6486–6493 *Nucleic Acids Research, 2017, Vol. 45, No. 11* *Published online 20 April 2017*
doi: 10.1093/nar/gkx266

Replication and repair of a reduced 2′-deoxyguanosine-abasic site interstrand cross-link in human cells

Nathan E. Price[1], Lin Li[1], Kent S. Gates[2,*] and Yinsheng Wang[1,*]

[1]Department of Chemistry, University of California, Riverside, CA 92521-0403, USA and [2]Department of Chemistry, University of Missouri, 125 Chemistry Building, Columbia, MO 65211, USA

Received March 05, 2017; Revised March 24, 2017; Editorial Decision March 27, 2017; Accepted April 13, 2017

ABSTRACT

Apurinic/apyrimidinic (AP) sites, or abasic sites, which are a common type of endogenous DNA damage, can forge interstrand DNA–DNA cross-links via reaction with the exocyclic amino group on a nearby 2′-deoxyguanosine or 2′-deoxyadenosine in the opposite strand. Here, we utilized a shuttle vector method to examine the efficiency and fidelity with which a reduced dG–AP cross-link-containing plasmid was replicated in cultured human cells. Our results showed that the cross-link constituted strong impediments to DNA replication in HEK293T cells, with the bypass efficiencies for the dG- and AP-containing strands being 40% and 20%, respectively. While depletion of polymerase (Pol) η did not perturb the bypass efficiency of the lesion, the bypass efficiency was markedly reduced (to 1–10%) in the isogenic cells deficient in Pol κ, Pol ι or Pol ζ, suggesting the mutual involvement of multiple translesion synthesis polymerases in bypassing the lesion. Additionally, replication of the cross-linked AP residue in HEK293T cells was moderately error-prone, inducing a total of ∼26% single-nucleobase substitutions at the lesion site, whereas replication past the cross-linked dG component occurred at a mutation frequency of ∼8%. Together, our results provided important insights into the effects of an AP-derived interstrand cross-link on the efficiency and accuracy of DNA replication in human cells.

INTRODUCTION

Covalent interstrand cross-links (ICLs) can be induced in DNA by chemotherapeutic drugs such as mechlorethamine and endogenous chemicals like 4-hydroxynonenal (1–6). ICLs prevent strand separation of the DNA double helix and, as a result, are strong impediments to replication and transcription (3,7). When an ICL is encountered during DNA replication, complex repair processes are initiated that lead to unhooking of the cross-link by structure-specific endonucleases such as XPF-ERCC1, MUS81-EME1 and SLX1-SLX3 (Figure 1) (8,9). The unhooked lesion may be bypassed by translesion synthesis (TLS) polymerases (9–11), which are frequently error-prone and may introduce mutations into the genetic code.

An interesting new class of ICLs are formed in duplex DNA from the reaction of an apurinic/apyrimidinic (AP) site, also known as abasic site, with nucleobases on the opposing strand (12–16). AP sites, which arise from the loss of a nucleobase due to spontaneous or enzyme-catalyzed hydrolysis of the *N*-glycosidic bond, are a very common form of endogenous DNA damage (17). Equilibrium amounts of the AP aldehyde generate imine-derived ICLs via reactions with exocyclic amino groups of adenine or guanine residues in the complementary strand (12–16). Although formally reversible, these ICLs are chemically stable in duplex DNA (13,15), as manifested by the observations that the dA–AP ICL blocks DNA replication by the ϕ29 polymerase, which possesses strong helicase-like activity (18), and by the DNA replication machinery in *Xenopus* egg extracts (19).

ICLs can also form between the N^2-amino group of 2′-deoxyguanosine and an AP aldehyde situated on the 3′ side of a cytosine residue (Figure 2) (12–14). The dG–AP ICL can be reduced by $NaCNBH_3$ to yield a cross-link (dG–AP_{red}) that is chemically stable under physiological conditions (Figure 2) (14). Along these lines, it is interesting to note that adducts derived from the reaction of low-molecular weight aldehydes with the N^2-position of dG are subject to similar intracellular reduction, presumably involving agents such as ascorbic acid, NAD(P)H, and glutathione (20).

The ubiquitous occurrence of AP sites in genomic DNA suggests that AP-derived cross-links could be significant in biology, toxicology, and medicine. As a result, it is important to understand how these lesions are processed in hu-

*To whom correspondence should be addressed. Tel: +1 951 827 2700; Fax: +1 951 827 4713; Email: Yinsheng.Wang@ucr.edu
Correspondence may also be addressed to Kent S. Gates. Tel: +1 573 882 6763; Fax: +1 573 882 2754; Email: gatesk@missouri.edu

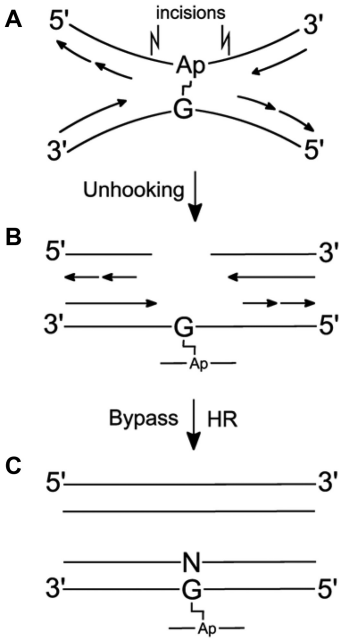

Figure 1. Model for replication-dependent repair of a dG–AP interstrand cross-link lesion. (**A**) Cross-link stalls replication. (**B**) Structure-specific endonucleases unhook the cross-link. (**C**) TLS and extension past the dG–AP adduct remnant (bottom duplex). Homologous recombination repairs the double-strand break (top duplex), and the NER machinery may remove cross-link remnant from the bottom duplex.

man cells, in particular, how these type of cross-link lesions modulate the efficiency and accuracy of DNA replication and how they are repaired. In this study, we constructed a double-stranded plasmid containing a dG–AP_{red} cross-link at a specific site, and assessed the degrees to which the dG–AP_{red} cross-link impedes DNA replication and induces mutations in cultured human cells. We also characterized the roles of TLS polymerases in the replicative bypass of the lesion.

