3242–3252 *Nucleic Acids Research, 2017, Vol. 45, No. 6* *Published online 23 January 2017*
doi: 10.1093/nar/gkx015

A novel role for transcription-coupled nucleotide excision repair for the *in vivo* repair of 3,N^4-ethenocytosine

Isaac A. Chaim[1,2], Alycia Gardner[2,3], Jie Wu[4], Teruaki Iyama[5], David M. Wilson, III[5] and Leona D. Samson[1,2,3,6,*]

[1]Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA, [2]Center for Environmental Health Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139, USA, [3]Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA, [4]The Barbara K. Ostrom (1978) Bioinformatics and Computing Facility in the Swanson Biotechnology Center, Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA, [5]Laboratory of Molecular Gerontology, National Institute on Aging, Intramural Research Program, National Institutes of Health, 251 Bayview Boulevard, Suite 100, Baltimore, MD 21224, USA and [6]The David H. Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Received November 03, 2016; Revised December 12, 2016; Editorial Decision January 03, 2017; Accepted January 20, 2017

ABSTRACT

Etheno (ε) DNA base adducts are highly mutagenic lesions produced endogenously via reactions with lipid peroxidation (LPO) products. Cancer-promoting conditions, such as inflammation, can induce persistent oxidative stress and increased LPO, resulting in the accumulation of ε-adducts in different tissues. Using a recently described fluorescence multiplexed host cell reactivation assay, we show that a plasmid reporter bearing a site-specific 3,N^4-ethenocytosine (εC) causes transcriptional blockage. Notably, this blockage is exacerbated in Cockayne Syndrome and xeroderma pigmentosum patient-derived lymphoblastoid and fibroblast cells. Parallel RNA-Seq expression analysis of the plasmid reporter identifies novel transcriptional mutagenesis properties of εC. Our studies reveal that beyond the known pathways, such as base excision repair, the process of transcription-coupled nucleotide excision repair plays a role in the removal of εC from the genome, and thus in the protection of cells and tissues from collateral damage induced by inflammatory responses.

INTRODUCTION

Etheno (ε) DNA adducts are highly mutagenic base lesions characterized by an exocyclic (imidazole) ring. They are produced endogenously through reactions with lipid peroxidation (LPO) products generated under oxidative stress conditions (1–4). ε-Adducts were first identified as products resulting from exposure of DNA bases to chloroacetaldehyde (CAA) (5). The role of these lesions in carcinogenesis became evident after occupational exposure to vinyl chloride (VC), a precursor of chloroethylene oxide and CAA (6), was correlated with the development of a rare tumor, angiosarcoma of the liver (7–9). Studies in rodents have shown comparable carcinogenic effects of both VC (10) and urethane (11) exposure. The toxicity and mutagenicity of ε-adducts have also been demonstrated in both bacteria (12–16) and mammalian cells (13,17,18).

With the development of more sensitive techniques to detect ε-adducts, background levels of 1,N^6-ethenoadenine (εA) and 3,N^4-ethenocytosine (εC) were first measured in rodents (19,20). Interestingly, a range of adduct levels was detected for different tissues, suggesting differential exposure to relevant DNA-damaging agents and/or variation in tissue DNA repair capacity (19). It was later shown that LPO products represent a major endogenous source of ε-adducts (1–3). LPO is triggered by the reaction of reactive oxygen and nitrogen species with polyunsaturated fatty acid residues of phospholipids. Conditions such as inflammation, impaired metal transport or dietary imbalance can induce persistent oxidative stress and LPO excess, which in turn can cause an accumulation of ε-adducts in different tissues. Notably, many of these conditions are characteristic of cancer-prone diseases (21).

Two DNA repair pathways have been shown to repair ε-adducts, namely base excision repair (BER) and direct reversal (DR). BER is a finely tuned process that begins with the excision of a damaged base by a DNA glycosylase, generating an abasic (AP) site. For responses involv-

*To whom correspondence should be addressed. Tel: +1 617 258 7813; Fax: +1 617 258 7813; Email: lsamson@mit.edu

ing a monofunctional DNA glycosylase, the AP site is processed by apurinic/apyriminidic endonuclease 1 (APE1), which produces a single strand break with a 3′-OH and 5′-deoxyribosephosphate (5′-dRP). DNA polymerase β removes the 5′-dRP moiety and fills in the gap by inserting a nucleotide with the complementary base to the undamaged strand. Repair is completed by the action of a DNA ligase to seal the nick (22). DR by the AlkB homolog (ALKBH) enzymes does not require any other proteins or DNA processing per se, but instead involves direct reversal of the lesion via removal of the etheno ring. Specifically, the repair reaction begins with epoxidation of the etheno ring after which the addition of water opens the ring to form a glycol intermediate that is spontaneously released as glyoxal, leaving behind a normal base (23).

