Published in final edited form as:

Mutat Res. 2013 ; 0: 1–10. doi:10.1016/j.mrfmmm.2013.02.005.

DNA-protein crosslinks processed by nucleotide excision repair and homologous recombination with base and strand preference in *E.coli* model system

Qingming Fang^{*,1}

Department of Cellular and Molecular Physiology, The Pennsylvania State University College of Medicine, P.O. Box 850, Hershey, Pennsylvania 17033 USA

Abstract

Bis-electrophiles including dibromoethane and epibromohydrin can react with *O*⁶-alkylguanine-DNA alkyltransferase (AGT) and form AGT-DNA crosslinks *in vitro* and *in vivo*. The presence of human AGT (hAGT) paradoxically increases the mutagenicity and cytotoxicity of bis-electrophiles in cells. Here we establish a bacterial system to study the repair mechanism and cellular responses to DNA-protein crosslinks (DPCs) *in vivo*. Results show that both nucleotide excision repair (NER) and homologous recombination (HR) pathways can process hAGT-DNA crosslinks with HR playing a dominant role. Mutation spectra show that HR has no strand preference but NER favors processing of the DPCs in the transcribed strand; UvrA, UvrB and Mfd can interfere with small size DPCs but only UvrA can interfere with large size DPCs in the transcribed strand processed by HR. Further, we found that DPCs at TA deoxynucleotide sites are very inefficiently processed by NER and the presence of NER can interfere with these DNA lesions processed by HR. These data indicate that NER and HR can process DPCs cooperatively and competitively and NER processes DPCs with base and strand preference. Therefore, the formation of hAGT-DNA crosslinks can be a plausible and specific system to study the repair mechanism and effects of DPCs precisely *in vivo*.

Keywords

DNA-protein crosslinks; Nucleotide excision repair; MGMT; Dibromoethane; Epibromohydrin

1. Introduction

DNA-protein crosslinks (DPCs) are common DNA lesions generated by endogenous and exogenous agents, including the enzymes topoisomerase I, topoisomerase II and DNA

© 2013 Elsevier B.V. All rights reserved.

^{*} Corresponding author. Address correspondence to: Qingming Fang, Tel 412-623-7807; fax: 412-623-7761. fangq@upmc.edu or qingmingfang@yahoo.com.

¹ Present address: Department of Pharmacology & Chemical Biology, University of Pittsburgh School of Medicine and Hillman Cancer Center, University of Pittsburgh Cancer Institute, Research Pavilion, 5117 Centre Avenue, Pittsburgh, PA 15213, USA

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Conflict of interest statement

The author declares no conflict of interest.

Supplementary Data.

Supplementary data including Supplementary Figures S1-2 and Supplementary Tables S1-3 are available at Mutation Research online.

polymerase β , endogenous aldehyde metabolites, physical factors such as ionizing radiation and UV light, chemical agents including formaldehyde and transition metals, and bifunctional chemotherapeutic drugs such as nitrogen mustards and platinum compounds^{1; 2; 3; 4}. The bulky nature of DPCs blocks normal physiological processes such as replication, transcription, DNA repair and chromatin remodeling¹. Since all known DPC-inducing agents can produce a variety of DNA lesions in addition to DPCs, it is difficult to precisely evaluate the cellular effects of DPCs as well as the DNA repair mechanisms required to process DPCs *in vivo*². In order to understand the mechanism of DPC formation and repair, several *in vitro* model systems have been developed^{5; 6; 7; 8}. Repair studies with these systems suggested that NER could efficiently remove small size DNA-peptide crosslinks^{9; 10}. Genetic and biochemical studies have shown that both NER and HR could remove DPCs and both pathways play different roles. In bacteria, NER repairs DPCs that contain proteins which are less than 15kDa whereas oversized DPCs are processed by HR¹¹. In mammalian cells, NER does not contribute to the repair of DPCs unless the proteins are <10 kDa whereas HR plays a pivotal role to process DPCs^{9; 12}. However, previous model systems cannot study the heterogeneity of DNA repair of DPCs at DNA base level and the precise effect of DPCs *in vivo*.

The repair protein O⁶-alkylguanine-DNA alkyltransferase (AGT or MGMT) protects the cells from the genotoxic effects caused by endogenous and exogenous alkylating agents^{13; 14}. Unexpectedly, AGT can enhance the cytotoxic and mutagenic effects induced by dibromoethane (DBE) in prokaryotic and eukaryotic cells^{15; 16; 17; 18}. Some bis-electrophiles including DBE and epibromohydrin (EBH) are environmental toxicants that have been used globally as pesticides, solvents and chemical intermediates in agriculture and industry^{18; 19}. It has been shown that DBE can react directly with human AGT (hAGT) to form a highly reactive half-mustard intermediate and then hAGT can facilitate the binding to DNA and form hAGT-DNA crosslinks^{18; 20}. *In vitro* experiments showed that the order of bases forming covalent hAGT-oligodeoxyribonucleotide (Oligo) complexes is G>T>C>A and the predominant crosslinked site is N⁷-G^{18; 20; 21}. Other epoxide compounds of bis-electrophiles such as EBH and 1,3-butadiene diepoxide (BDO) can produce AGT-DNA adducts via two mechanisms: (a) compounds react with AGT to form intermediates and then intermediates react with DNA to form AGT-DNA adducts; or (b) compounds react with DNA to form intermediates and then these intermediates react with AGT to form AGT-DNA adducts^{21; 22; 23}. The presence of hAGT can increase the mutagenicity and cytotoxicity of EBH and BDO in *E.coli* cells^{22; 23}. Human histones and glyceraldehyde 3-phosphate dehydrogenase are also identified as candidate proteins that can form DPCs with bis-electrophiles-diepoxymethane but not DBE as determined in a global proteomic screen besides AGT and glutathione (GSH). However, neither of them will enhance mutagenesis *in vivo*^{24; 25}. AGT is very likely the only nuclear protein that can enhance mutagenesis induced by DBE and EBH *in vivo*. The formation of hAGT-DNA crosslinks in cells may provide a good model to study the cellular response and precise repair mechanism of DPCs *in vivo*. Previous studies based on UV-induced cyclobutane pyrimidine dimer repaired by NER have discovered that the repair mechanism of NER and transcription-coupled NER (TC-NER) which prefers to repair the transcribed DNA strand²⁶. While a study in *Escherichia coli* (*E.coli*) suggested that TC-NER is not involved in the repair mechanism of DPCs²⁷. It is still unknown whether NER processes DPCs with strand preference. As DPCs are dominantly formed at GC pairs and processed by both HR and NER pathways in our model system, it may provide some novel data that has not been shown in previous studies. We generated several cell strains via knockout of the UvrA, UvrB, RecA and Mfd genes in GWR109 cells that harbor a deletion of the endogenous AGT genes *Ogt* and *Ada*. These cell strains were treated with DBE and EBH. Our data show that both NER and HR can remove hAGT-DNA crosslinks yet HR plays a dominant role. Mutation spectra show that NER cannot efficiently remove DPCs at TA sites and the presence of NER can obviously prevent

DPCs at TA sites and transcribed strand processed by HR. This study indicates that NER can remove DNA lesions with base and strand preference and NER can compete with HR to remove DPCs.

2. Materials and Methods

2.1 Materials

Oligodeoxyribonucleotides and all the primers were synthesized and purified by the Macromolecular Core Facility, Hershey Medical Center. *E. coli* XL1-blue bacterial strain was purchased from Stratagene (La Jolla, CA). Wild type *E. coli* and *E. coli* bacteriophage P1 were obtained from ATCC (Manassas, VA). QIAquick PCR Purification kit, QIAprep Spin Miniprep kit, Blood & Cell Culture DNA Maxi kit, the pQE30 plasmid and Anti-penta His antibody were obtained from Qiagen (Chatsworth, CA). Pfx DNA polymerase was purchased from Invitrogen (Carlsbad, CA). Ampicillin, kanamycin, tetracycline, chloramphenicol, isopropyl β -D-thiogalactopyranoside, L-arabinose, DBE, EBH, PMSF, rifampicin and most other biochemical reagents were purchased from Sigma Chemical Co. (St. Louis, MO). KOD DNA polymerase was purchased from Novagen (Madison, WI). HRP linked anti-mouse IgG and anti-rabbit IgG were purchased from Cell signal technology (Danvers, MA). Plasmids pKD46 and pKD3 were provided by Dr. B. Wanner (Department of Biological Sciences, Purdue University, West Lafayette, IN).