EXPERIMENTAL

Materials

Unmodified oligodeoxyribonucleotides (ODNs) were purchased from Integrated DNA Technologies (Coralville, IA). [γ-^{32}P]ATP was obtained from Perkin Elmer Life Sciences, enzymes were from New England Biolabs (NEB), and chemicals unless otherwise noted were from Sigma-Aldrich. The pSVRLuc plasmid backbone was kindly provided by Professor Orlando Schärer (21). HEK293T cells depleted of Pol η, Pol ι, Pol κ or Pol ζ were obtained by the CRISPR–Cas9 genome editing method, as described previously (22).

Preparation of cross-linked DNA inserts

DNA duplexes containing an AP site at defined locations were prepared using standard procedures (Figure 3A) (23). AP sites were generated by incubation of the annealed Duplex **C** (10 nmol) with uracil DNA glycosylase (UDG, 10 μl, 400 units/ml) in a mixture comprised of UDG buffer

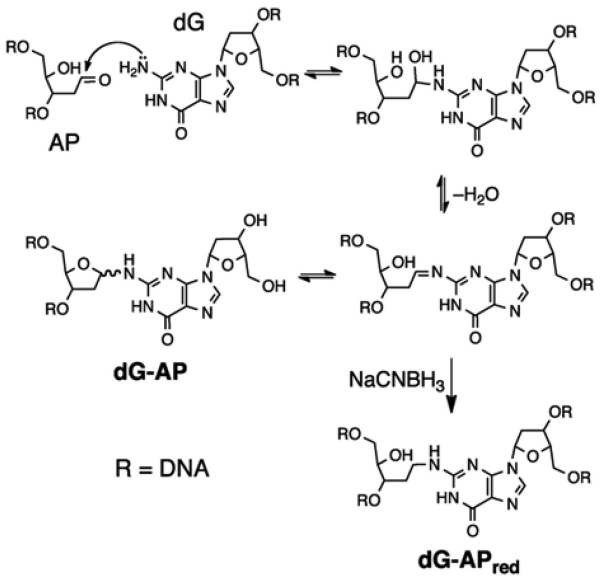

Figure 2. Formation and structure of the dG–AP_{red} cross-link.

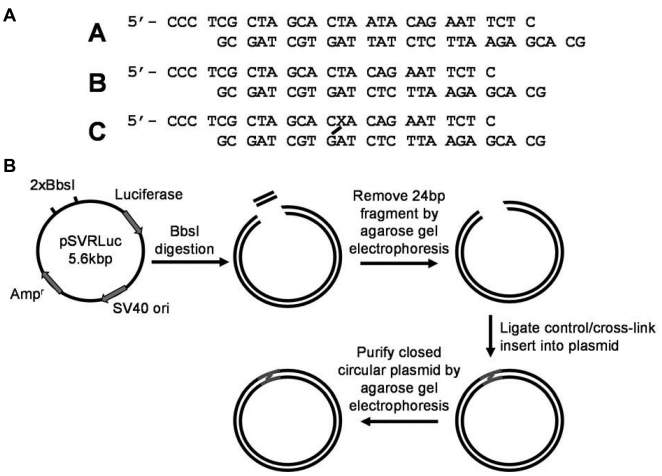

Figure 3. Construction of the dG–AP_{red}-containing plasmid for cellular replication studies. (**A**) Duplexes used for constructing the plasmids for the replication study. Duplexes A, B and C are the inserts used for construction of the competitor, control, and dG–AP_{red} cross-link-bearing plasmids, respectively. 'X' represents uracil in the initial duplex and is converted to an abasic site prior to the cross-linking reaction. (**B**) Workflow for the preparation of the lesion-bearing plasmid.

(30 μl) and H_2O (96 μl) at 37°C for 2 h, followed by extraction with phenol:chloroform:isoamyl alcohol (25:24:1; pH 8) and precipitation with ethanol. The cross-linking reaction was conducted at 37°C for 12 h in a buffer containing sodium acetate (750 mM, pH 5.2) and 200 mM $NaCNBH_3$. The cross-linked duplex was purified on a 20% denaturing polyacrylamide gel. The structure of the cross-link was confirmed by nuclease P1 digestion followed by LC-MS/MS analysis of the resulting cross-link-bearing tetranucleotide (Supplementary Figure S1) (13).

Construction of pSVRLuc genomes harboring a site-specifically inserted dG–AP$_{red}$ (Figure 3b)

Plasmid DNA (100 μg) was mixed with 20 μl 10 × NEB buffer 2, 60 units BbsI (20 000 U/ml), and water to give a 200-μl solution. After incubation at 37°C for 8 h, another aliquot of BbsI (60 units) was added and the mixture was incubated at 37°C overnight. Complete BbsI cleavage was confirmed by BamHI and EarI digestion, followed by PAGE analysis (Supplementary Figure S2) (21). The mixture was extracted with an equal volume of phenol–chloroform–isoamyl alcohol (25:24:1, v/v/v) to remove the BbsI. The aqueous phase was resolved on a 0.5% low-melt agarose gel containing 0.1% ethidium bromide. The band corresponding to the linearized plasmid was cut from the gel and the plasmid recovered using QIAquick gel extraction kit (Qiagen). Ligations were conducted using 1 nM linearized plasmid, 1.5 molar equivalents of control or cross-linked DNA inserts, 1 ml 10 × T4–DNA ligase buffer, 10 μl (4000 units) T4–DNA ligase and water added to give a total reaction volume of 10 ml. The ligation mixture was incubated at 16°C overnight, followed by lyophilization and resuspension in 200 μl water, and then desalted using a G-25 Sephadex column. The desalted solution was again resolved on a 0.5% low-melt agarose gel containing 0.1% ethidium bromide, and the supercoiled plasmid recovered from the gel using QIAquick gel extraction kit. A representative gel including marker lanes is shown in Supplementary Figure S3. The dG–AP$_{red}$ cross-link-containing and control plasmids contained a C:C mismatch to allow for independent analyses of replication products from the cross-linked-AP- and dG-containing strands, whereas no mismatch was incorporated into the competitor plasmid.

To confirm the successful incorporation of the dG–AP$_{red}$-containing duplex DNA into the plasmid, the cross-linked plasmid (200 ng) was digested with BbsI (5 units) and shrimp alkaline phosphatase (1 unit) at 37°C for 1 h, heat inactivated at 80°C for 20 min, and phosphorylated with 5 units of T4 polynucleotide kinase (PNK) and 1.25 μCi (0.5 pmol) of [γ-^{32}P]ATP for 30 min. The T4 PNK was heat inactivated at 80°C for 20 min and finally digested with 20 units of EcoRI-HF. The reaction mixtures were diluted with 30 μl of formamide gel loading buffer. Approximately 5 μl of the mixture was loaded onto a 30% polyacrylamide gel (acrylamide:bis-acrylamide 19:1) and products monitored by phosphorimager analysis (Supplementary Figure S4). The newly constructed lesion-containing double-stranded vectors were then normalized against the lesion-free competitor vector following published procedures (24).

