plasmid by exonuclease III and prohibit the subsequent PCR amplification of the parent vector.

PCR and polyacrylamide gel electrophoresis (PAGE) analyses

The progeny genomes arising from cellular replication were amplified by PCR with the use of GoTaq Hot Start DNA polymerase (Promega, Madison, WI). The two primers were 5′-GCTAGCGGATGCATCGACTCAATTACAG-3′ and 5′-GCTGATTATGATCTAGAGTTGCGGCCGC-3′, and the PCR amplifications started at 95°C for 2 min, followed by 30 cycles at 95°C for 30 s, 64°C for 30 s, and 72°C for 1 min, and a final 5-min extension at 72°C. The PCR products were purified using Cycle Pure Kit (Omega, Norcross, GA) and stored at -20°C until use.

For polyacrylamide gel electrophoresis (PAGE) analysis, a portion of the PCR products was treated with 5 U NcoI and 1 U shrimp alkaline phosphatase (SAP) at 37°C in 10 μl of NEB buffer 3 for 1 h, followed by heating at 80°C for 20 min to deactivate the SAP. The above mixture was then treated with 5 U of T4 polynucleotide kinase (T4 PNK) in 15 μl of NEB buffer 3 containing 5 mM dithiothreitol and ATP (10 pmol cold, premixed with 1.66 pmol [γ-^{32}P]ATP). The reaction was continued at 37°C for 0.5 h, followed by heating at 65°C for 20 min to deactivate the polynucleotide kinase. To the reaction mixture was subsequently added 2 U of SfaNI in 5 μl NEB buffer 3, and the solution was incubated at 37°C for 1.5 h, followed by quenching with 20 μl of formamide gel-loading buffer containing xylene cyanol FF and bromophenol blue dyes. The mixture was loaded onto 30% native polyacrylamide gel (acrylamide/bis-acrylamide = 19:1), and the gel band intensities were quantified using phosphorimager analysis. The effects of DNA lesions on replication efficiency and fidelity were determined by the relative bypass efficiency (RBE) and mutation frequency values, respectively. The mutation frequency was calculated from the percentage of mutagenic product (i.e. the product with T→C mutation) among the sum of all products arising from replication of the lesion-containing genome. The RBE value was calculated using the following formula, %RBE = (lesion signal/competitor signal)/(non-lesion control signal/competitor signal) × 100% (22,30).

Identification of mutagenic products by LC-MS/MS

The PCR products were digested with 30 U SfaNI restriction endonuclease and 15 U SAP in 150-μl NEB buffer 3 at 37°C for 2 h, followed by deactivation of the phosphatase at 80°C for 20 min. To the mixture was added 50 U NcoI restriction endonuclease in 5 μl NEB buffer 3 and the solution was incubated at 37°C for another 2 h. The resulting solution was extracted once with phenol/chloroform/isoamyl alcohol (25:24:1, v/v). To the aqueous layer were subsequently added 2.5 volumes of 100% ethanol and 0.1 volume of 3.0 M sodium acetate, and the solution was incubated at −20°C overnight to precipitate the DNA. The DNA pellet was then dissolved in sterile water for LC-MS/MS analysis.

An Agilent 1200 capillary HPLC system (Agilent Technologies, Santa Clara, CA, USA) and an LTQ linear ion trap mass spectrometer (Thermo Fisher Scientific, San Jose, CA, USA) were used for all the LC-MS and MS/MS experiments. An Agilent Zorbax SB-C18 column (0.5 × 250 mm, 5 μm in particle size) was employed, and the gradient for LC-MS/MS analysis was 5 min of 5–20% methanol followed by 50 min of 20–45% methanol in 400 mM HFIP (pH was adjusted to 7.0 with the addition of trimethylamine). The temperature for the ion-transport tube was maintained at 300°C. The mass spectrometer was set up for monitoring the fragmentation of the $[M-3H]^{3-}$ ions of the 13-mer [d(AATTACAGCNCGC), where 'N' represents A, T, C or G]. The fragment ions detected in MS/MS were manually assigned.

RESULTS

The objectives of the present study are to assess systematically how the eight O^4-alkyldT lesions with varying sizes and structures of the alkyl group (Figure 1A) compromise the efficiencies and fidelities of DNA replication in human cells, and to define the roles of TLS polymerases in bypassing these lesions. As noted above, unambiguous determination of the involvement of a specific TLS polymerase in bypassing a structurally defined DNA lesion in human cells often requires the availability of human cells with complete depletion of that polymerase. Recent advances in CRISPR-Cas9 genome editing method enabled the facile knockout of individual genes in cultured human cells (24,25,28). Thus, we employed this method and successfully knocked out, in HEK293T cells, the *POLH*, *POLI*, *POLK* and *REV3L*