2.2 Bacterial strains and media

The bacterial strains GWR109, FC326, FC218, CJM1 as well as CJM2 with the deletion of the endogenous AGT genes *Ada* and *Ogt* were generously provided by Dr. L. Samson (Biological Engineering Division and Center for Environmental Health Sciences, MIT, Cambridge, MA)^{28; 29}. The BW25113 cell strain lacking exonuclease V of the RecBCD pathway was provided by Dr. B. Wanner³⁰. CJM1 and CJM2 with inactive *UvrB* gene were derived from FC326 and FC218 respectively. FC326 and FC218 cells with a chromosomal deletion of the lactose operon carry mutated lacZ sequences at codon 461(GAG) on an F' episome. These *E. coli* cells with the mutated lacZ alleles are lac⁻ and cannot grow in M9 minimum medium with lactose as a sole carbon source; however, the lac⁻ cells can be reverted to the lac⁺ cells (wild type) via AT to GC (FC326: AAG to GAG) and GC to AT (FC218: GGG to GAG) mutations respectively and the mutated cells can grow in the medium with lactose as the sole carbon source^{28; 29}. Cell mediums were prepared as described previously^{29; 31}.

2.3 Knockout UvrA, UvrB, RecA and Mfd in GWR109 cell strain

Gene disruption was carried out as described previously³⁰. Plasmid pKD3 containing chloramphenicol (Cam)-resistance gene and FLP recognition target site was used as template. Primers UvrA-P1: UvrA-P2, UvrB-P1: UvrB-P2, RecA-P1: RecA-P2, and Mfd-P1: Mfd-P2 containing FLP sequence, ribosome binding site and targeted genes sequence were used for PCR to obtain UvrA, UvrB, RecA and Mfd PCR products respectively and their sequence is listed in Supplementary Table S1. PCR was carried out using KOD DNA polymerase and the PCR products were digested by Dpn I for 1 h at 37°C and purified with QIAquick Gel Extraction kit. PCR products were electroporated into competent BW25113 cells and transformants were selected in LB plates containing ampicillin (Amp) and Cam. BW25113 cells with targeted gene disruption were used as donor strain for P1 transduction.

P1 transduction was performed using the protocol as described previously with modifications³² and <http://biology4.wustl.edu/levin/protocols.php>. Positive colonies were isolated and verified by PCR and UV sensitivity assay.

2.4 PCR verification of gene knockout

The correct gene disruption for all the mutants was confirmed by PCR. The isolated positive colonies were suspended in 50 μ l H₂O, 5 μ l portions were taken from them and used for 50 μ l PCR reaction mixture. Two pairs of primers were used for each set of disrupted genes. The first pair of primers were named TP1-1 and TP1-2 and their sequence was designed to recognize DNA outside of the modification of the knockout genes. The PCR reaction can produce the correct fragments of targeted genes and replaced sequences. The second pair of primers were named TP2-1 and TP2-2 and their sequence lies within the modified sequence of knockout genes. So if the targeted genes were knocked out, then the correct size PCR product cannot be obtained. The sequence of primers is listed in Supplementary Table S1. 1% Agarose gel with ethidium bromide was used to identify the PCR products.

2.5 Determination of AGT expression

The plasmids pQE30 or pQE-hAGT (which has an N-terminal (His)₆-tag replacing the terminal M- with the sequence MRGS(H)₆GS-) ³³ with pREP4 were co-transformed into cell strains including GWR109, GWR109-UvrA- (UA-), GWR109-UvrB- (UB-), GWR109-Mfd- (Mfd-) and GWR109-RecA- (RA-). The expression level of hAGT in these cell strains was determined by Western blot as described previously ³⁴. For the transformation of FC326, FC218, CJM1 and CJM2 cells with pQE30 or pQE-hAGT or pQE-Ogt and the expression level of AGT proteins in these cell strains see reference ³¹.

2.6 Measurement of cell survival and mutations in *E. coli* cell strains

GWR109, UA-, UB-, Mfd- and RA-cells transformed with plasmid pQE30 or pQE-hAGT were used to measure cell survival and mutations as described previously ^{20; 22; 35}. Diluted cell suspension was plated onto LB plates supplemented with respective antibiotics to measure cell survival and undiluted cell suspension was plated onto LB plates with 100 μ g/ml rifampicin and other antibiotics to estimate mutation frequency. Mutagenesis assay in FC326, FC218, CJM1 and CJM2 cell strains with plasmids pQE30 or pQE30-hAGT or pQE30-Ogt was carried out as described previously ^{29; 31}. The plates were incubated for 72 h (M9 minimal medium) at 37°C, the colonies were scored and the mutation frequency was expressed as the number of mutants per 10⁸ survivors.

2.7 Mutation spectra analysis of rifampicin-resistant mutants

In order to analyze the mutation spectra induced by DBE and EBH at the *rpoB* gene locus in different cell strains, rifampicin-resistant cell clones from 4-5 repeated experiments were randomly selected and suspended in 50 μ l of deionized water and vortexed vigorously. 5 μ l of these suspensions were used as DNA templates in PCR. The highly mutable region of *rpoB* gene was amplified by PCR with Pfx DNA polymerase. Primers sequence is listed in Supplementary Table S1. For the PCR cycling conditions see reference ²². The purified PCR products were sequenced.

2.8 Analysis of crosslinked proteins

Genomic DNA of cells was prepared with Qiagen blood and cell culture DNA Maxi kit. 0.2mg/ml RNase A and 3mg/ml lysozyme were added to the lysis buffer but protease K was excluded to keep the crosslinked proteins. The detailed procedures can be found in the handbook provided by Qiagen company. As it is unsuccessful to extract genomic DNA from RA-cells with the Qiagen blood and cell culture DNA Maxi kit, CsCl ultracentrifugation was used to extract genomic DNA for repair assay. The procedure was performed as described previously ¹¹. To study the size distribution of crosslinked proteins in cells treated with DBE or EBH, 10 μ g DNA isolated from cells expressing no hAGT or 2.5 μ g DNA isolated from cells expressing hAGT was digested with DNase I at 37°C for 1 h. SYPRO

Ruby Protein Gel Stain was used to detect the presence of crosslinked proteins in SDS-PAGE gels. Western blot was used to determine the presence of crosslinked hAGT with antibodies ATO-1 (which can recognize amino acid 8-20 of hAGT, prepared by our lab) and Anti-penta His tag while crosslinked GSH was detected with antibody D8 (Enzo Life Science).

3. Results

3.1 Generation of UA-, UB-, Mfd- and RA- cell strains

To study the role of NER and HR in the processing of DPCs, we generated several cell strains deficient in these DNA repair pathways. The mechanism of gene knockout is chromosomal genes in recombination-proficient *E. coli* cells lacking exonuclease V of the RecBCD pathway can be replaced by linear DNA^{30; 36}. UvrA, UvrB, Mfd and RecA gene, with 2.8, 2.0, 3.5 and 1.0 kb size respectively, were disrupted in BW25113 cells by replacing with a 1.1 kb PCR product containing Cam resistance gene. The replaced genes in BW25113 cell strain were transferred into GWR109 cells by P1 transduction. The GWR109 cells with correct gene knockout were confirmed by PCR with two pairs of primers and UV sensitivity assay. The correct size of gene sequence can be detected by PCR with primers TP1-1 and TP1-2 in normal cells (Fig.S1. A: RA-: lane 1; UA-: lane 6; UB-: lane 10 and Mfd-: lane 14) whereas a 1.1 kb fragment was detected in cells with gene disruption (Fig. S1. A: RA-: lane 2; UA-: lane 7; UB-: lane 11 and Mfd-: lane 15). No PCR product can be detected in cells with correct gene disruption (Fig. S1. A: RA-: Lane 3; UA-: lane 8; UB-: lane 12 and Mfd-: lane 16) and a slightly smaller size of the targeted gene fragment can be detected in normal cells (Fig. S1. A: RA-: Lane 4; UA-: lane 9; UB-: lane 13 and Mfd-: lane 17) with primers TP2-1 and TP2-2. The results indicate that all the derivatives of the GWR109 cell strain contain the correct gene disruption. The UV sensitivity assay showed no colony in cells with UvrA, UvrB or RecA knockout while a few colonies in cells with Mfd knockout and a considerable number of colonies in GWR109 cells could be detected after these cell strains were exposed to UV light. This experiment showed that these cell strains possess a repair deficiency resulting from the specific gene knockout.