Cell culture, transfection and cellular DNA replication

Replication of the newly constructed genomes was performed using a competitive replication and adduct bypass (CRAB) assay (Figure 4) (24,25). The lesion-bearing and the corresponding non-lesion control plasmids were premixed individually with the competitor genome and transfected into HEK293T cells and the isogenic polymerase-deficient cells. The molar ratios of the competitor to control and lesion-bearing genome were 1:1 and 1:30, respectively. The cells (1 × 10^5) were seeded in 24-well plates and

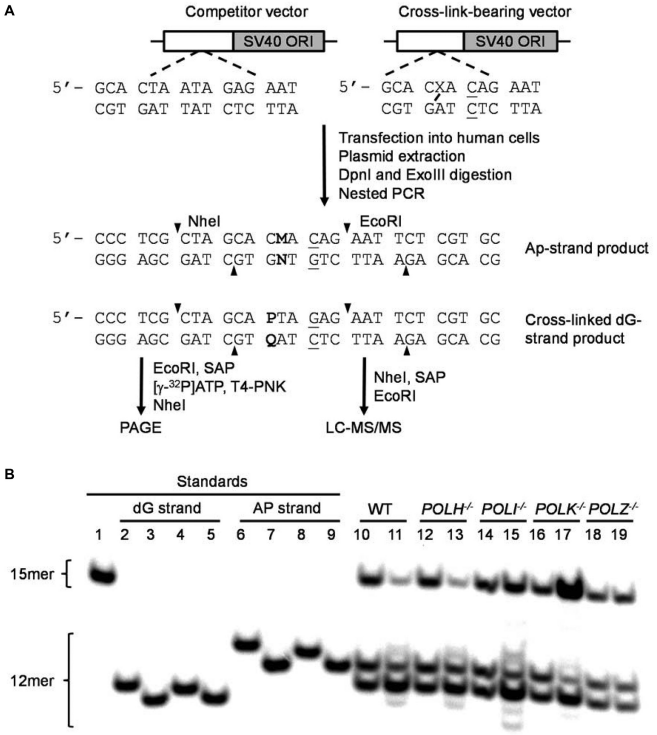

Figure 4. The competitive replication and adduct bypass (CRAB) assay for examining how a dG–AP$_{red}$ crosslink perturbs the efficiency and fidelity of DNA replication in human cells. Shown in (**A**) is a schematic diagram for the assay workflow. Displayed in (**B**) is the results from native PAGE analysis of restriction fragments of replication products from HEK293T cells (WT) and the isogenic polymerase-deficient cells. The bottom strand was radio-labeled. Lane 1, competitor standard 5′-AAT TCT CTA TTA GTG-3′; lanes 2–5, standards of 5′-AAT TCT CTA QTG-3′, with 'Q' being G, A, T and C, respectively; lanes 6–9, standards of 5′-AAT TCT GTN GTG-3′, with 'N' being G, A, T and C, respectively; lanes 10, 12, 14, 16 and 18 correspond to replication products of control and competitor plasmids in parental HEK293T cells (WT), *POLH*$^{-/-}$, *POLI*$^{-/-}$, *POLK*$^{-/-}$ and *POLZ*$^{-/-}$ cells, respectively; lane 11, 13, 15, 17 and 19 correspond to replication products of the dG–AP$_{red}$ and competitor plasmids in parental HEK293T cells (WT), *POLH*$^{-/-}$, *POLI*$^{-/-}$, *POLK*$^{-/-}$ and *POLZ*$^{-/-}$ cells, respectively.

cultured overnight at 37°C in a 5% CO_2 atmosphere, after which they were transfected with 500 ng carrier plasmid and 32 ng of the mixed genomes by using Lipofectamine 2000 (Invitrogen) following the manufacturer's recommended procedures. The cells were harvested at 24 h after transfection, and the progenies of the plasmid were isolated using the Qiagen Spin Kit (Qiagen), following previously described procedures with minor modifications (24,26). The residual unreplicated plasmid was further removed by DpnI digestion, followed by digesting the resulting linear DNA with exonuclease III as described elsewhere (24).

PCR, PAGE and LC–MS/MS analyses

The progeny genomes arising from cellular replication were amplified using a nested PCR protocol with a pair of primers whose products covered the initial lesion site and included nine DpnI recognition sites. The primers were 5′-GAT ATC GCC CTG ATC AAG AGC GAA-3′ and 5′-

CGA GGA AGC GGA TCC AGA CAT GAT-3′, and the PCR amplification started at 98°C for 2 min; then, 35 cycles at 98°C for 30 s, 63°C for 30 s, and 72°C for 30 s; and a final 5-min extension at 72°C. The PCR products were purified using an E.Z.N.A. Cycle Pure Kit (Omega Bio-tek) and a nested PCR step was performed with primers 5′-GGT GGC TAT AAA GAG GTC ATC AGT-3′ and 5′-CGG CCT CGG CCT CTG CAT AAA TAA-3′. The PCR products were again purified using an E.Z.N.A. Cycle Pure Kit and stored at −20°C until use.

For PAGE analysis, 200 ng of the PCR fragments were treated with 20 units of EcoRI-HF and 1 unit of shrimp alkaline phosphatase at 37°C in 2 μl of NEB CutSmart buffer for 1 h, followed by heating at 80°C for 20 min to deactivate the shrimp alkaline phosphatase. To the above mixture were then added 1.25 μCi (0.5 pmol) of [γ-^{32}P]ATP and 5 units of T4 PNK. The reaction was continued at 37°C for 1 h, followed by heating at 85°C for 20 min to deactivate the T4 PNK. To the reaction mixture was added 20 units of NheI-HF, and the solution was incubated at 37°C for 1 h, followed by quenching with 30 μl of formamide gel loading buffer. The resulting restriction digestion yielded a 15mer radiolabeled fragment for the competitor genome, i.e. 5′- p*AAT TCT CTA TTA GTG-3′, and 12mer fragments for the dG–AP-cross-link-bearing genome, i.e. 5′-p*AAT TCT GTN GTG-3′ and 5′-p*AAT TCT CTA QTG-3′ for the cross-linked AP and dG-bearing strands, respectively, where 'N' and 'Q' designate A, T, C or G, and 'p*' indicates the radiolabeled phosphate (Figure 4). The mixture was again resolved using 30% polyacrylamide gel (acrylamide:bis-acrylamide 19:1) and products quantified by phosphorimager analysis. The extent to which the dG–AP$_{red}$ impedes DNA replication was characterized by bypass efficiency (%), calculated from the ratio of (lesion signal/competitor signal)/(non-lesion control signal/competitor signal) (24).