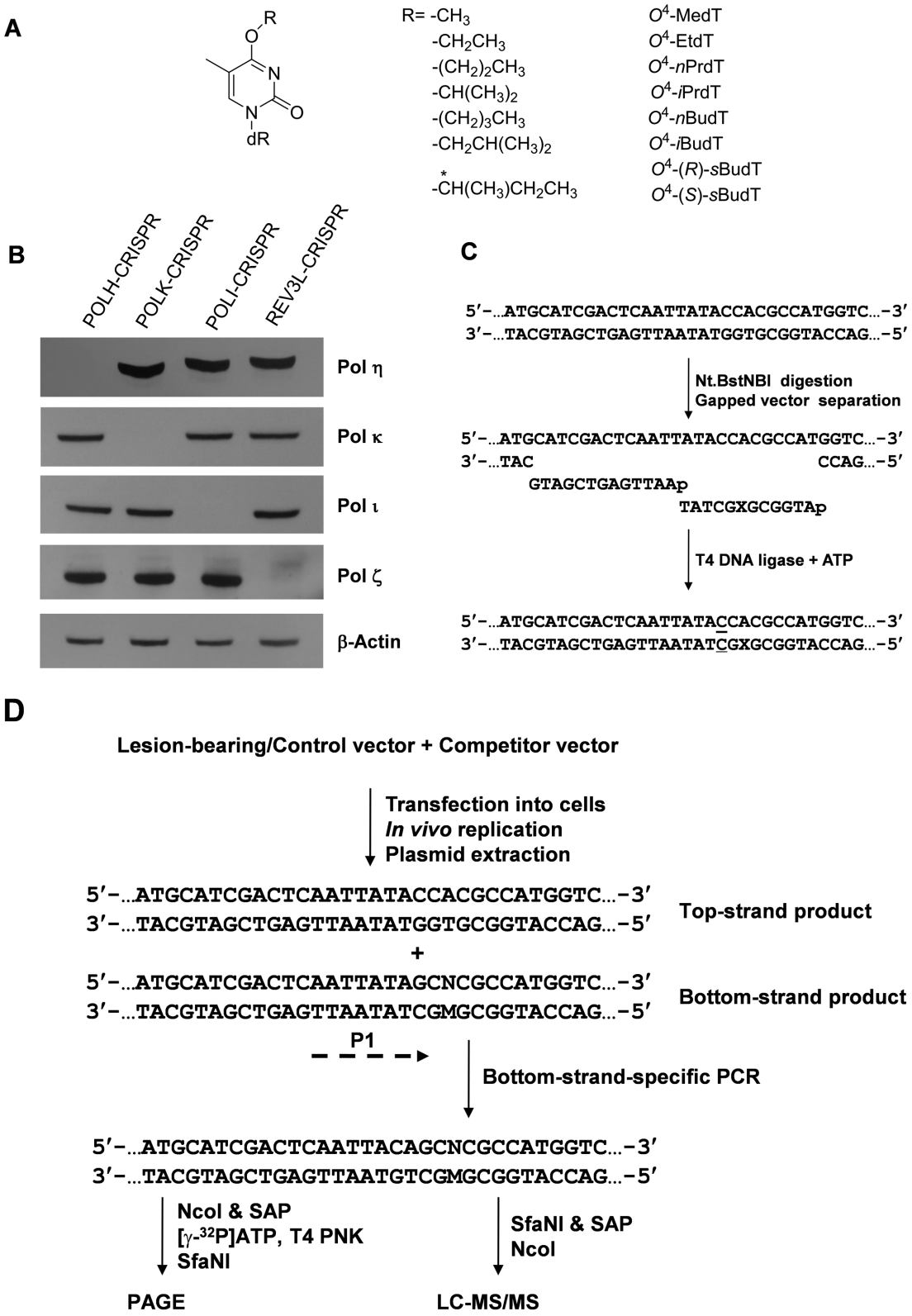

Figure 1. Experimental procedures for the cell-based replication assay with the use of translesion synthesis (TLS) polymerase-deficient HEK293T cells generated by the CRISPR/Cas9 genome editing method. (**A**) The structures of the O^4-alkyldT lesions employed in this study. (**B**) Western blot results showing the selective and complete knockout of individual TLS polymerases in HEK293T cells. (**C**) Experimental procedures for the construction of lesion-containing double-stranded vectors for replication studies. 'X' indicates the location of the O^4-alkyldT lesions and the C/C mismatch site is underlined. (**D**) Procedures for cellular replication studies. 'P1' (5′-GCTAGCGGATGCATCGACTCAATTACAG-3′) represents one of PCR primers and contains a G as the terminal 3′-nt corresponding to the C/C mismatch site of the lesion-bearing genome. It also carries a C/A mismatch 2 nt away from its 3′-end for improving the PCR specificity. 'M' and 'N' represent the nucleotide incorporated at the lesion site during DNA replication and the paired nucleotide of 'M' in the complementary strand, respectively.