Plasmids pQE-30 and pQE-hAGT with pREP4 were co-transferred into GWR109 cells and established derivative cells. hAGT expression level in these cell strains was determined with Western blot with antibody-Anti-penta His and no obvious difference can be detected (Fig. S1.B).

3.2 DPCs are formed in cells treated with DBE and EBH and removed by HR and NER in vivo

It has been shown that hAGT and GSH can react with DNA to form DPCs in the presence of DBE and EBH *in vitro*^{18; 22; 37; 38}. In order to confirm these DPCs can be formed *in vivo*, we used two methods to extract DPCs. SYPRO Ruby Protein Gel Stain and Western blot were used to determine the size distribution of crosslinked proteins and specific crosslinked protein respectively. Without DBE or EBH treatment, no crosslinked proteins can be detected in cells (lane 5 in Fig. S2.A, lane 1 in Fig. S2.B, lane 3 and lane 6 in Fig. S2.C, lane 1 in Fig. 1.C); with DBE or EBH treatment, crosslinked proteins can be detected in cells with or without hAGT expression (Fig. S2 and Fig. 1). Results show that two methods can successfully extract DPCs without significant contamination of noncrosslinked proteins (Fig. S2 and Fig. 1). Only one band can be detected by Western blot in samples with less DPCs formation (Fig. S2 B and Fig. 1A and 1C), this indicates that the DNA in DPCs can very efficiently be digested by DNase I in our experimental condition and our methods are reliable. In cells without hAGT expression treated with DBE or EBH, the size of the most crosslinked proteins are small (<10 KD) (lanes 6-9 in Fig. S2. A) and the result of Western

blot confirmed crosslinked GSH can be determined (Fig. S2.B). In cells expressing hAGT, most of the crosslinked proteins are more than 10 kDa (Fig. S2.C). As the number of base pairs between crosslinked proteins will determine whether DNaseI possesses enough space to bind to and cut the DNA between them efficiently, more DPCs formed in cells, more bands with large sizes detected (Fig. S2 and Fig.1). SYPRO Ruby Protein Gel Stain can only be used to estimate the size distribution of crosslinked proteins. Dependent on the intensity of DPCs, the bands corresponding to specific crosslinked protein can be detected at any higher position than the normal band position. As most of the bands of DPCs in cells expressing no hAGT are smaller than 10kDa (Fig. S2.A), Western blot can detect the crosslinked GSH (Fig. S2.B) and the MW of other potential crosslinked proteins is more than 10kDa²⁵, this suggests that these DPCs determined in cells expressing no hAGT are mainly DNA-GSH crosslinks. In cells expressing hAGT treated with DBE or EBH, the size of most crosslinked proteins are larger (>10 kDa) than in cells expressing no hAGT (Fig. S2.C), and large amount of crosslinked hAGT can be determined (Fig. 1A and 1C). Currently, only GSH and AGT have been found to react with DBE to form DPCs in *E.coli*. As GWR109 and derived cells have no endogenous AGT expression, the large size DPCs formed in cells expressing hAGT are mainly hAGT-DNA crosslinks. Our result shows that more DPCs are formed in cells treated with 1mM EBH than 10mM DBE *in vivo* (Fig. S2.A-B and Fig. 1.A-B). This result is reasonable, for previous reports have shown that EBH can react with hAGT more efficiently and formed intermediates are more stable³⁹. The reason will be explained in detail in next section.

Reports showed that both NER and HR can mitigate the genotoxic effects induced by DPCs, it is due to the formed DPCs processed by NER and HR^{9; 11; 12}. Determining their removal can provide direct evidence to explain genotoxic effects. GWR109, UA- and Mfd-cells treated with 10mM DBE for different time course of repair incubation show that more than 70% of hAGT-DNA crosslinks can be removed in 1h repair incubation and Mfd- cells can remove DPCs as efficiently as GWR109 cells (Fig. 1.A-B). The result of SYPRO Ruby Protein Gel Stain also shows that more than 90% of crosslinked proteins are removed in 1h repair incubation in both Mfd- and GWR109 cells treated with 10mM DBE (Fig. S2.C). Therefore we select 1h repair incubation as the time course to study the role of NER and HR in processing DPCs. For 1h repair incubation, approximately 90%, 88%, 80% and 65% of crosslinked hAGT were released in GWR109, Mfd-, UA- and UB-cells respectively while less than 25% released in RA- cells (Fig. 1), this shows that HR can efficiently process the DPCs whereas NER has a very limited capacity to process the crosslinks and Mfd deficiency cannot significantly affect the process of DPCs compared to GWR109 cells. Results of SYPRO Ruby Protein Gel Stain show more DPCs can be detected in RA-cells than in Mfd-cells and GWR109 cells (Fig. S2.C) after 1h repair incubation, it also confirms that HR plays a dominant role in processing large size DPCs *in vivo*. For small size GSH-DNA crosslinks formed in cells expressing no hAGT, both NER and HR can efficiently release crosslinked GSH (Fig. S2.A-B).

3.3 Both HR and NER are involved in mitigating genotoxic effects of DPCs induced by DBE and EBH

In order to understand how the large size hAGT-DNA crosslinks produced *in vivo* affect the genotoxic effects, we treated GWR109 cells and derived GWR109 cells with TC-NER (Mfd-) or NER (UA-, UB-) or HR (RA-) deficiency with DBE or EBH. Results show that DBE and EBH mediated toxicity in GWR109 cells and its derivatives are greatly increased by the expression of hAGT in a dose-dependent manner (Fig. 2 and 3). This result is consistent with previous reports^{20; 22}. In cells without hAGT expression, 10mM DBE treatment can only increase mutation frequency 6-9 fold and cell death 1.2-3.5 fold in HR and NER deficient cells compared to GWR109 cells (Fig. 2.A-B). The genotoxic effects in

RA-cells is slightly more potent than UA-, UB- and Mfd-cells. However, compared to GWR109 cells, 1mM EBH treatment can induce very potent genotoxic effects in RA-cells but not in UA-, UB- or Mfd-cells (Fig. 3.A-B). These results suggest that EBH treatment can produce DNA lesions that can be processed by HR efficiently but not NER, and DBE treatment can produce a limited amount of DNA damage that can be processed by both HR and NER efficiently. For RA-cells expressing no hAGT treated with 10mM DBE and 1mM EBH, the cell survival is decreased by 70% and 97%, and the mutation frequency is increased by 9-fold and 243-fold respectively (Fig. 2.A-B and Fig. 3.A-B). This shows that HR deficiency yields a much more significant effect on the toxicity caused by EBH than DBE. In *E.coli*, GSH conjugation can activate DBE and EBH to form DPCs and induce the mutagenic effect^{37; 38; 40; 41}. These small size peptide-DNA crosslinks can be efficiently processed by both NER and HR^{9; 10; 11} (Fig. S2.A-B). There is no evidence to suggest that DBE can directly react with DNA but the epihalohydrins including EBH can induce a wide variety of DNA damages directly^{20; 22; 42; 43}. Due to the stability of DNA adducts induced by EBH, several of these adducts undergo chemical transformations such as depurination and deamination to produce secondary lesions such as apurinic sites, ring-opened products and uracil⁴⁴. These secondary lesions play an important role to cause genotoxic effects of EBH and they are mainly processed by the base excision repair (BER) and HR but not NER⁴⁵. This difference may provide a reasonable explanation for the result.