LC-MS/MS was used to identify and quantify the products arising from replication of dG–AP$_{red}$ cross-link-bearing plasmid (Supplementary Figures S5–S17) (27,28). Briefly, 5 μg of PCR products were treated with 40 units of EcoRI-HF in 10 μl of NEB CutSmart buffer at 37°C for 1 h. To the resulting solution was added 40 units of NheI-HF, and the reaction mixture was incubated at 37°C for 1 h, followed by addition of 2 units of shrimp alkaline phosphatase and incubation at 37°C for 1 h. The mixture was extracted with phenol/chloroform/isoamyl alcohol (25:24:1, v/v/v). The aqueous portion was dried with Speed-Vac and ethanol precipitated. The ODN mixture was subjected to LC–MS/MS analysis, as described elsewhere (27,28). An LTQ linear ion-trap mass spectrometer (Thermo Electron) was set up for monitoring the fragmentation of the [M-3H]$^{3-}$ ions of the 12-mer ODNs of the strand initially containing the cross-linked AP site, i.e. 5′-CTA GCA CMA CAG-3′ or the 12-mer ODNs initially harboring the cross-linked dG, i.e. 5′-CTA GCA PTA GAG-3′, where 'M' and 'P' designate A, T, C or G (Supplementary Figures S5–S16). The mutagenic products were quantified by the peak area found in the selected-ion chromatogram (SIC) of each restriction fragment corresponding to the wild-type and mutated products by choosing two unique fragment ions, and finding the peak area ratio of mutant to wild-type product. The peak area ratios were then converted to the corresponding molar ratios by employing calibration curves constructed from the analyses of mixtures of synthetic ODNs in a range of defined molar ratios (Supplementary Figure S17).

RESULTS

Construction of cross-linked and control shuttle vectors

We began with construction of a double-stranded plasmid containing a site-specific dG–AP$_{red}$ (Figure 3) as well as the corresponding control and competitor plasmids by using a protocol adapted from that reported by Schärer and coworkers (21). The three duplex inserts employed for plasmid construction are shown in Figure 3A, where the cross-linked duplex **C**, harboring the dG–AP$_{red}$ lesion, was prepared from the corresponding AP-containing duplex as described previously (13). Duplex **B** is a control, damage-free duplex, with the sequence being the same as Duplex **C** except that the AP-site in **C** is replaced with a thymine in **B**. Duplexes **B** and **C** contained a C:C mismatch, which, as noted previously (24), enabled us to assess independently the replication efficiencies of the two strands. Duplex **A**, which is three base pairs longer than duplex **B** or **C**, was employed for the construction of the competitor plasmid. The competitor plasmid, which is co-transfected with the dG–AP$_{red}$-bearing plasmid or control plasmid in cellular replication experiments, serves as an internal reference for determining the degree to which the dG–AP$_{red}$ ICL impedes DNA replication in cells. The three duplexes contained two different sticky ends that can be ligated into the linearized plasmid following dual cleavages by BbsI. The successfully ligated, supercoiled plasmid was purified from the ligation mixture using agarose gel electrophoresis, and the ligation yields were ∼10% (Supplementary Figure S3).

The amounts of the control and lesion-containing genomes were normalized against that of the competitor genome by digestion with BbsI and alkaline phosphatase, followed by radiolabeling the nascent 5′ termini with T4-PNK and a final cleavage with EcoRI. PAGE analysis of the resulting products showed a 24 mer for the unmodified duplex (which does not contain EcoRI or NheI restriction sites and will therefore not produce a 12 mer band in the final analysis), a 21 mer band for the competitor, a 18 mer band for the control, and a cross-linked 18 mer band for the cross-linked plasmid (Supplementary Figure S4).

Effect of the dG–AP$_{red}$ cross-link on the efficiency and fidelity of DNA replication in human cells

The dG–AP$_{red}$ cross-link-containing and control genomes were mixed individually with the competitor genome and transfected into HEK293T cells. We chose to use the HEK293T cells for the replication experiments because these cells were transformed with SV40 large T antigen, which supports the replication of the SV40 origin-bearing plasmids that are used in the present study. After 24 h, the progeny genomes were extracted, and a 1500 bp region surrounding the cross-link was amplified by nested PCR. The final PCR amplicon (500 bp) was digested with NheI and