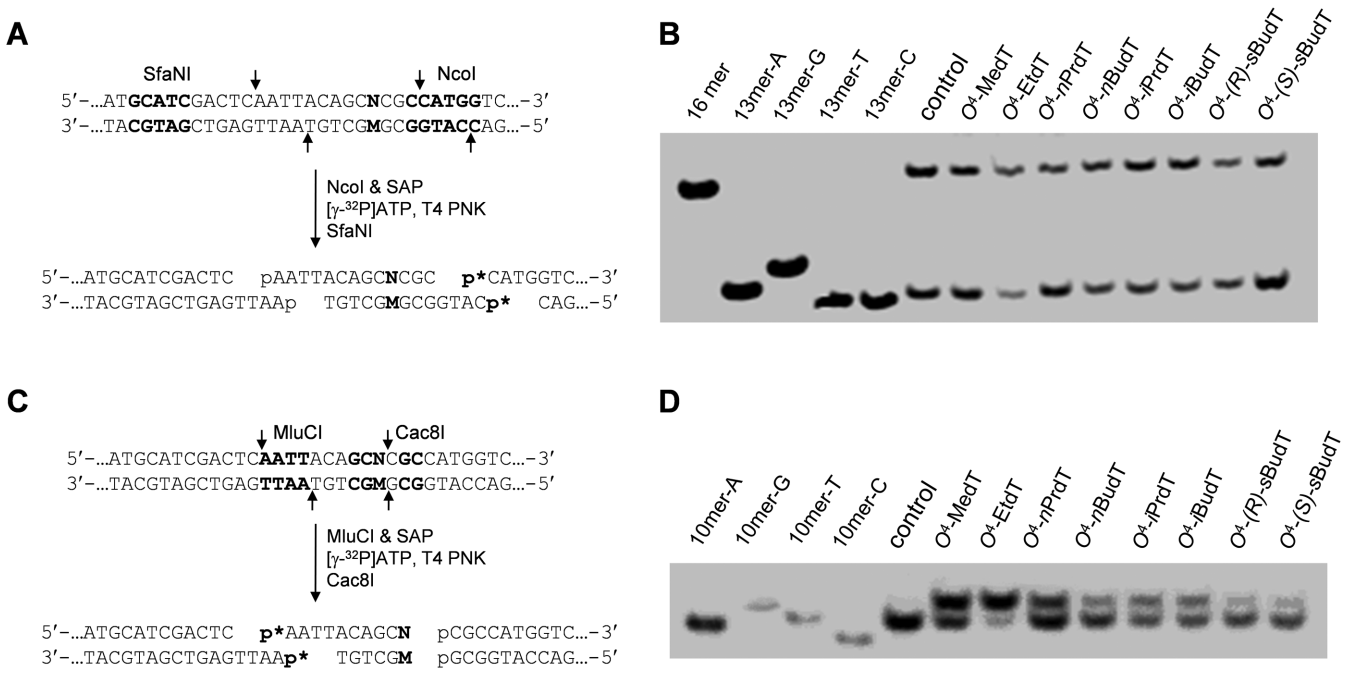

Figure 2. Restriction digestion and post-labeling method for determining the bypass efficiencies and mutation frequencies of the O^4-alkyldT lesions in HEK293T cells that are competent in TLS. (**A**) Sample processing for restriction digestion using NcoI and SfaNI, and post-labeling assay ('p*' indicates a ^{32}P-labeled phosphate group). (**B**) Representative gel images showing the NcoI/SfaNI-produced restriction fragments of interest. The restriction fragment arising from the competitor vector, i.e. d(CATGGCGATATGCTGT), is designated as '16-mer'; '13-mer A', '13-mer G', '13-mer T' and '13-mer C' represent the standard synthetic ODNs d(CATGGCGMGCTGT), where 'M' is A, G, T and C, respectively. (**C**) Sample processing for restriction cleavage using MluCI and Cac8I, and post-labeling assay. (**D**) Representative gel images showing the MluCI/Cac8I-generated restriction fragments of interest. '10-mer A', '10-mer G', '10-mer T' and '10-mer C' designate the standard synthetic ODNs d(AATTACAGCN), where 'N' is A, G, T and C, respectively. The recognition sequences for restriction enzymes are highlighted in bold, and their cleavage sites are indicated with arrows in (A) and (C). The control/competitor genome ratio was 1:1, and lesion/competitor genome ratio was 3:1 for replication experiments in the parental HEK293T cells.

genes, which encode DNA polymerases η, ι, κ and the catalytic subunit of DNA polymerase ζ, respectively (20). Our western blot results showed that the knockout of the target TLS polymerase genes is complete (Figure 1B). Although the CRISPR-Cas9 method is known to also give rise to undesirable mutations at off-target sites that resemble the on-target sequence (31,32), the western blot results revealed that the depletion of one TLS polymerase does not affect the expression levels of other TLS polymerases (Figure 1B).

We next constructed double-stranded shuttle vectors containing site-specifically inserted O^4-alkyldT lesions and the corresponding lesion-free control vector (Figure 1C). We subsequently mixed individually the lesion-bearing or undamaged control vectors with a damage-free competitor vector at fixed molar ratios and co-transfected the lesion/competitor or control/competitor plasmid mixtures into human cells. The progeny genomes were extracted from cells and amplified by PCR using a pair of primers spanning the site that initially housed the lesion. In this vein, one primer (P1) carries a G as the terminal 3′-nt corresponding to the C/C mismatch site (Figure 1D), which allows for the selective amplification of the progeny genomes emanating from the replication of the strand that initially contained the lesion (22). Additionally, P1 carries a C/A mismatch 2 nt away from its 3′-end to increase the specificity of strand-specific PCR (22,33). We then digested the resultant PCR products with two restriction enzymes, i.e. NcoI and SfaNI, or MluCI and Cac8I, and subjected the digestion mixture to LC-MS/MS and PAGE analyses (Figure 1D). In this vein, our previous study showed that parallel PAGE and LC-MS/MS analyses yield very similar bypass efficiencies and mutation frequencies for other DNA lesions (34). While we always use LC-MS/MS for the identification of the replication products, PAGE, owing to its relatively high throughput, was employed for the quantification of the replication products for determining the bypass efficiencies and mutation frequencies of the O^4-alkyldT lesions.

Our LC-MS/MS and PAGE analysis results revealed that the replication across all the O^4-alkyldT lesions could give rise to T→C mutation, and no other single-base substitution products at the lesion site could be detected (Figures 2 and 3; Supplementary Figures S2–5). In this respect, sequential digestion with NcoI and SfaNI and selective radio-labeling of the nascent 5′ termini generated from NcoI digestion give a 13-mer radiolabeled fragment, 5′-p*CATGGCGMGCTGT-3′, where 'M' designates the nucleotide inserted at the site where the O^4-alkyldT lesions were initially situated and 'p*' indicates the ^{32}P-labeled phosphate (Figure 1D). Native PAGE analysis allows for the differentiation of the non-mutagenic 13mer radiolabeled fragment (13mer-T) from the corresponding fragments harboring a T→A (13mer-A) or T→G (13mer-G) mutation. The 13-mer fragment with a T→C mutation (13mer-C), however, cannot be resolved from the non-mutagenic 13mer-T (Figure 2A and B; Supplementary Figures S2–5A and B). On the other hand, a separate sequen-