Previous studies have shown that hAGT can react with DNA to form hAGT-DNA crosslinks in the presence of DBE and EBH *in vitro*^{18; 22}. Our experiments *in vivo* also confirmed that high level of hAGT-DNA crosslinks could be formed in cell strains expressing hAGT treated with DBE and EBH (Fig. 1.). The cell sensitivity to DPCs can obviously be affected by the status of the NER and HR pathways: cell survival is 1.1 ± 0.3 , $5.0 \times 10^{-3} \pm 4.0 \times 10^{-3}$, $6.0 \times 10^{-6} \pm 3.0 \times 10^{-6}$ and the mutation frequency is 9.7 ± 3.8 , $3.5 \times 10^2 \pm 1.4 \times 10^2$, $3.1 \times 10^4 \pm 1.4 \times 10^4$ in 10^6 surviving GWR109 cells, UA-cells and RA-cells expressing hAGT treated with 10mM DBE respectively (Fig. 2.C-D). Both NER and HR deficiency can greatly decrease the survival and increase the mutation frequency, and RA-cells are much more sensitive. As the genotoxic effects induced by DBE can more efficiently be prevented in GWR109 cells with the presence of both NER and HR than in HR proficient UA- cells or NER proficient RA-cells, it shows that NER and HR can process hAGT-DNA crosslinks cooperatively with HR playing a dominant role. This genotoxic effects is consistent with the capability of HR and NER to process DPCs determined in the previous section and shown in Fig. 1. The genotoxic effects of NER and HR in cells treated with EBH is consistent with cells treated with DBE, and the cells are much more sensitive to EBH treatment (Fig. 3.C-D). Previous reports show that DBE can form DPCs only at the C145 site and DPCs are formed through a highly unstable intermediate produced by reacting hAGT with DBE firstly. However, EBH can form DPCs at both C150 and C145 sites and DPCs can also be formed through a stable intermediate produced by reacting with DNA directly^{22; 39}. EBH can react with hAGT more efficiently and form intermediates that are more stable³⁹. Our data confirms that cells treated with 1mM EBH can form more hAGT-DNA crosslinks than 10mM DBE *in vivo* (Fig. S2 and Fig. 1). The formation of more DPCs results in the reduced cell survival and higher mutation frequency.

Compared to GWR109 cells, mutation frequency but not cell survival in Mfd-cells containing small or large size DPCs (Fig.S2.A,C and Fig.1. A-B) is significantly changed (Fig. 2 and 3), it also shows the involvement of TC-NER in DNA repair of DPCs. As Mfd-cells can process DPCs as efficiently as GWR109 cells *in vivo* (Fig. S2.C and Fig. 1.A-B) and cell survival isn't significantly changed in Mfd- cells compared to GWR109 cells (Fig. 2.A, C and Fig. 3.A, C), it suggests that cell survival is not a very sensitive biomarker to evaluate the genotoxic effects induced by DPCs.

3.4 Mutation spectra suggest that NER processes DPCs with base preference and the presence of UvrA interferes with DPCs formed in the transcribed strand and TA sites of any strand processed by HR

In *E.coli*, gene mutation depends on the activated DNA repair pathways induced by SOS response. Mutation spectra can show the strand or the base pair of DPCs that cannot be efficiently repaired by error-free repair pathways to avoid base mutations. Previous studies have shown that TC-NER prefers to repair transcribed DNA strand and strand specific repair is associated with strand specific mutation^{26; 46}. We measured the mutation spectra at the β subunit of RNA polymerase (rpoB) locus to evaluate the DNA repair and cellular responses. The expression of hAGT caused a large increase in mutation frequency in response to DBE and EBH (Fig. 2.D, 3.D). Table S2 lists the mutation types, mutation sites, base changes and amino acid changes of the mutation spectra and Table S3 lists the mutation frequency, total analyzed samples, total mutants and mutation number of specific base pair. The mutation spectra in GWR109 cells (Table 1) are similar to those reported previously²⁰. A previous study showed that the mutations at G:C sites are produced by hAGT crosslinked with G base of DNA¹⁸. In *E.coli*, GSH conjugation can also activate DBE to form DPCs (Fig. S2.B) and enhance the mutagenesis of DBE^{37; 38; 40}. Although 5-6 adducts crosslinked with GSH can be detected *in vitro*, only GSH crosslinked with N⁷-G and N¹-A adducts can be detected in *E.coli* cells treated with DBE, and GSH crosslinked with G is dominant (>95%)^{47; 48; 49; 50}. These DNA lesions can induce GC to AT and AT to GC mutations^{15; 51}. Mutations at GC sites are due to the G crosslinked with GSH and a similar trend can also be examined in AGT-DNA crosslinks induced by DBE and EBH^{18; 20; 22; 39}. In this paper we will neglect the potential trivial contribution from GSH or AGT crosslinked with C base of DNA for the presence of GSH and AGT crosslinked with C base cannot be detected *in vitro* and *in vivo*^{18; 47; 48; 49; 50}. So we define a mutation in the transcribed strand if a mutation detected at G base of the transcribed strand and we define a mutation in the non-transcribed strand if a mutation detected at C base of the transcribed strand.

Mutation spectra of mutants in cell strains without hAGT expression caused by 10mM DBE is shown in Table 1. Currently, only GSH and AGT have been found to react with DBE to form DPCs in *E.coli*. Our result also shows that small size peptides can react with DBE to form DPCs (Fig. S2. A) and these peptide-DNA crosslinks can be efficiently processed by both NER and HR (Fig. S2.A-B and Fig. 2.A-B). Since GWR109 cells without the expression of Ogt and Ada and the small size crosslinked peptide can be recognized by anti-GSH antibody (Fig. S2.B), the majority of the crosslinked peptide are GSH. Mutation spectra show that dominant mutations at GC pairs were only detected in HR proficient cells and there is no difference between mutations that occurred at GC and AT pairs in RA-cells (Table 1). The relative mutation ratio at TA sites in NER proficient cell strains RA- (48%) and GWR109 (24%) is significantly higher than that in NER deficient cell strains UA- (4%) and UB- (10%) ($p < 0.05$), and in RA-cells is significantly higher than in GWR109 cells ($p < 0.05$) (Table 1). The result indicates the relative mutation ratio at AT sites induced by small size DPCs can be significantly increased by the presence of NER but decreased by the presence of HR, suggesting that DNA-peptide crosslinks of AT sites are processed by NER very inefficiently and the presence of NER may interfere with these DNA lesions processed by HR. Compared to GWR109 cells, the relative mutation ratio at AT sites in UA-cells and UB-cells but not Mfd- cells can significantly be decreased. This indicates that GC-NER but not TC-NER can interfere with the process of the DPCs at AT sites by HR. Mutation spectrum also show the relative mutation ratio in the transcribed strand in Mfd-, UA- and UB-cells is significantly higher than in GWR109 cells (Table 1), suggesting TC-NER plays an important role to process small size of DPCs in the transcribed strand.

The mutation spectra in cell strains with hAGT expression treated with 10mM DBE were listed in Table 1. Compared to cells expressing no hAGT, the presence of hAGT can

significantly reduce the relative mutation ratio at TA sites and significantly increase the ratio of mutations at GC sites in GWR109 and RA-cells ($p < 0.05$) (Table 1). As HR pathway can efficiently process the DPCs at TA sites and the formation of hAGT-DNA crosslinks at TA sites is very rare, the change of relative mutation ratio at TA sites isn't significant in UA-, UB- and Mfd- cells compared with GWR109 cells (Table 1). In GWR109 cells, 79% (23/29) and 21% (6/29) mutations are at G base of the transcribed and non-transcribed strand respectively (Table 1 and Table S3). This shows that cells with both NER and HR proficiency prefer to process the DNA lesions in the non-transcribed strand. In UA- cells, there are 52% and 48% mutations at G base of the transcribed and non-transcribed strand respectively (Table 1) and the relative mutation ratio is not significantly different between transcribed and non-transcribed strand, suggesting that HR has no preference to process the DPCs of both strands in these cells. However, there are 27% (7/26) and 73% (19/26) mutations occurred at G base of the transcribed and non-transcribed strand in RA-cells respectively (Table 1 and Table S3), it shows that cells with HR deficiency can significantly decrease the relative mutation ratio at G base of the transcribed strand, suggesting that NER prefers to process the hAGT-DNA crosslinks in the transcribed strand. Compared to GWR109 cells, the mutations in the transcribed strand decreased significantly in UA-cells but not in Mfd- or UB-cells (Table 1). This shows that the presence of UvrA may interfere with the hAGT-DNA crosslinks in the transcribed strand processed by HR. The relative mutation ratio at G base of the transcribed strand in GWR109 cells expressing no hAGT (11%) is significantly lower than that in cells expressing hAGT (79%) (Table 1). This difference may be due to formed small size DPCs but not large size DPCs can be efficiently repaired by NER. The reason will be explained in detail in discussion part.