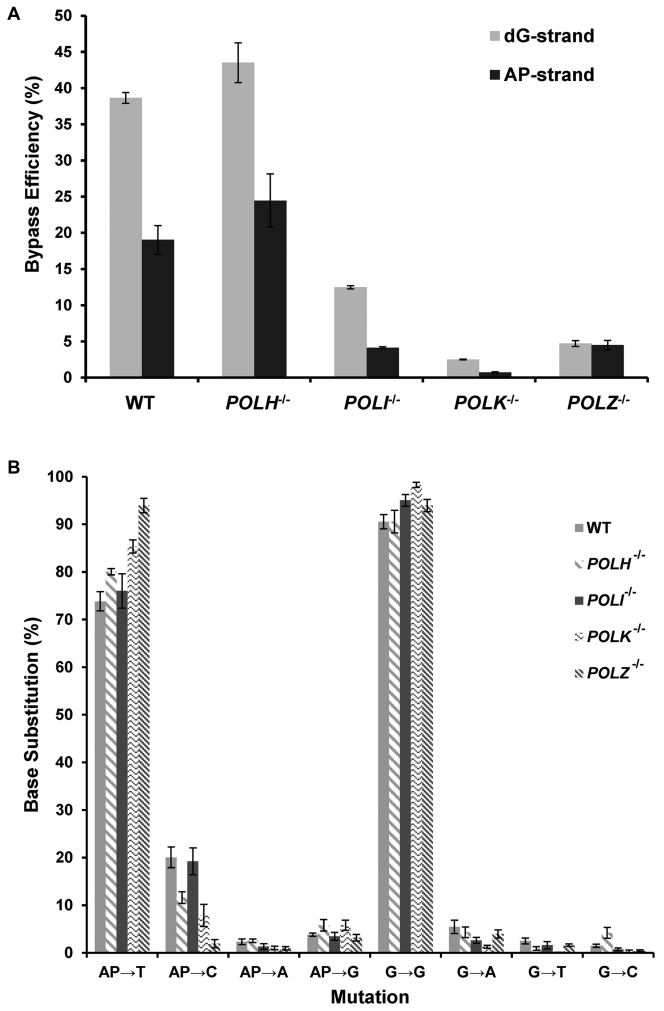

Figure 5. Bypass efficiencies (**A**) and frequencies of base substitutions (**B**) observed at crosslinked AP and dG sites after the dG–AP_{red} interstrand cross-link was replicated in HEK293T cells (WT) and the isogenic cells depleted of Pol η, Pol κ, Pol ι or Pol ζ. The data represent the means and standard deviations of results obtained from three independent transfection experiments.

EcoRI, which yielded a ^{32}P-labeled 12 mer fragment for the PCR products of progeny of the lesion-containing vector corresponding to the region where the dG–AP_{red} ICL was initially situated, and a 15 mer for the competitor vector (Figure 4). Bypass efficiency, which reflects the extent to which cellular repair systems unhook the ICL and the degree to which the unhooked dG–AP_{red} impedes DNA replication, was determined by comparing the intensities of the 15 mer competitor band with the 12 mer lesion bands (Figure 4). Our results showed that, in HEK293T cells, the dG–AP_{red} ICL decreases substantially the replication efficiencies, which were approximately 40% and 20% for the strands harboring the cross-linked dG and AP site, respectively (Figure 5A).

Multiple bands were observed in the region of the gel where the 12 mer products migrate (Figure 4). This suggested that replication in wild-type HEK293T cells generated progenies with multiple, distinct sequences in the region surrounding the ICL as a result of error-prone lesion bypass. Sequences of the progeny were identified and quantified by LC–MS/MS (Supplementary Figures S5–S17). The PCR amplification product was digested with NheI and EcoRI (Figure 4 and Supplementary Figure S8), and the digestion mixture was subjected to LC–MS and MS/MS analyses, where the fragmentations of the $[M-3H]^{3-}$ ions of the eight 12 mer products with all four possible canonical nucleotides being incorporated at the initially cross-linked dG or AP sites were monitored. The MS/MS of each product displayed a series of 'w_n' and 'a_n—Base' ions as shown in Supplementary Figures S9–S16, and the nomenclature for fragment ions follows that described previously (29). We plotted the selected-ion chromatograms (SICs) by monitoring the formation of two fragment ions from each ODN (e.g. the w_5 and w_6^{2-} ions in Supplementary Figure S9B). The results showed that replication of the dG–AP_{red} ICL yields progenies primarily bearing a T-residue at the site where the AP side of the cross-link was initially located (AP→T, 74 ± 2%, Figure 5B). Smaller, but clearly detectable amounts of AP→C (10–20%), AP→G (4–6%), and AP→A (1–2%) were observed at this site in the progeny (Figure 5B). Replication primarily generates progeny genomes that have retained a G-residue at the site of the dG cross-link attachment in the dG–AP_{red} ICL (G→G, 91 ± 2%, Figure 5B). Low frequencies of G→A (2–5%), G→T (1–2%) and G→C (1%) mutations were observed at this site. Clearly, replication past the cross-linked AP-site was more mutagenic than that past the cross-linked guanine residue.

The roles of Pol η, Pol κ, Pol ι or Pol ζ in the replication across the dG–AP_{red} cross-link in cells

As described in the Introduction, TLS mediated by Pol η, Pol κ, Pol ι and/or Pol ζ enables cells to tolerate cross-links in their genome, but this may come at the expense of errors introduced into the genetic code. To explore these issues, the replication of cross-linked DNA in HEK293T cells was compared to that in the isogenic cells depleted of Pol η, Pol κ, Pol ι or Pol ζ, where the polymerase-deficient cells were generated by CRISPR–Cas9 and confirmed by genomic DNA sequencing along with Western blot analysis (22). After replication, progeny genomes were isolated and analyzed as described above to determine replication bypass efficiency and fidelity.

The deletion of Pol η in HEK293T cells did not exert significant effects on relative bypass efficiencies. In contrast, deletion of any one of the other three TLS polymerases (i.e. Pol κ, Pol ι and Pol ζ) caused substantial decreases in replication efficiency. Similar to what we observed for the wild-type HEK293T cells, the dG strand of the cross-linked vector was replicated more efficiently than the AP strand in *POLH*$^{-/-}$, *POLK*$^{-/-}$ and *POLI*$^{-/-}$ cells, whereas the two strands were replicated at nearly the same efficiencies in *POLZ*$^{-/-}$ cells (Figure 5A).

The mutations resulting from replication of the dG–AP_{red} ICL in *POLH*$^{-/-}$, *POLI*$^{-/-}$, *POLK*$^{-/-}$, *POLZ*$^{-/-}$ cells are, for the most part, similar to those found in the wild-type cells (Figure 5B). Perhaps the most striking difference was in the *POLK*$^{-/-}$ and *POLZ*$^{-/-}$ cells, where, compared to the parental HEK293T cells, there is a 10–20% increase

in the progeny bearing a T-residue at the initial AP site of the dG–AP_{red} vector (Figure 5B). This is accompanied by a concomitant decrease in T → C mutation found in the two polymerase-deficient backgrounds.

DISCUSSION

There are many structurally diverse types of ICLs. Differences in the chemical structures and reactivities of various cross-links may evoke different unhooking and bypass mechanisms; however, the structure–activity relationships governing the fate of ICL processing in cells are not well understood. Given the relevance of ICLs in biology and medicine, it is important to explore how variations in cross-link structure lead to differences in their biological properties.

AP-derived cross-links may be biologically important and also constitute a structurally novel class of ICLs. Most ICLs can be defined as residing in either the major or minor groove of duplex DNA. For example, the prototypical cross-links derived from nitrogen mustards and mitomycin C bridge the two guanine residues on the opposing DNA strands through the major-groove N^7 and the minor-groove N^2 positions, respectively (1,3). In contrast, the dG–AP_{red} cross-link studied here and the dA–AP cross-link both originate on the backbone of one strand, but terminate on the other strand at the minor-groove N^2 position of guanine (Figure 2) and the major-groove N^6 position of adenine, respectively (13,15). Three-dimensional structures of the AP-derived ICLs are not yet available; nevertheless, the cross-linkages in these lesions transect the helical axis and, in this regard, are distinct from both minor- and major-groove ICLs.