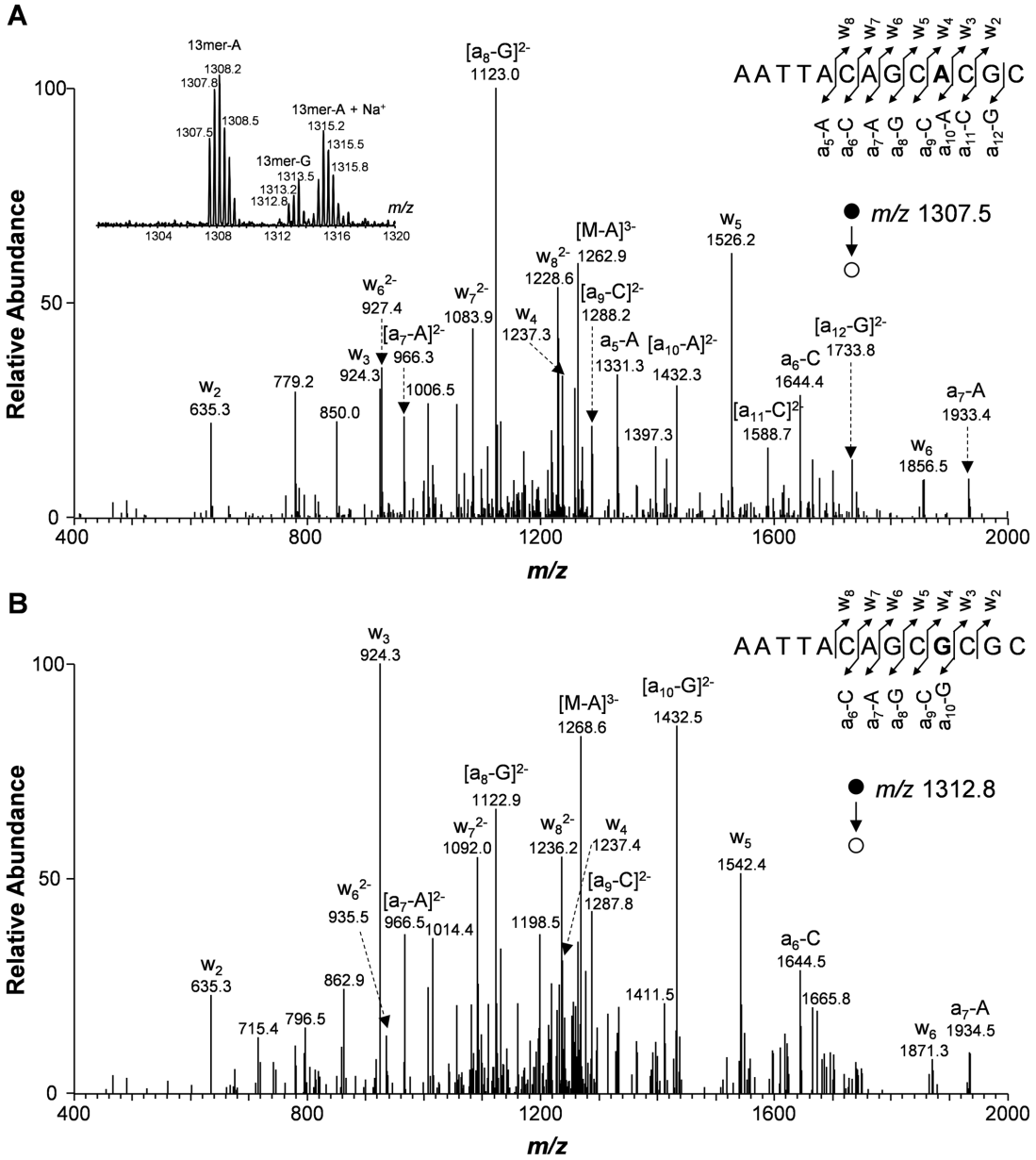

Figure 3. Restriction digestion followed by LC-MS/MS for the identification of restriction digestion products (with NcoI and SfaNI) of PCR products of progeny genome arising from the replication of O^4-*n*BudT lesions in HEK293T cells. Shown are the MS/MS for monitoring the fragmentations of the $[M–3H]^{3–}$ ions of 5′-AATTACAGC**A**CGC-3′ (wild-type product, **A**) and 5′-AATTACAGC**G**CGC-3′ (with T→C mutation, **B**). The restriction digestion method was the same as what was described in Figure 2A except that the shrimp alkaline phosphatase was added after the addition of the two restriction enzymes and that the [^{32}P]-post-labeling step was omitted. Shown in the insets are schemes summarizing the observed fragment ions and a higher-resolution 'ultra-zoon scan' ESI-MS for monitoring the $[M–3H]^{3–}$ ions of the wild-type and T→C mutation products. The nomenclature for fragment ions follow that described previously (41).

tial digestion with MluCI and Cac8I, and selective radiolabeling yield 10-mer radiolabeled fragments that permit the differentiation of the non-mutagenic product from the respective product with a T→C mutation (i.e. 10mer-A versus 10mer-G, Figure 2C and D; Supplementary Figures S2–5C and D). With the use of this method, we were able to determine quantitatively the degrees to which the O^4-alkyldT lesions inhibit DNA replication and induce mutations in human cells.

Our results revealed that the O^4-alkyldT lesions are moderate blocks to DNA replication, with bypass efficiencies varying from 20 to 35% in HEK293T cells (Figure 4A). In addition, the O^4-alkyldT lesions with a straight-chain alkyl group were more readily bypassed than those carrying the corresponding branched-chain alkyl group. In this vein, O^4-*n*PrdT and O^4-*i*PrdT were bypassed in HEK293T cells at efficiencies of 34.9 and 30.3%, respectively, and the bypass efficiencies for O^4-*n*BudT, O^4-*i*BudT, O^4-(*R*)-*s*BudT and O^4-(*S*)-*s*BudT were 31.5, 24.3, 19.9 and 21.7%, respec-