As described previously, EBH can directly react with DNA to form DNA adducts and produce secondary DNA lesions through chemical transformation. These secondary DNA lesions including apurinic sites, ring-opened products and uracil⁴⁴ can efficiently cause mutations and they are good substrates of BER and HR but not NER⁴⁴. The mutation spectrum of cells expressing no hAGT induced by EBH cannot show the formation and repair of DPCs by NER. The mutation spectrum also support this conclusion (Table 1). In cells expressing hAGT treated by EBH, formed hAGT-DNA crosslinks can induce more than 1,000-fold higher mutation frequency than that induced by EBH reacting with DNA directly (Fig. 3). The mutation spectrum induced by EBH and hAGT interaction can show the cellular response to DPCs. The mutation spectrum in cells expressing hAGT treated with EBH also show that cells with both NER and HR proficiency prefer to process the large size DPCs in the non-transcribed strand, HR has no base preference while NER prefers to process large size DPCs in the transcribed strand in HR deficient cells and the presence of UvrA may interfere with the hAGT-DNA crosslinks in the transcribed strand processed by HR (Table 1). This indicates that the similar result in cells treated with 10mM DBE can be proved in cells treated with 1mM EBH (Table 1).

3.5 The presence of NER significantly increases T:A to C:G mutations in cells treated with DBE

As hAGT-DNA crosslinks are rarely formed at TA sites by reacting with DBE and hAGT, the mutation spectrum in cells expressing hAGT treated by DBE and EBH cannot clearly confirm the finding that NER cannot efficiently repair the DPCs at TA sites induced by DBE and the presence of NER impedes DPCs at TA sites processed by HR, which suggested by the mutation spectrum in cells expressing no hAGT treated with DBE. In order to confirm the finding, plasmids pQE-Ogt, pQE-hAGT and pQE30 with pREP4 were co-transferred into FC326, FC218, CJM1 and CJM2 cells. There is no significant difference in the expression level of Ogt and hAGT among these cell strains³¹. Cells were treated with 0-10mM DBE and the survival and mutation frequency were obtained in M9 minimal

medium containing glucose or lactose respectively. Results show: (1) In cells expressing hAGT, the mutation frequency of TA to C:G mutation is very low and it is not significantly higher in FC326 cells than in CJM1 cells (Fig. 4A). This finding is consistent with the mutation spectra result (Table 1) and previous results showing that hAGT has a very limited capacity to form crosslinks with T of oligos in the presence of DBE¹⁸. (2) The mutation frequency of G:C to A:T mutation in NER-proficient FC218 cells expressing Ogt is significantly lower than that in NER-deficient CJM2 cells expressing Ogt, suggesting that Ogt-DNA crosslinks at GC sites can be processed by NER efficiently. However, the mutation frequency of T:A to C:G in NER-proficient FC326 cells expressing Ogt is significantly higher than that in NER-deficient CJM1 cells expressing Ogt (Fig. 4). This finding confirms that NER cannot efficiently repair DPCs at TA sites and the presence of NER can interfere with the process of these DPCs by HR.

4. Discussion

DPCs can interfere with the progression of DNA polymerases and RNA polymerases to disrupt the replication and transcription process. In *E.coli*, unrepaired or undergoing repair DNA lesions can lead to the stalling of DNA replication at replication forks to trigger the SOS response. Error-free DNA replication plays a dominant role in the early phase of the SOS response. If error-free repair pathways cannot efficiently process DNA damage and restart replication, error-prone DNA replication is triggered to bypass DNA lesions via translesion DNA synthesis (TLS) in the later phase of SOS response⁵². Cell survival and gene mutation depend on the activated DNA repair pathways induced by SOS response. Our experiments show that NER and HR can process DPCs cooperatively (Fig. 1, 2 and 3). This synergistic effect is partially due to RecA-mediated excision repair pathway which effectively provides the second opportunity for NER to process DPCs to avoid detrimental effects⁵³. Results indicate that GWR109 cells with a different gene knockout of the NER pathway show a different sensitivity to the genotoxic effects induced by hAGT-DNA crosslinks (Fig. 2.C-D, 3.C-D). This difference may result from the role of NER subpathways and the interaction of NER with HR. UA⁻ cells are more sensitive to genotoxic effects induced by DBE and EBH than UB⁻ and Mfd⁻ cells (Fig. 2.C-D, 3.C-D). The plausible explanation is that the presence of UvrA plays a role to interact with HR to cooperatively mitigate genotoxic effects induced by DPCs. However, it cannot show which strand or which base pair benefits from this interaction.

Studies have shown that a major DBE adduct is formed by conjugation to GSH in cells without AGT expression and this is confirmed by our data (Fig. S2.A-B). This adduct can be miscoded with different polymerases to cause mutations *in vivo* and *in vitro*^{51; 54}. hAGT expression can significantly increase the mutation frequency in cells treated with DBE and EBH (Fig. 2 and 3). One possible explanation is that they are due to the apurinic sites formed through the depurination of hAGT-DNA crosslinks at the N⁷-G sites or similar lesions, which are known to miscode with A to cause GC to AT mutations³⁹. Another explanation is that the extremely bulky nature of formed DPCs blocks the DNA polymerases and requires error-prone TLS to bypass the lesions to restart the replication forks⁵². As NER or HR deficiency results in more than 1000-fold higher mutation frequency (Fig. 2 and 3), it suggests that the TLS triggered by DPCs plays a dominant role to induce mutations. Mutations are the result of DNA repair and error-prone TLS of DPCs. Mutation spectra can show DPCs at which strand or which base pair cannot be efficiently repaired by error-free repair pathways to avoid base mutations. It is known that DNA lesions are preferentially repaired in the transcribed template strand by TC-NER and Mfd plays important roles to remove a stalled RNAP, recruit repair proteins and detect DNA lesions^{26; 55; 56}. The absence of Mfd can significantly increase mutation frequency in cells expressing hAGT treated with DBE or EBH (Fig. 2 and 3) and the relative mutation ratio in the transcribed

strand in cells expressing no AGT treated with DBE (Table 1), showing that TC-NER plays a role in processing DPCs. This finding argues against previous report which evaluated the damage tolerance and repair of DPCs with the insensitive biomarker-cell survival and claimed that TC-NER and TLS are not involved²⁷. The finding that NER could significantly decrease mutations in the transcribed strand has been shown in mammalian cells previously^{39; 46; 57}. Our data provide evidence to support this conclusion in prokaryotic cells attacked by DPCs. However, the mutation frequency in the transcribed strand in GWR109 cells treated with DBE and EBH is 79% (23/29) and 86% (25/29) respectively (Table 1 and Table S3). Similar results have been reported for UV-induced hypoxanthine phosphoribosyltransferase (hprt) mutations in NER deficient Chinese hamster ovary (CHO) cells and DBE-induced hprt mutations in CHO cells^{39; 57}. The mutations lacking strand specificity were attributed to the effect of neighbouring genes, the different polarity and fidelity of DNA replication in the transcribed and non-transcribed strand in those reports^{46; 57}. However, our mutation spectra show that the absence of UvrA but not Mfd and UvrB can significantly decrease mutation frequency in the transcribed strand induced by large size DPCs (Table 1). Since the mutation spectrum were examined in the same gene and same parental cell line with the knockout of different NER gene in this assay, the neighbouring genes, the polarity and fidelity of DNA replication cannot affect the mutation specificity of rpoB gene. The plausible explanation for this phenomenon is due to the protein-protein interaction of different DNA repair pathways to interfere with the repair of DNA lesions in the transcribed strand. UvrA can be efficiently recruited by TC-NER and loaded at damaged sites. As steric hindrance of large size DPCs can efficiently block UvrB loading to UvrA to form preincision complex¹¹, NER cannot be efficiently initiated. The presence of UvrA at damaged sites in the transcribed strand will affect the DNA lesions recognized and processed by HR and cause higher mutation frequency in the transcribed strand (Table 1). So UA- cells can significantly decrease the relative mutation rate in the transcribed strand compared with GWR109 cells for the absence of UvrA. RA- cells show the preferential repair of transcribed strand for the presence of UvrA cannot interfere with HR but NER can preferentially process a small amount of DPCs in the transcribed strand (Fig. 2, 3 and Table 1). The mutation frequency in the transcribed strand in UB- cells is not significantly different from GWR109 cells or Mfd-cells for the presence of UvrA (Table 1). When cells expressing no hAGT treated with DBE, GSH-DNA crosslinks can be processed by NER and HR efficiently (Fig. S2. A-B and Fig. 2.A) and TC-NER will process DNA adducts in the transcribed strand preferentially, so each of Mfd, UvrA and UvrB deficiency will significantly increase the relative mutation ratio in the transcribed strand (Table 1). This result is consistent with the previous report which examined 6/17 and 10/11 of UV-induced hprt mutation in the transcribed strand in NER normal and deficient CHO cells respectively⁵⁷. The NER deficient CHO cells in the report can recognize but cannot efficiently incise UV damages⁵⁸, so the presence of NER protein associated with UV damages will interfere with the lesions processed by other error-free DNA repair pathways and this results in higher mutation frequency in the transcribed strand.