Here, we developed a shuttle vector method to examine the effects of the dG–AP_{red} cross-link on the efficiency and fidelity of DNA replication in human cells. We found that the ICL significantly decreased DNA replication efficiency in HEK293T cells. Interestingly, the two strands of the cross-linked vector were not replicated at equal efficiencies, where the AP-strand (20%) was replicated at a significantly lower efficiency than the dG strand (40%).

Importantly, our approach enabled us to explore the translesion synthesis step of the dG–AP_{red} ICL repair. A number of mechanisms were previously shown to be involved in the repair of the ICL preceding replication, including: (i) Replication-independent repair (9), (ii) replication traverse of the ICL involving FANCM (30), (iii) replication-dependent repair involving the Fanconi anemia pathway and unhooking by structure-specific endonucleases (Figure 1) (8,9), and, (iv) replication-dependent repair involving unhooking by a DNA glycosylase (19). The glycosylase-dependent pathway seems less likely here because there are no known glycosylase enzymes displaying activity on the glycosidic bond of N^2-alkyl-dG residues like the one found in the dG–AP_{red} cross-link studied here. On the other hand, *Escherichia coli* AlkB was shown to be capable of removing alkyl group from N^2-alkylguanine in DNA *in vitro* (31), though such activity has yet been demonstrated in cells (32). It would be interesting to examine whether the human orthologs of AlkB (e.g. ALKHB2 and ALKBH3) are able to unhook the dG–AP_{red} cross-link through an oxidative dealkylation mechanism.

All currently proposed mechanisms for ICL repair include a common step involving TLS past an unhooked cross-link remnant (e.g. Figure 1) (10), and previous biochemical studies showed that translesion synthesis across DNA interstrand cross-link lesions depends highly on their structures. For instance, Ho *et al*. (33) showed that the TLS efficiency across several major-groove N^7–N^7 guanine interstrand cross-links is affected by the length of the cross-link bridging the two N^7 positions, with the longer and flexible interstrand cross-links being more readily bypassed. In addition, *in-vitro* studies showed that Pol ι misincorporated thymine nucleotide across the psoralen interstrand cross-link and Pol κ extended the mispaired primer termini, indicating that DNA Pol ι and Pol κ may serve as inserter and extender, respectively, in the TLS past the psoralen DNA interstrand cross-links (34). Moreover, Pol η exhibited a moderate efficiency and a high error rate when bypassing an oxidized AP site-derived interstrand cross-link; however, human Pol κ, Pol ι, Pol *ν*, REV1, and yeast pol ζ stalled at the position before the cross-linked nucleotide (35). Pol ζ, in complex with Rev1, was recently found to be responsible for extension after the insertion step of TLS across a cisplatin-induced DNA interstrand cross-link (11), and cells deficient in Pol ζ are hypersensitive to cross-linking agents (36,37). Consistent with the critical role of TLS in ICL repair, we found that replication efficiency is dramatically decreased in isogenic cells that are deficient in Pol ι, κ or ζ, whereas no significant alteration in replication efficiency was observed in cells lacking Pol η (Figure 5A). The observation that deficiency in any *one* of three polymerases (i.e. Pol ι, κ, or ζ) reduces markedly the replication efficiency suggests that these proteins work together to bypass the cross-link remnant. The mutual involvement of Pols ι and κ in the bypass of the dG–AP_{red} ICL is reminiscent of earlier work showing that both polymerases are essential for the high-fidelity bypass of the minor-groove N^2-(1-carboxyethyl)-2′-deoxyguanosine lesions in mammalian cells (38). In light of the previously reported role of the Rev1–Pol ζ complex in the extension step of TLS past the cisplatin-induced ICL (11), we reason that Pol ζ, perhaps in conjunction with Rev1, may also function in the extension step of TLS across the dG–AP_{red} crosslink. In addition, our observation that the replication efficiency was not significantly altered in the isogenic cells lacking Pol η is consistent with precedents showing that this enzyme does not catalyze the bypass of minor-groove N^2-alkyl-dG lesions (38). The requirement of multiple TLS polymerases (i.e. Pols ι, κ and ζ) seen in the present work parallels previous studies showing that the replication across oxidatively induced 8,5′-cyclopurine 2′-deoxynucleosides in human cells involves Pols η, ι and ζ (24), and that the bypass of TT cyclobutane pyrimidine dimers in Pol η-deficient human cells requires the cooperation of Pols ι, κ and ζ (39).

TLS can be error-prone (mutagenic); hence, it is important to assess the fidelity associated with replication across the dG–AP_{red} ICL. We found that replication past the guanine residue in the cross-link in wild-type HEK293T cells was ∼90% error-free, with dCMP being predominantly incorporated opposite the cross-linked nucleobase

(Figure 5B). The AP residue in the cross-link primarily directs dAMP insertion, giving rise to progenies carrying a T residue (74%) at this location (Figure 5B). The dAMP incorporation can be considered as error-free replication if one envisions that the endogenous AP-derived cross-link arises mainly from the misincorporation of a dU across from an adenine residue, followed by the generation of an AP site by the action of uracil DNA glycosylase (40).

We propose a model to rationalize all the data presented above. First, stalling of a replication fork at the cross-link results in activation of the Fanconi anemia pathway, which promotes the unhooking of the ICL. Although little is known about the factors that dictate the locations of incisions during unhooking, we envision that unhooking of the cross-link by selective incisions on the AP-strand gives rise to an N^2-alkyl-dG cross-link remnant (Figure 1), which is subsequently bypassed in a largely error-free manner by the combined action of Pols ι, κ and ζ. Precedents support the involvement of Pol κ in the error-free bypass of N^2-alkyl-dG lesions (28,38,41,42), an acrolein-derived DNA-peptide cross-link formed on N^2 position of guanine (43), and a model acrolein-derived N^2-N^2-guanine interstrand cross-link (44). In addition, Pol ι was found to be essential for the accurate bypass of some N^2-alkyl-dG lesions in mammalian cells (38). In the cross-linked vector, the G-containing strand is replicated at a higher efficiency (40%) than the AP-containing strand (20%). This perhaps can be rationalized from the suggestion that homologous recombination, which is required for the repair of the AP-strand in this model, fails to engage or succeed in approximately half the cases where TLS processes successfully complete replication of the dG-strand in the cross-linked vector.