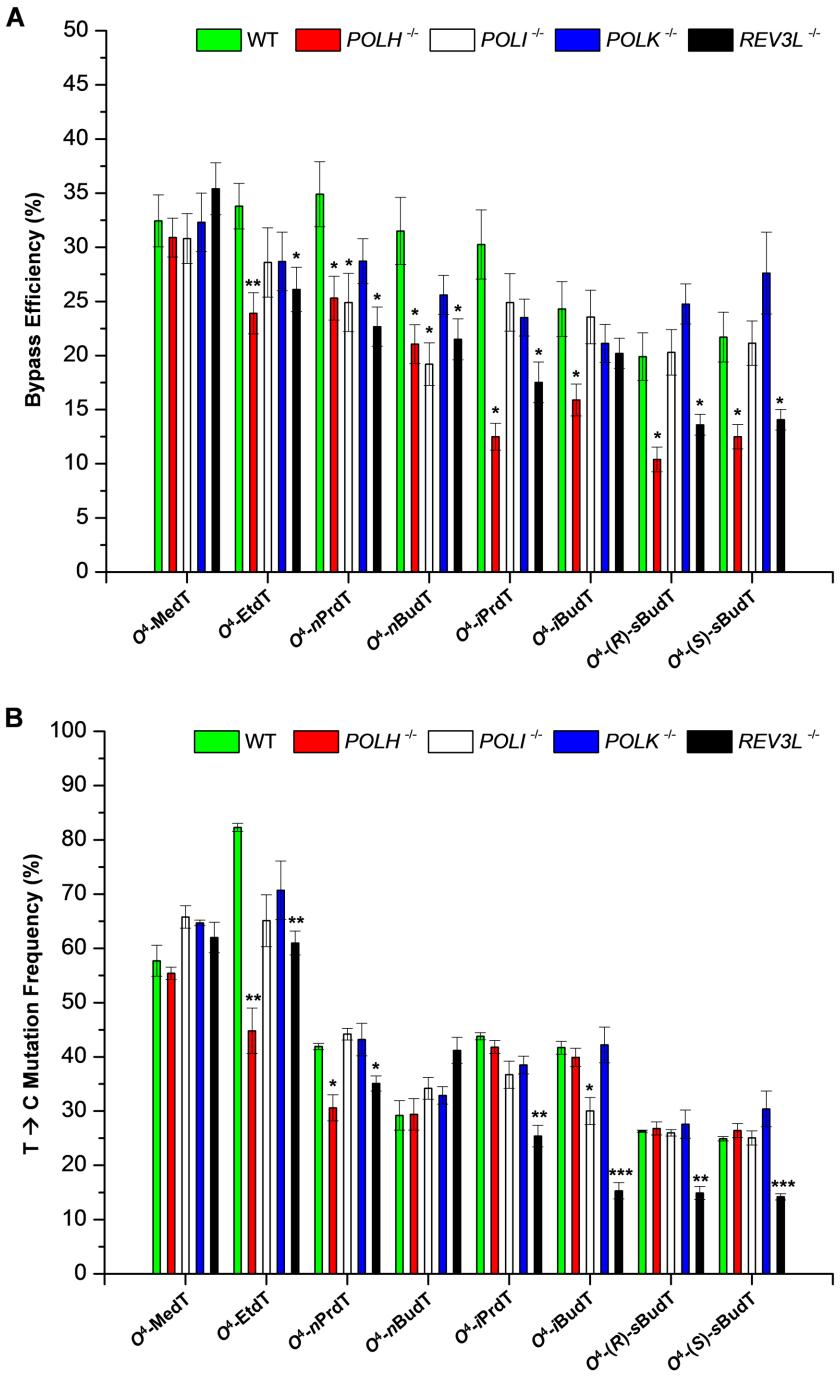

Figure 4. The bypass efficiencies (**A**) and mutation frequencies (**B**) of the O^4-alkyldT lesions in HEK 293T cells and isogenic cells that are deficient in Pol η, Pol ζ, Pol ι or Polκ. The data represent the means and standard errors of the mean of results from three independent replication experiments. $^*P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$. The *P*-values were calculated by using unpaired two-tailed Student's *t*-test.

tively (Figure 4A). The latter result also showed that the two diastereomers of O^4-*s*BudT displayed very similar bypass efficiencies in HEK293T cells (Figure 4A).

Replication studies in the aforementioned TLS polymerase-deficient HEK293T cells provided important insights into the roles of these polymerases in bypassing the O^4-alkyldT lesions. In particular, we found that the bypass efficiencies for all the O^4-alkyldT lesions except O^4-MedT were markedly reduced in HEK293T cells lacking Pol η or Pol ζ. While knockout of Pol κ did not exert any apparent effect on the bypass of any of O^4-alkyldT lesions, knockout of Pol ι led to moderate, albeit statistically significant, drops in the bypass efficiencies of O^4-*n*PrdT and O^4-*n*BudT.

Our results also demonstrated that the O^4-alkyldT lesions are highly mutagenic in HEK293T cells, where these lesions induce exclusively the T→C transition mutation and the

mutation frequencies are affected by the structure of the alkyl group. Specifically, we found that O^4-EtdT exhibited the highest frequency of T→C mutation (82.3%), followed by O^4-MedT (57.7%) and then O^4-*n*PrdT (41.9%) and O^4-*n*BudT (29.2%). While O^4-*n*PrdT and O^4-*i*PrdT (43.8%) displayed very similar frequencies of T→C mutation, O^4-*n*BudT is more mutagenic than the two diastereomers of O^4-*s*BudT, but less mutagenic than O^4-*i*BudT (Figure 4B). In addition, we found that depletion of Pol η, Pol κ or Pol ι did not give rise to significant alterations in the T→C mutation for any of the O^4-alkyldT lesions, with the exceptions that loss of Pol η resulted in significant decreases in this type of mutation for O^4-EtdT and O^4-*n*PrdT and depletion of Pol ι led to a significant drop in this mutation for O^4-*i*BudT (Figure 4B). On the other hand, depletion of Pol ζ gave rise to significant drops in T→C mutation frequencies for all O^4-alkyldT lesions except O^4-MedT and O^4-*n*BudT (Figure 4B).