Mutagenic results show that Mfd-dependent TC-NER plays a role to decrease mutation frequency in cells containing hAGT-DNA crosslinks (Fig. 2 and 3). However, mutation spectra show that Mfd is dispensable when facing large size DPCs (Table 1). This result is reasonable, for large size DPCs are processed by HR dominantly. Compared to HR, NER and TC-NER play a very limited role to decrease mutation frequency and the trivial effect cannot be shown in mutation spectrum in Mfd-cells with HR proficiency. However, in HR deficient RA-cells, mutation spectrum show that mutations in the transcribed strand is significantly less than that in the non-transcribed strand (Table 1), confirming TC-NER plays a role in processing hAGT-DNA crosslinks.

The presence of NER can significantly decrease the mutation frequency at GC pairs but increase the mutation frequency at AT pairs, which indicates that NER can efficiently process DPCs at GC sites but not AT sites and the presence of NER can interfere with DPCs at TA sites processed by HR (Fig. 4 and Table 1). The interference results from the presence of UvrA and UvrB but not Mfd. This finding clearly shows the base preference of NER pathway. This finding is very surprising and interesting. Since NER pathway is very similar between bacteria and eukaryotes⁵⁹, this phenomenon very likely exists in eukaryotes too. Common fragile sites and some nonfolate-sensitive rare fragile sites are characterized by highly AT-rich sequences or expanded AT-rich minisatellite repeats⁶⁰. These fragile sites are very sensitive to replication stress and their instability plays important roles in many aspects of cancer^{60; 61}. It is unclear how the stability of these common fragile sites is regulated and how the lesions of these sites are repaired. Our finding may provide some clues to explore the repair mechanism of these highly AT-rich sequences and the effects of replication stress on DNA damage and cancer. Further studies will be performed to confirm this finding in mammalian cells and explore the repair mechanism of base preference and its physiological role and clinical significance.

In conclusion, we describe a new *in vivo* model system to investigate the DNA repair mechanism and genotoxic effects of DPCs precisely. Our results show that NER prefers to process DPCs at G base and the presence of NER can interfere with DPCs in the transcribed strand and TA base pairs of any strand processed by HR. We provide evidence to show the heterogeneity of NER repair at the base level and TC-NER is involved in the repair of DPCs in the prokaryotic cells firstly.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

I thank Dr. Anthony E. Pegg for suggestions to design this project and financial support. I thank Dr. Robert W. Sobol (University of Pittsburgh School of Medicine and University of Pittsburgh Cancer Institute, Hillman Cancer center, Pittsburgh, PA) for his critical comments on the manuscript. I thank Ms. Andrea Braganza for her proofreading. I thank Dr. Barry L. Wanner for providing BW25113 cell line and plasmids pKD46 and pKD3. I thank Dr. Leona D. Samson for providing cell lines FC218, FC 326, CJM1 and CJM2.

Role of the funding source

This project was supported by the National Institutes of Health (R01 CA-097209 and CA-071976 to AEP). The funding sources had no involvement in study design, data collection, data analysis and interpretation, the writing of the article or the decision to submit the article for publication.

Abbreviations

AGT (MGMT)	8-alkylguanine-DNA alkyltransferase
hAGT	human wild type AGT
NER	nucleotide excision repair
TC-NER	transcription-coupled NER
GC-NER	global genome-coupled NER
BER	base excision repair
DSB	double-strand break

RNAP	RNA polymerase
HR	homologous recombination
TLS	translesion synthesis
GSH	glutathione
PMSF	phenylmethylsulfonyl fluoride
PCR	polymerase chain reaction
SDS-PAGE	polyacrylamide gel electrophoresis in the presence of sodium dodecyl sulfate
MW	Molecular Weight
DPC	DNA-protein crosslink
Oligo	oligodeoxyribonucleotide
DTT	dithiothreitol
DBE	1,2-dibromoethane
EBH	epibromohydrin
BDO	1,3-butadiene diepoxide
E. coli	Escherichia coli
Rif^{res}	rifampicin resistance
rpoB	gene encoding β -subunit of RNA polymerase in <i>E. coli</i>
HPRT	hypoxanthine phosphoribosyltransferase
CHO	Chinese hamster ovary

References

1. Barker S, Weinfeld M, Murray D. DNA-protein crosslinks: their induction, repair, and biological consequences. *Mutat Res.* 2005; 589:111–35. [PubMed: 15795165]
2. Reardon JT, Cheng Y, Sancar A. Repair of DNA-protein cross-links in mammalian cells. *Cell Cycle.* 2006; 5:1366–70. [PubMed: 16775425]
3. Connelly JC, Leach DR. Repair of DNA covalently linked to protein. *Mol Cell.* 2004; 13:307–16. [PubMed: 14967139]
4. Loeber R, Michaelson E, Fang Q, Campbell C, Pegg AE, Tretyakova N. Cross-linking of the DNA repair protein Omcron6-alkylguanine DNA alkyltransferase to DNA in the presence of antitumor nitrogen mustards. *Chem Res Toxicol.* 2008; 21:787–95. [PubMed: 18324787]
5. Copeland KD, Lueras AM, Stemp ED, Barton JK. DNA cross-linking with metallointercalator-peptide conjugates. *Biochemistry.* 2002; 41:12785–97. [PubMed: 12379121]
6. Kane SA, Lippard SJ. Photoreactivity of platinum(II) in cisplatin-modified DNA affords specific cross-links to HMG domain proteins. *Biochemistry.* 1996; 35:2180–8. [PubMed: 8652559]
7. Sastry SS, Spielmann HP, Hoang QS, Phillips AM, Sancar A, Hearst JE. Laser-induced protein-DNA cross-links via psoralen furanside monoadducts. *Biochemistry.* 1993; 32:5526–38. [PubMed: 8504073]
8. Johansen ME, Muller JG, Xu X, Burrows CJ. Oxidatively induced DNA-protein cross-linking between single-stranded binding protein and oligodeoxynucleotides containing 8-oxo-7,8-dihydro-2'-deoxyguanosine. *Biochemistry.* 2005; 44:5660–71. [PubMed: 15823024]
9. Reardon JT, Sancar A. Repair of DNA-polypeptide crosslinks by human excision nuclease. *Proc Natl Acad Sci U S A.* 2006; 103:4056–61. [PubMed: 16537484]