In summary, our results showed that the dG–AP_{red} cross-link constitutes a strong block to DNA replication, and the TLS step of the cross-link repair involves three polymerases (i.e. Pol κ, Pol ι and Pol ζ). The shuttle vector technology described here, coupled with genetic manipulation using CRISPR–Cas9, provides a powerful approach for studying the repair and biological consequences of ICL lesions in human cells. It will be interesting to use this methodology to gain detailed insight regarding the repair and replicative bypass of other AP-derived cross-links, as well as other cross-links formed from endogenous processes or cancer chemotherapeutic agents. Detailed understanding of repair mechanisms of drug-derived cross-links, in combination with genotyping of individual cancer patients, may enable personalized therapies that deploy specific cross-linking agents against genetically susceptible cancers. Characterizing the mechanisms by which endogenous cross-links are repaired will help define genetic predispositions to cancer and aging-related illnesses.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

The authors would like to thank Professor Orlando Schärer for providing the pSVRLuc plasmid, and Drs. Changjun You and Zhijie Chen for helpful discussion.

FUNDING

National Institutes of Health (NIH) [ES021007 and ES018827]. Funding for open access charge: NIH [ES021007].
Conflict of interest statement. None declared.

REFERENCES

1. Povirk,L.F. and Shuker,D.E. (1994) DNA damage and mutagenesis induced by nitrogen mustards. *Mutat. Res.*, **318**, 205–226.
2. Lawley,P.D. and Phillips,D.H. (1996) DNA adducts from chemotherapeutic agents. *Mutat. Res.*, **355**, 13–40.
3. Deans,A.J. and West,S.C. (2011) DNA interstrand crosslink repair and cancer. *Nat. Rev. Cancer*, **11**, 467–480.
4. Nejad,M.I., Johnson,K.M., Price,N.E. and Gates,K.S. (2016) A new cross-Link for an old cross-linking drug: the nitrogen mustard anticancer agent mechlorethamine generates cross-links derived from abasic sites in addition to the expected drug-bridged cross-links. *Biochemistry*, **55**, 7033–7041.
5. Wang,H., Kozekov,I.D., Harris,T.M. and Rizzo,C.J. (2003) Site-specific synthesis and reactivity of oligonucleotides containing stereochemically defined 1,N^2-deoxyguanosine adducts of the lipid peroxidation product trans-4-hydroxynonenal. *J. Am. Chem. Soc.*, **125**, 5687–5700.
6. Huang,H., Kozekov,I.D., Kozekova,A., Wang,H., Lloyd,R.S., Rizzo,C.J. and Stone,M.P. (2010) DNA cross-link induced by trans-4-hydroxynonenal. *Environ. Mol. Mutagen.*, **51**, 625–634.
7. Niedernhofer,L.J., Lalai,A.S. and Hoeijmakers,J.H. (2005) Fanconi anemia (cross)linked to DNA repair. *Cell*, **123**, 1191–1198.
8. Zhang,J. and Walter,J.C. (2014) Mechanism and regulation of incisions during DNA interstrand cross-link repair. *DNA Repair*, **19**, 135–142.
9. Clauson,C., Scharer,O.D. and Niedernhofer,L. (2013) Advances in understanding the complex mechanisms of DNA interstrand cross-link repair. *Cold Spring Harb. Perspect. Biol.*, **5**, a012732.
10. Roy,U. and Scharer,O.D. (2016) Involvement of translesion synthesis DNA polymerases in DNA interstrand crosslink repair. *DNA Repair*, **44**, 33–41.
11. Budzowska,M., Graham,T.G., Sobeck,A., Waga,S. and Walter,J.C. (2015) Regulation of the Rev1-pol ζ complex during bypass of a DNA interstrand cross-link. *EMBO J.*, **34**, 1971–1985.
12. Dutta,S., Chowdhury,G. and Gates,K.S. (2007) Interstrand cross-links generated by abasic sites in duplex DNA. *J. Am. Chem. Soc.*, **129**, 1852–1853.
13. Catalano,M.J., Liu,S., Andersen,N., Yang,Z., Johnson,K.M., Price,N.E., Wang,Y. and Gates,K.S. (2015) Chemical structure and properties of interstrand cross-links formed by reaction of guanine residues with abasic sites in duplex DNA. *J. Am. Chem. Soc.*, **137**, 3933–3945.
14. Johnson,K.M., Price,N.E., Wang,J., Fekry,M.I., Dutta,S., Seiner,D.R., Wang,Y. and Gates,K.S. (2012) On the formation and properties of interstrand DNA–DNA cross-links forged by reaction of an abasic site with the opposing guanine residue of 5′-CAp sequences in duplex DNA. *J. Am. Chem. Soc.*, **135**, 1015–1025.
15. Price,N.E., Catalano,M.J., Liu,S., Wang,Y. and Gates,K.S. (2015) Chemical and structural characterization of interstrand cross-links formed between abasic sites and adenine residues in duplex DNA. *Nucleic Acids Res.*, **43**, 3434–3441.
16. Price,N.E., Johnson,K.M., Wang,J., Fekry,M.I., Wang,Y. and Gates,K.S. (2014) Interstrand DNA–DNA cross-link formation between adenine residues and abasic sites in duplex DNA. *J. Am. Chem. Soc.*, **136**, 3483–3490.
17. Nakamura,J. and Swenberg,J.A. (1999) Endogenous apurinic/apyrimidinic sites in genomic DNA of mammalian tissues. *Cancer Res.*, **59**, 2522–2526.
18. Yang,Z., Price,N.E., Johnson,K.M. and Gates,K.S. (2015) Characterization of interstrand DNA–DNA cross-Links derived from abasic sites using bacteriophage φ29 DNA polymerase. *Biochemistry*, **54**, 4259–4266.
19. Semlow,D.R., Zhang,J., Budzowska,M., Drohat,A.C. and Walter,J.C. (2016) Replication-dependent unhooking of DNA interstrand cross-links by the NEIL3 glycosylase. *Cell*, **167**, 498–511.