DISCUSSION

O^4-alkyldT lesions are poorly repaired (8–10); thus, it is important to examine how they are recognized by DNA replication machinery in human cells. This study is the first systematic interrogation about the impact of the O^4-alkyldT lesions on the fidelity and efficiency of DNA replication in human cells. In addition, this represents the first application of the CRISPR-Cas9 genome editing tool for the selective knockout of individual TLS polymerases in cultured human cells. While the siRNA method provides efficient knockdown of TLS polymerase, the presence of residual levels of the TLS polymerase following siRNA knockdown sometimes prevents the clear-cut delineation of the roles of that polymerase in modulating the bypass efficiencies and mutation frequencies of DNA lesions in cells. For instance, previous replication studies from this laboratory and others showed that it is sometimes difficult to distinguish whether the failure in observing a difference in bypass efficiency or mutation frequency after siRNA knockdown of a TLS polymerase is attributed to the lack of involvement of that polymerase in the lesion bypass or due to the presence of remaining TLS polymerase after incomplete knockdown (22,23). The complete knockout of the TLS polymerases by the CRISPR-Cas9 method removes such ambiguity (Figure 1B).

The combination of our quantitative shuttle vector-based method with genetic manipulation of TLS polymerases using the CRISPR-Cas9 method allowed us to make several important conclusions about how the O^4-alkyldT lesions are recognized by DNA replication machinery in human cells (Figure 4). First, we found that the O^4-alkyldT lesions are moderately blocking to DNA replication machinery in human cells (Figure 4A). Second, we found that Pol η and Pol ζ promote the bypass of all the O^4-alkyldT lesions except O^4-MedT in HEK293T cells, whereas depletion of Pol ι or Pol κ did not affect appreciably the bypass efficiencies for any of the O^4-alkyldT lesions (Figure 4A). This finding is in contrast with the observations made in *E. coli* cells, where single depletion of any SOS-induced DNA polymerases did not result in apparent drop in bypass efficiencies for any O^4-alkyldT lesions (14). This may reflect the differences in bacterial and mammalian TLS polymerases.

Third, we found that the O^4-alkyldT lesions only induce one type of mutation, i.e. the T→C transition mutation (Figure 4B). This is consistent with previous replication studies about the same set of O^4-alkyldT lesions in *E. coli* cells and O^4-MedT, O^4-EtdT and O^4-*n*PrdT lesions in mammalian cells (14–19). Our studies with the use of a comprehensive set of O^4-alkyldT lesions also revealed that the frequencies of the T→C mutation are modulated by the length and branching of the alkyl chain conjugated to the O^4 position of thymine (Figure 4B). Furthermore, we found that the genetic knockout of Pol ζ led to marked reduction in T→C mutation for all O^4-alkyldT lesions except O^4-MedT and O^4-*n*BudT (Figure 4B). On the other hand, depletion of Pol η only resulted in significant drop in T→C mutation frequency for O^4-EtdT and O^4-*n*PrdT, and deletion of Pol ι led to a significant drop in T→C mutation frequency for O^4-*i*BudT alone. Moreover, removal of Pol κ did not influence the mutagenic properties of any of the O^4-alkyldT lesions. These results support that the O^4-alkyldT lesions are differentially recognized by TLS polymerases.

It is also worth discussing the findings made from the present study in the context of recent *in-vitro* adduct bypass studies with the use of purified DNA polymerases (35). In this vein, O^4-alkyldT lesions were found to block strongly the primer extension mediated by human Pol κ, Pol ι and *Saccharomyces cerevisiae* Pol ζ, though these lesions could be readily bypassed by human Pol η *in vitro* (35). These results are largely consistent with the observations made in the present study with a few exceptions. First, depletion of Pol η did not appreciably affect the bypass efficiency of O^4-MedT in human cells, suggesting that this lesion may be bypassed by replicative and/or other TLS polymerases in cells. Second, loss of Pol ζ was found to lead to decreases in bypass efficiencies and mutation frequencies for most O^4-alkyldT lesions in human cells. While this could be attributed in part to the differences between *S. cerevisiae* and human Pol ζ, it is also possible that human Pol ζ participates in the extension of the nascent DNA strand after Pol η inserts one or a few nucleotides at and near the lesion site; the efficiencies for such extension may be influenced by the identities of the nucleotides inserted opposite the lesions (i.e. A or G).

Alkylating agents are commonly employed in cancer chemotherapy; however, therapeutic resistance and therapy-derived secondary tumor are among the major challenges in cancer chemotherapy, and both limitations are thought to be driven, in part, by mutations arising from TLS across therapy-induced DNA adducts (36). Thus, understanding the TLS of alkylated DNA lesions may also guide the development of cancer chemotherapy that can minimize therapeutic resistance and secondary tumor development. In this vein, previous cell-based and *in-vivo* studies have demonstrated that genetic depletion of TLS polymerases can improve the efficacy and reduce the therapeutic resistance of several chemotherapeutic agents, including cisplatin, *N*-methyl-*N*'-nitro-*N*-nitrosoguanidine and cyclophosphamide (36–38). Our finding that depletion of polymerases ζ could lead to marked reduction in bypass efficiency and T→C mutation induced by most O^4-

alkyldT lesions suggests the possibility of combination therapy with the use of inhibitors of this TLS polymerase along with DNA alkylating agents. Along this line, it has been demonstrated that inhibitors of poly(ADP-ribose) polymerase were effective in treating those breast cancer patients deficient in BRCA1 or BRCA2 (39,40).

Taken together, our systematic shuttle vector-based study on a group of structurally defined O^4-alkyldT lesions provided important new knowledge about the impact of this group of DNA lesions on the efficiency and accuracy of DNA replication and revealed unambiguously, for the first time, the roles of the TLS polymerases in bypassing these lesions in human cells.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

FUNDING

National Institutes of Health [R01 ES025121]. Funding for open access charge: NIH [ES025121].
Conflict of interest statement. None declared.