10. Minko IG, Zou Y, Lloyd RS. Incision of DNA-protein crosslinks by UvrABC nuclease suggests a potential repair pathway involving nucleotide excision repair. *Proc Natl Acad Sci U S A*. 2002; 99:1905–9. [PubMed: 11842222]
11. Nakano T, Morishita S, Katafuchi A, Matsubara M, Horikawa Y, Terato H, Salem AM, Izumi S, Pack SP, Makino K, Ide H. Nucleotide excision repair and homologous recombination systems commit differentially to the repair of DNA-protein crosslinks. *Mol Cell*. 2007; 28:147–58. [PubMed: 17936711]
12. Nakano T, Salem AM, Terato H, Pack SP, Makino K, Ide H. Comparison of the activities of bacterial and mammalian nucleotide excision repair systems for DNA-protein crosslinks. *Nucleic Acids Symp Ser (Oxf)*. 2009:225–6.
13. Pegg AE, Dolan ME, Moschel RC. Structure, function, and inhibition of O6-alkylguanine-DNA alkyltransferase. *Prog Nucleic Acid Res Mol Biol*. 1995; 51:167–223. [PubMed: 7659775]
14. Gerson SL. MGMT: its role in cancer aetiology and cancer therapeutics. *Nat Rev Cancer*. 2004; 4:296–307. [PubMed: 15057289]
15. Abril N, Luque-Romero FL, Prieto-Alamo MJ, Margison GP, Pueyo C. ogt alkyltransferase enhances dibromoalkane mutagenicity in excision repair-deficient *Escherichia coli* K-12. *Mol Carcinog*. 1995; 12:110–7. [PubMed: 7662116]
16. Abril N, Luque-Romero FL, Christians FC, Encell LP, Loeb LA, Pueyo C. Human O(6)-alkylguanine-DNA alkyltransferase: protection against alkylating agents and sensitization to dibromoalkanes. *Carcinogenesis*. 1999; 20:2089–94. [PubMed: 10545410]
17. Liu H, Xu-Welliver M, Pegg AE. The role of human O(6)-alkylguanine-DNA alkyltransferase in promoting 1,2-dibromoethane-induced genotoxicity in *Escherichia coli*. *Mutat Res*. 2000; 452:1–10. [PubMed: 10894884]
18. Liu L, Hachey DL, Valadez G, Williams KM, Guengerich FP, Loktionova NA, Kanugula S, Pegg AE. Characterization of a mutagenic DNA adduct formed from 1,2-dibromoethane by O6-alkylguanine-DNA alkyltransferase. *J Biol Chem*. 2004; 279:4250–9. [PubMed: 14645247]
19. Kolman A, Chovanec M, Osterman-Golkar S. Genotoxic effects of ethylene oxide, propylene oxide and epichlorohydrin in humans: update review (1990-2001). *Mutat Res*. 2002; 512:173–94. [PubMed: 12464351]
20. Liu L, Pegg AE, Williams KM, Guengerich FP. Paradoxical enhancement of the toxicity of 1,2-dibromoethane by O6-alkylguanine-DNA alkyltransferase. *J Biol Chem*. 2002; 277:37920–8. [PubMed: 12151404]
21. Valadez JG, Liu L, Loktionova NA, Pegg AE, Guengerich FP. Activation of bis-electrophiles to mutagenic conjugates by human O6-alkylguanine-DNA alkyltransferase. *Chem Res Toxicol*. 2004; 17:972–82. [PubMed: 15257623]
22. Kalapila AG, Loktionova NA, Pegg AE. Effect of O6-alkylguanine-DNA alkyltransferase on genotoxicity of epihalohydrins. *Environ Mol Mutagen*. 2009; 50:502–14. [PubMed: 19472322]
23. Kalapila AG, Loktionova NA, Pegg AE. Alkyltransferase-mediated toxicity of 1,3-butadiene diepoxide. *Chem Res Toxicol*. 2008; 21:1851–61. [PubMed: 18712882]
24. Loecken EM, Dasari S, Hill S, Tabb DL, Guengerich FP. The bis-electrophile diepoxybutane cross-links DNA to human histones but does not result in enhanced mutagenesis in recombinant systems. *Chem Res Toxicol*. 2009; 22:1069–76. [PubMed: 19364102]
25. Loecken EM, Guengerich FP. Reactions of glyceraldehyde 3-phosphate dehydrogenase sulfhydryl groups with bis-electrophiles produce DNA-protein cross-links but not mutations. *Chem Res Toxicol*. 2008; 21:453–8. [PubMed: 18163542]
26. Mellon I, Hanawalt PC. Induction of the *Escherichia coli* lactose operon selectively increases repair of its transcribed DNA strand. *Nature*. 1989; 342:95–8. [PubMed: 2554145]
27. Salem AM, Nakano T, Takuwa M, Matoba N, Tsuboi T, Terato H, Yamamoto K, Yamada M, Nohmi T, Ide H. Genetic analysis of repair and damage tolerance mechanisms for DNA-protein cross-links in *Escherichia coli*. *J Bacteriol*. 2009; 191:5657–68. [PubMed: 19617358]
28. Mackay WJ, Han S, Samson LD. DNA alkylation repair limits spontaneous base substitution mutations in *Escherichia coli*. *J Bacteriol*. 1994; 176:3224–30. [PubMed: 8195077]
29. Samson L, Han S, Marquis JC, Rasmussen LJ. Mammalian DNA repair methyltransferases shield O4MeT from nucleotide excision repair. *Carcinogenesis*. 1997; 18:919–24. [PubMed: 9163676]

30. Datsenko KA, Wanner BL. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A*. 2000; 97:6640–5. [PubMed: 10829079]
31. Fang Q, Kanugula S, Tubbs JL, Tainer JA, Pegg AE. Repair of O4-alkylthymine by O6-alkylguanine-DNA alkyltransferases. *J Biol Chem*. 2010; 285:8185–95. [PubMed: 20026607]
32. Thomason LC, Costantino N, Court DL. *E. coli* genome manipulation by P1 transduction. *Curr Protoc Mol Biol*. 2007 **Chapter 1**, Unit 1 17.
33. Liu L, Xu-Welliver M, Kanugula S, Pegg AE. Inactivation and degradation of O(6)-alkylguanine-DNA alkyltransferase after reaction with nitric oxide. *Cancer Res*. 2002; 62:3037–43. [PubMed: 12036910]
34. Fang Q, Noronha AM, Murphy SP, Wilds CJ, Tubbs JL, Tainer JA, Chowdhury G, Guengerich FP, Pegg AE. Repair of O6-G-alkyl-O6-G interstrand cross-links by human O6-alkylguanine-DNA alkyltransferase. *Biochemistry*. 2008; 47:10892–903. [PubMed: 18803403]
35. Fang Q, Kanugula S, Pegg AE. Function of domains of human O6-alkylguanine-DNA alkyltransferase. *Biochemistry*. 2005; 44:15396–405. [PubMed: 16285744]
36. Cosloy SD, Oishi M. Genetic transformation in *Escherichia coli* K12. *Proc Natl Acad Sci U S A*. 1973; 70:84–7. [PubMed: 4630612]
37. Ozawa N, Guengerich FP. Evidence for formation of an S-[2-(N7-guanyl)ethyl]glutathione adduct in glutathione-mediated binding of the carcinogen 1,2-dibromoethane to DNA. *Proc Natl Acad Sci U S A*. 1983; 80:5266–70. [PubMed: 6577422]
38. Thier R, Taylor JB, Pemble SE, Humphreys WG, Persmark M, Ketterer B, Guengerich FP. Expression of mammalian glutathione S-transferase 5-5 in *Salmonella typhimurium* TA1535 leads to base-pair mutations upon exposure to dihalomethanes. *Proc Natl Acad Sci U S A*. 1993; 90:8576–80. [PubMed: 8378332]
39. Kalapila AG, Pegg AE. Alkyltransferase-mediated toxicity of bis-electrophiles in mammalian cells. *Mutat Res*. 2010; 684:35–42. [PubMed: 19941875]
40. van Bladeren PJ, Breimer DD, Rotteveel-Smijds GM, Mohn GR. Mutagenic activation of dibromomethane and diiodomethane by mammalian microsomes and glutathione S-transferases. *Mutat Res*. 1980; 74:341–6. [PubMed: 7010125]
41. Commandeur JN, Stijntjes GJ, Vermeulen NP. Enzymes and transport systems involved in the formation and disposition of glutathione S-conjugates. Role in bioactivation and detoxication mechanisms of xenobiotics. *Pharmacol Rev*. 1995; 47:271–330. [PubMed: 7568330]
42. Ehrenberg L, Hussain S. Genetic toxicity of some important epoxides. *Mutat Res*. 1981; 86:1–113. [PubMed: 7015118]
43. Kolman A, Spivak I, Naslund M, Dusinska M, Cedervall B. Propylene oxide and epichlorohydrin induce DNA strand breaks in human diploid fibroblasts. *Environ Mol Mutagen*. 1997; 30:40–6. [PubMed: 9258328]
44. Koskinen M, Plna K. Specific DNA adducts induced by some mono-substituted epoxides in vitro and in vivo. *Chem Biol Interact*. 2000; 129:209–29. [PubMed: 11137062]
45. Fu D, Calvo JA, Samson LD. Balancing repair and tolerance of DNA damage caused by alkylating agents. *Nat Rev Cancer*. 2012; 12:104–20. [PubMed: 22237395]
46. Vrieling H, Venema J, van Rooyen ML, van Hoffen A, Menichini P, Zdzenicka MZ, Simons JW, Mullenders LH, van Zeeland AA. Strand specificity for UV-induced DNA repair and mutations in the Chinese hamster HPRT gene. *Nucleic Acids Res*. 1991; 19:2411–5. [PubMed: 1674998]
47. Koga N, Inskeep PB, Harris TM, Guengerich FP. S-[2-(N7-guanyl)ethyl]glutathione, the major DNA adduct formed from 1,2-dibromoethane. *Biochemistry*. 1986; 25:2192–8. [PubMed: 3707941]
48. Kim DH, Humphreys WG, Guengerich FP. Characterization of S-[2-(N1-adenyl)ethyl]glutathione as an adduct formed in RNA and DNA from 1,2-dibromoethane. *Chem Res Toxicol*. 1990; 3:587–94. [PubMed: 1715767]
49. Ross MK, Pegram RA. [35S]-labeling of the *Salmonella typhimurium* glutathione pool to assess glutathione-mediated DNA binding by 1,2-dibromoethane. *Chem Biol Interact*. 2003; 146:39–49. [PubMed: 12902151]