20. Fang,J.L. and Vaca,C.E. (1995) Development of a ^{32}P-postlabelling method for the analysis of adducts arising through the reaction of acetaldehyde with 2′-deoxyguanosine-3′-monophosphate and DNA. *Carcinogenesis*, **16**, 2177–2185.
21. Enoiu,M., Ho,T.V., Long,D.T., Walter,J.C. and Scharer,O.D. (2012) Construction of plasmids containing site-specific DNA interstrand cross-links for biochemical and cell biological studies. *Methods Mol. Biol.*, **920**, 203–219.
22. Wu,J., Li,L., Wang,P., You,C., Williams,N.L. and Wang,Y. (2016) Translesion synthesis of O^4-alkylthymidine lesions in human cells. *Nucleic Acids Res.*, **44**, 9256–9265.
23. Lindahl,T., Ljungquist,S., Siegert,W., Nyberg,B. and Sperens,B. (1977) DNA *N*-glycosidases: properties of uracil-DNA glycosidase from *Escherichia coli*. *J. Biol. Chem.*, **252**, 3286–3294.
24. 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.
25. Delaney,J.C. and Essigmann,J.M. (2006) Assays for determining lesion bypass efficiency and mutagenicity of site-specific DNA lesions in vivo. *Methods Enzymol.*, **408**, 1–15.
26. 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.
27. Hong,H., Cao,H. and Wang,Y. (2007) Formation and genotoxicity of a guanine cytosine intrastrand cross-link lesion in vivo. *Nucleic Acids Res.*, **35**, 7118–7127.
28. 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.
29. 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.
30. Huang,J., Liu,S., Bellani,M.A., Thazhathveetil,A.K., Ling,C., de Winter,J.P., Wang,Y., Wang,W. and Seidman,M.M. (2013) The DNA translocase FANCM/MHF promotes replication traverse of DNA interstrand crosslinks. *Mol. Cell*, **52**, 434–446.
31. Li,D., Fedeles,B.I., Shrivastav,N., Delaney,J.C., Yang,X., Wong,C., Drennan,C.L. and Essigmann,J.M. (2013) Removal of *N*-alkyl modifications from N^2-alkylguanine and N^4-alkylcytosine in DNA by the adaptive response protein AlkB. *Chem. Res. Toxicol.*, **26**, 1182–1187.
32. Shrivastav,N., Fedeles,B.I., Li,D., Delaney,J.C., Frick,L.E., Foti,J.J., Walker,G.C. and Essigmann,J.M. (2014) A chemical genetics analysis of the roles of bypass polymerase DinB and DNA repair protein AlkB in processing N^2-alkylguanine lesions *in vivo*. *PLoS One*, **9**, e94716.
33. Ho,T.V., Guainazzi,A., Derkunt,S.B., Enoiu,M. and Scharer,O.D. (2011) Structure-dependent bypass of DNA interstrand crosslinks by translesion synthesis polymerases. *Nucleic Acids Res.*, **39**, 7455–7464.
34. Smith,L.A., Makarova,A.V., Samson,L., Thiesen,K.E., Dhar,A. and Bessho,T. (2012) Bypass of a psoralen DNA interstrand cross-link by DNA polymerases β, ι, and κ in vitro. *Biochemistry*, **51**, 8931–8938.
35. Xu,W., Ouellette,A., Ghosh,S., O'Neill,T.C., Greenberg,M.M. and Zhao,L. (2015) Mutagenic bypass of an oxidized abasic lesion-induced DNA interstrand cross-link analogue by human translesion synthesis DNA polymerases. *Biochemistry*, **54**, 7409–7422.
36. McHugh,P.J., Sones,W.R. and Hartley,J.A. (2000) Repair of intermediate structures produced at DNA interstrand cross-links in Saccharomyces cerevisiae. *Mol. Cell. Biol.*, **20**, 3425–3433.
37. Sarkar,S., Davies,A.A., Ulrich,H.D. and McHugh,P.J. (2006) DNA interstrand crosslink repair during G1 involves nucleotide excision repair and DNA polymerase ζ. *EMBO J.*, **25**, 1285–1294.
38. Yuan,B., You,C., Andersen,N., Jiang,Y., Moriya,M., O'Connor,T.R. and Wang,Y. (2011) The roles of DNA polymerases κ and ι in the error-free bypass of N^2-carboxyalkyl-2′-deoxyguanosine lesions in mammalian cells. *J. Biol. Chem.*, **286**, 17503–17511.
39. Ziv,O., Geacintov,N., Nakajima,S., Yasui,A. and Livneh,Z. (2009) DNA polymerase ζ cooperates with polymerases κ and ι in translesion DNA synthesis across pyrimidine photodimers in cells from XPV patients. *Proc. Natl. Acad. Sci. U.S.A.*, **106**, 11552–11557.
40. Guillet,M. and Boiteux,S. (2003) Origin of endogenous DNA abasic sites in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.*, **23**, 8386–8394.
41. Suzuki,N., Ohashi,E., Kolbanovskiy,A., Geacintov,N.E., Grollman,A.P., Ohmori,H. and Shibutani,S. (2002) Translesion synthesis by human DNA polymerase κ on a DNA template containing a single stereoisomer of dG-(+)- or dG-(-)-*anti*-N^2-BPDE (7,8-dihydroxy-*anti*-9,10-epoxy-7,8,9,10-tetrahydrobenzo[*a*]pyrene). *Biochemistry*, **41**, 6100–6106.
42. Jarosz,D.F., Godoy,V.G., Delaney,J.C., Essigmann,J.M. and Walker,G.C. (2006) A single amino acid governs enhanced activity of DinB DNA polymerases on damaged templates. *Nature*, **439**, 225–228.
43. Minko,I.G., Yamanaka,K., Kozekov,I.D., Kozekova,A., Indiani,C., O'Donnell,M.E., Jiang,Q., Goodman,M.F., Rizzo,C.J. and Lloyd,R.S. (2008) Replication bypass of the acrolein-mediated deoxyguanine DNA-peptide cross-links by DNA polymerases of the DinB family. *Chem. Res. Toxicol.*, **21**, 1983–1990.
44. Minko,I.G., Harbut,M.B., Kozekov,I.D., Kozekova,A., Jakobs,P.M., Olson,S.B., Moses,R.E., Harris,T.M., Rizzo,C.J. and Lloyd,R.S. (2008) Role for DNA polymerase κ in the processing of N^2-N^2-guanine interstrand cross-links. *J. Biol. Chem.*, **283**, 17075–17082.