REFERENCES

1. Shrivastav,N., Li,D. and Essigmann,J.M. (2010) Chemical biology of mutagenesis and DNA repair: cellular responses to DNA alkylation. *Carcinogenesis*, **31**, 59–70.
2. Gerson,S.L. (2004) MGMT: its role in cancer aetiology and cancer therapeutics. *Nat. Rev. Cancer*, **4**, 296–307.
3. Liu,S. and Wang,Y. (2015) Mass spectrometry for the assessment of the occurrence and biological consequences of DNA adducts. *Chem. Soc. Rev.*, **44**, 7829–7854.
4. Swann,P.F. (1990) Why do O^6-alkylguanine and O^4-alkylthymine miscode? The relationship between the structure of DNA containing O^6-alkylguanine and O^4-alkylthymine and the mutagenic properties of these bases. *Mutat. Res.*, **233**, 81–94.
5. 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.
6. You,C., Wang,J., Dai,X. and Wang,Y. (2015) Transcriptional inhibition and mutagenesis induced by *N*-nitroso compound-derived carboxymethylated thymidine adducts in DNA. *Nucleic Acids Res.*, **43**, 1012–1018.
7. You,C., Wang,P., Dai,X. and Wang,Y. (2014) Transcriptional bypass of regioisomeric ethylated thymidine lesions by T7 RNA polymerase and human RNA polymerase II. *Nucleic Acids Res.*, **42**, 13706–13713.
8. Bronstein,S.M., Skopek,T.R. and Swenberg,J.A. (1992) Efficient repair of O^6-ethylguanine, but not O^4-ethylthymine or O^2-ethylthymine, is dependent upon O^6-alkylguanine-DNA alkyltransferase and nucleotide excision repair activities in human cells. *Cancer Res.*, **52**, 2008–2011.
9. Den Engelse,L., De Graaf,A., De Brij,R.J. and Menkveld,G.J. (1987) O^2- and O^4-ethylthymine and the ethylphosphotriester dTp(Et)dT are highly persistent DNA modifications in slowly dividing tissues of the ethylnitrosourea-treated rat. *Carcinogenesis*, **8**, 751–757.
10. Swenberg,J.A., Dyroff,M.C., Bedell,M.A., Popp,J.A., Huh,N., Kirstein,U. and Rajewsky,M.F. (1984) O^4-ethyldeoxythymidine, but not O^6-ethyldeoxyguanosine, accumulates in hepatocyte DNA of rats exposed continuously to diethylnitrosamine. *Proc. Natl. Acad. Sci. U.S.A.*, **81**, 1692–1695.
11. Godschalk,R., Nair,J., van Schooten,F.J., Risch,A., Drings,P., Kayser,K., Dienemann,H. and Bartsch,H. (2002) Comparison of multiple DNA adduct types in tumor adjacent human lung tissue: effect of cigarette smoking. *Carcinogenesis*, **23**, 2081–2086.
12. Chen,H.J., Wang,Y.C. and Lin,W.P. (2012) Analysis of ethylated thymidine adducts in human leukocyte DNA by stable isotope dilution nanoflow liquid chromatography-nanospray ionization tandem mass spectrometry. *Anal. Chem.*, **84**, 2521–2527.
13. Hecht,S.S. (2008) Progress and challenges in selected areas of tobacco carcinogenesis. *Chem. Res. Toxicol.*, **21**, 160–171.
14. Wang,P., Amato,N.J., Zhai,Q. and Wang,Y. (2015) Cytotoxic and mutagenic properties of O^4-alkylthymidine lesions in Escherichia coli cells. *Nucleic Acids Res.*, **43**, 10795–10803.
15. You,C. and Wang,Y. (2016) Mass spectrometry-based quantitative strategies for assessing the biological consequences and repair of DNA adducts. *Acc. Chem. Res.*, **49**, 205–213.
16. Zhai,Q., Wang,P. and Wang,Y. (2014) Cytotoxic and mutagenic properties of regioisomeric O^2-, $N3$- and O^4-ethylthymidines in bacterial cells. *Carcinogenesis*, **35**, 2002–2006.
17. Altshuler,K.B., Hodes,C.S. and Essigmann,J.M. (1996) Intrachromosomal probes for mutagenesis by alkylated DNA bases replicated in mammalian cells: a comparison of the mutagenicities of O^4-methylthymine and O^6-methylguanine in cells with different DNA repair backgrounds. *Chem. Res. Toxicol.*, **9**, 980–987.
18. Klein,J.C., Bleeker,M.J., Lutgerink,J.T., van Dijk,W.J., Brugghe,H.F., van den Elst,H., van der Marel,G.A., van Boom,J.H., Westra,J.G., Berns,A.J. *et al.* (1990) Use of shuttle vectors to study the molecular processing of defined carcinogen-induced DNA damage: mutagenicity of single O^4-ethylthymine adducts in HeLa cells. *Nucleic Acids Res.*, **18**, 4131–4137.
19. Klein,J.C., Bleeker,M.J., Roelen,H.C., Rafferty,J.A., Margison,G.P., Brugghe,H.F., van den Elst,H., van der Marel,G.A., van Boom,J.H., Kriek,E. *et al.* (1994) Role of nucleotide excision repair in processing of O^4-alkylthymines in human cells. *J. Biol. Chem.*, **269**, 25521–25528.
20. Guo,C., Kosarek-Stancel,J.N., Tang,T.S. and Friedberg,E.C. (2009) Y-family DNA polymerases in mammalian cells. *Cell Mol. Life Sci.*, **66**, 2363–2381.
21. Ohmori,H., Friedberg,E.C., Fuchs,R.P.P., Goodman,M.F., Hanaoka,F., Hinkle,D., Kunkel,T.A., Lawrence,C.W., Livneh,Z., Nohmi,T. *et al.* (2001) The Y-family of DNA polymerases. *Mol. Cell*, **8**, 7–8.
22. You,C., Swanson,A.L., Dai,X., Yuan,B., Wang,J. and Wang,Y. (2013) Translesion synthesis of 8,5′-cyclopurine-2′-deoxynucleosides by DNA polymerases η, ι, and ζ. *J. Biol. Chem.*, **288**, 28548–28556.
23. Weerasooriya,S., Jasti,V.P., Bose,A., Spratt,T.E. and Basu,A.K. (2015) Roles of translesion synthesis DNA polymerases in the potent mutagenicity of tobacco-specific nitrosamine-derived O^2-alkylthymidines in human cells. *DNA Repair*, **35**, 63–70.
24. Cong,L., Ran,F.A., Cox,D., Lin,S., Barretto,R., Habib,N., Hsu,P.D., Wu,X., Jiang,W., Marraffini,L.A. *et al.* (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science*, **339**, 819–823.
25. Mali,P., Yang,L., Esvelt,K.M., Aach,J., Guell,M., DiCarlo,J.E., Norville,J.E. and Church,G.M. (2013) RNA-guided human genome engineering via Cas9. *Science*, **339**, 823–826.
26. Baker,D.J., Wuenschell,G., Xia,L., Termini,J., Bates,S.E., Riggs,A.D. and O'Connor,T.R. (2007) Nucleotide excision repair eliminates unique DNA-protein cross-links from mammalian cells. *J. Biol. Chem.*, **282**, 22592–22604.
27. You,C. and Wang,Y. (2015) Quantitative measurement of transcriptional inhibition and mutagenesis induced by site-specifically incorporated DNA lesions in vitro and in vivo. *Nat. Protoc.*, **10**, 1389–1406.
28. Sakuma,T., Nishikawa,A., Kume,S., Chayama,K. and Yamamoto,T. (2014) Multiplex genome engineering in human cells using all-in-one CRISPR/Cas9 vector system. *Sci. Rep.*, **4**, 5400.
29. Ziegler,K., Bui,T., Frisque,R.J., Grandinetti,A. and Nerurkar,V.R. (2004) A rapid in vitro polyomavirus DNA replication assay. *J. Virol. Methods*, **122**, 123–127.
30. Yuan,B., O'Connor,T.R. and Wang,Y. (2010) 6-Thioguanine and S^6-methylthioguanine are mutagenic in human cells. *ACS Chem. Biol.*, **5**, 1021–1027.
31. Hsu,P.D., Scott,D.A., Weinstein,J.A., Ran,F.A., Konermann,S., Agarwala,V., Li,Y., Fine,E.J., Wu,X., Shalem,O. *et al.* (2013) DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat. Biotechnol.*, **31**, 827–832.
32. Kim,D., Bae,S., Park,J., Kim,E., Kim,S., Yu,H.R., Hwang,J., Kim,J.I. and Kim,J.S. (2015) Digenome-seq: genome-wide profiling of CRISPR-Cas9 off-target effects in human cells. *Nat. Methods*, **12**, 237–243.
33. Newton,C.R., Graham,A., Heptinstall,L.E., Powell,S.J., Summers,C., Kalsheker,N., Smith,J.C. and Markham,A.F. (1989) Analysis of any