50. Raney KD, Meyer DJ, Ketterer B, Harris TM, Guengerich FP. Glutathione conjugation of aflatoxin B1 exo- and endo-epoxides by rat and human glutathione S-transferases. *Chem Res Toxicol.* 1992; 5:470–8. [PubMed: 1391613]
51. Foster PL, Wilkinson WG, Miller JK, Sullivan AD, Barnes WM. An analysis of the mutagenicity of 1,2-dibromoethane to *Escherichia coli*: influence of DNA repair activities and metabolic pathways. *Mutat Res.* 1988; 194:171–81. [PubMed: 3054522]
52. Patel M, Jiang Q, Woodgate R, Cox MM, Goodman MF. A new model for SOS-induced mutagenesis: how RecA protein activates DNA polymerase V. *Crit Rev Biochem Mol Biol.* 2010; 45:171–84. [PubMed: 20441441]
53. Bichara M, Meier M, Wagner J, Cordonnier A, Lambert IB. Postreplication repair mechanisms in the presence of DNA adducts in *Escherichia coli*. *Mutat Res.* 2011; 727:104–22. [PubMed: 21558018]
54. Kim MS, Guengerich FP. Polymerase blockage and misincorporation of dNTPs opposite the ethylene dibromide-derived DNA adducts S-[2-(N7-guanyl)ethyl]glutathione, S-[2-(N2-guanyl)ethyl]glutathione, and S-[2-(O6-guanyl)ethyl]glutathione. *Chem Res Toxicol.* 1998; 11:311–6. [PubMed: 9548801]
55. Manelyte L, Kim YI, Smith AJ, Smith RM, Savery NJ. Regulation and rate enhancement during transcription-coupled DNA repair. *Mol Cell.* 2010; 40:714–24. [PubMed: 21145481]
56. Selby CP, Sancar A. Molecular mechanism of transcription-repair coupling. *Science.* 1993; 260:53–8. [PubMed: 8465200]
57. Vrieling H, Van Rooijen ML, Groen NA, Zdzienicka MZ, Simons JW, Lohman PH, van Zeeland AA. DNA strand specificity for UV-induced mutations in mammalian cells. *Mol Cell Biol.* 1989; 9:1277–83. [PubMed: 2725498]
58. Zdzienicka MZ, van der Schans GP, Westerveld A, van Zeeland AA, Simons JW. Phenotypic heterogeneity within the first complementation group of UV-sensitive mutants of Chinese hamster cell lines. *Mutat Res.* 1988; 193:31–41. [PubMed: 3336370]
59. Hanawalt PC, Spivak G. Transcription-coupled DNA repair: two decades of progress and surprises. *Nat Rev Mol Cell Biol.* 2008; 9:958–70. [PubMed: 19023283]
60. Durkin SG, Glover TW. Chromosome fragile sites. *Annu Rev Genet.* 2007; 41:169–92. [PubMed: 17608616]
61. Glover TW, Arlt MF, Casper AM, Durkin SG. Mechanisms of common fragile site instability. *Hum Mol Genet.* 2005; 14:R197–205. Spec No. 2. [PubMed: 16244318]

Highlights

DPCs are processed by NER and HR with base preference.

TC-NER is involved in processing DPCs in prokaryotic cells.

Inefficient process of DPCs by NER results in mutations lacking strand preference.

NER and HR can process DPCs cooperatively and competitively.

HR plays a dominant role to process large size DPCs *in vivo*.

Figure. 1. hAGT-DNA crosslinks can be processed by NER and HR with HR playing dominant role

2.5 μ g of genomic DNA was isolated with Qiagen blood and cell culture DNA Maxi Kit (A) or with CsCl ultracentrifugation (C) as described in "Materials and Methods" and digested with DNase I at 37°C for 1 h. Crosslinked hAGT was determined by Western blot with antibody-ATO-1 and shown in A and C. Bands of Western blot were quantified with ImageQuant and the relative amount of remaining crosslinked hAGT was shown in B and D respectively. Results indicate the mean $\pm$ S.D. of two independent experiments.

Figure 2. Effects of hAGT expression on survival and mutations in *E. coli* cell strains treated with DBE

GWR109, UA-, UB-, Mfd- and RA-cells transformed empty pQE30 vector (A and B) or pQE30-WT hAGT (C and D) were treated with 0-10mM DBE for 90 min. Cell survival (A and C) and mutation frequency (B and D) were determined as described in "Materials and Methods". The results represent the means of 3 measurements $\pm$ SD. The results for GWR109 (○), UA- (⊙), Mfd- (⊕), UB- (□) and RA- (△) were shown.

Figure 3. Effects of hAGT expression on survival and mutations in *E. coli* cell strains treated with EBH

GWR109, UA-, UB-, Mfd- and RA-cells transformed empty pQE30 vector (A and B) or pQE30-WT hAGT (C and D) were treated with 0-5 mM EBH for 90 min. Cell survival (A and C) and mutation frequency (B and D) were determined as described in "Materials and Methods". The results represent the means of 3 measurements $\pm$ SD. The results for GWR109 (○), UA- (⊙), UB- (□), Mfd- (⊕) and RA- (△) cells were shown.

A.

B.

Figure 4. Effects of NER status on T:A to C:G and G:C to A:T mutations in *E. coli* cell strains treated with DBE

Experiments were performed as described in “Materials and Methods”. T:A to C:G mutations in FC326 cells with pQE30 (○), pQE-hAGT(□) and pQE-Ogt(Δ) and CJM1 cells with pQE30 (●), pQE-hAGT (■) and pQE-Ogt (▲) are shown in A. G:C to A:T mutations in FC218 with pQE30 (○) and pQE-Ogt(Δ) and CJM2 cells with pQE30 (●) and pQE-Ogt (▲) are shown in B. Results are shown with mean $\pm$ SD for at least 3 separate experiments. ** $p < 0.01$ compared to NER proficient cell lines expressing the same protein with t test.

Table 1

Relative mutation ratio in cells treated with DBE and EBH

Cell strains	Relative mutation ratio (%) [#]					
	DBE(10mM)			EBH(1mM)		
	G:C ^d	C:G ^d	T:A+T:T ^e	G:C ^d	C:G ^d	T:A+T:T ^e
-hAGT						
GWR109	11	89	24	35	65	8
Mfd-	48 ^{a**b***}	52 ^{a**b***}	17 ^{b*}			
UA-	71 ^{a**b***}	29 ^{a**b***}	4 ^{a*b***}	29	71	4
UB-	63 ^{a**b***}	27 ^{a**b***}	10 ^{a*b***}			
RA-	7	93	48 ^{a*}	22	78	18
+hAGT						
GWR109	79	21	3	86	14	3
Mfd-	83 ^{b**c***}	17 ^{b**c***}	0 ^{b*}	76 ^{b**c**}	24 ^{b**c*}	3
UA-	52 ^{a**b*}	48 ^{a**b*}	0 ^{b*}	57 ^{a**b***}	43 ^{a**b***}	0 ^{b*}
UB-	71 ^{b**c*}	29 ^{b**c*}	7	78 ^{b**c*}	22 ^{b**c*}	10
RA-	27 ^{a***}	73 ^{a***}	13	26 ^{a***}	74 ^{a***}	13

[#] One tailed chi square test was used to calculate p value.

^{a*} Compared to GWR109 cells, p<0.05

^{b*} Compared to RA -cells, p<0.05

^{c*} Compared to UA-cells, p<0.05

^{a**} Compared to GWR109 cells, p<0.01.

^{b**} Compared to RA-cells, p<0.01.

^{c**} Compared to UA-cells, p<0.01.

^d The ratio was calculated by the number of G:C or C:G mutants/total number of mutants at GC base pairs.

The ratio was calculated by the total number of mutants at AT base pairs/total number of mutants.