point mutation in DNA. The amplification refractory mutation system (ARMS). *Nucleic Acids Res.*, **17**, 2503–2516.
34. Yuan,B., Cao,H., Jiang,Y., Hong,H. and Wang,Y. (2008) Efficient and accurate bypass of N^2-(1-carboxyethyl)-2′-deoxyguanosine by DinB DNA polymerase *in vitro* and *in vivo*. *Proc. Natl. Acad. Sci. U.S.A.*, **105**, 8679–8684.
35. Williams,N.L., Wang,P., Wu,J. and Wang,Y. (2016) In vitro lesion bypass studies of O^4-alkylthymidines with human DNA polymerase η. *Chem. Res. Toxicol.*, **29**, 669–675.
36. Xie,K., Doles,J., Hemann,M.T. and Walker,G.C. (2010) Error-prone translesion synthesis mediates acquired chemoresistance. *Proc. Natl. Acad. Sci. U.S.A.*, **107**, 20792–20797.
37. Wu,F., Lin,X., Okuda,T. and Howell,S.B. (2004) DNA polymerase ζ regulates cisplatin cytotoxicity, mutagenicity, and the rate of development of cisplatin resistance. *Cancer Res.*, **64**, 8029–8035.
38. Doles,J., Oliver,T.G., Cameron,E.R., Hsu,G., Jacks,T., Walker,G.C. and Hemann,M.T. (2010) Suppression of Rev3, the catalytic subunit of Polζ, sensitizes drug-resistant lung tumors to chemotherapy. *Proc. Natl. Acad. Sci. U.S.A.*, **107**, 20786–20791.
39. Farmer,H., McCabe,N., Lord,C.J., Tutt,A.N., Johnson,D.A., Richardson,T.B., Santarosa,M., Dillon,K.J., Hickson,I., Knights,C. *et al.* (2005) Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. *Nature*, **434**, 917–921.
40. Bryant,H.E., Schultz,N., Thomas,H.D., Parker,K.M., Flower,D., Lopez,E., Kyle,S., Meuth,M., Curtin,N.J. and Helleday,T. (2005) Specific killing of BRCA2-deficient tumours with inhibitors of poly(ADP-ribose) polymerase. *Nature*, **434**, 913–917.
41. McLuckey,S.A., Van Berkel,G.J. and Glish,G.L. (1992) Tandem mass spectrometry of small, multiply charged oligonucleotides. *J. Am. Soc. Mass Spectrom.*, **3**, 60–70.