In mammalian cells, εA is excised by the alkyladenine DNA glycosylase (AAG, a.k.a. MPG) (24,25) or reversed by ALKBH2 (26). There are conflicting reports in the literature on whether ALKBH3 participates in the repair of εA in single-stranded DNA (26–28). Interestingly, εC can be bound by AAG, yet is not excised by the enzyme (25). This protein–DNA complex has been shown to block the repair of εC by ALKBH2 *in vitro* (29). Counterintuitively, $Aag^{-/-}$ mice show greater accumulation of εC than wild-type (WT) mice. This finding might implicate the recruitment of additional repair proteins by the εC–Aag complex, possibly through a transcription-associated process (30). εC has also been shown to be excised *in vitro* by other DNA glycosylases, such as SMUG1 (31) and TDG (32), and weakly by MBD4 (33). 1,N^2-εG is bound and excised by AAG, but in a less efficient manner than εA (25); repair of 1,N^2-εG by ALKBH3 was also recently reported (34). No repair mechanisms for N^2,3-εG have been described in mammals so far.

Nucleotide excision repair (NER) is another DNA repair pathway that has been proposed to play a role in the removal of some DNA lesions formed by LPO products, such as bulky *trans*-4-hydroxy-2-nonenal-DNA (HNE-DNA) and malondialdehyde-DNA adducts (4,35,36). NER has two distinct sub-pathways: transcription-coupled NER (TC-NER), which recognizes lesions that stall RNA Polymerase II (RNAPII) during active transcription, and global genome NER (GG-NER), which acts on bulky DNA lesions in both active and silent regions of the genome (37). These sub-pathways differ only in their method of lesion recognition, after which they converge into the same downstream enzymatic steps. Deficiencies in TC-NER primarily result in the premature aging disorder, Cockayne Syndrome (CS) (38), while deficiencies in GG-NER and any of the common downstream steps of NER result in xeroderma pigmentosum (XP), a disease characterized by up to a 2000-fold increased risk of developing skin cancer as well as neurological degeneration (39,40). Interestingly, cells deficient in TC-NER (specifically, deficient in the protein CSB) are hypersensitive to physiological concentrations of HNE and develop a higher level of sister chromatid exchanges than wild-type cells. Furthermore, HNE–DNA adducts formed endogenously in mammalian cells can block transcription and can be processed by TC-NER (4,35).

Here, we show that εC, a DNA lesion produced endogenously by LPO products, does indeed cause *in vivo* transcriptional blockage and, when bypassed, can promote the incorporation of an inappropriate ribonucleotide into the transcript opposite the lesion site, in a process known as transcriptional mutagenesis (41). By taking advantage of our recently described fluorescence multiplexed host cell reactivation (FM-HCR) assays for measuring DNA repair capacity, we further demonstrate that εC is repaired by TC-NER.

MATERIALS AND METHODS

Plasmids

As described previously (42), EGFP, mOrange and tagBFP reporter genes subcloned into the pmaxCloning Vector (Lonza) between the *Kpn*I and *Sac*I restriction sites in the multiple cloning site were employed. Plasmids were amplified in *Escherichia coli* DH5α (Invitrogen) and purified using Qiagen endotoxin-free Maxi and Giga kits.

UV treatment. Plasmids were irradiated in TE buffer (10 mM Tris–HCl, 1 mM EDTA, pH 7.0) at a DNA concentration of 50 ng/μl in a volume of 1.5 ml in 10-cm polystyrene Petri dishes (without lids) with 800 J/m^2 UV-C light generated by a Stratalinker 2000 box.

Site-specific thymine dimer reporter. pMax:GFP 615AA plasmid (described previously (42)) was prepared using a Giga Prep kit according to manufacturer instructions (Qiagen). The plasmid was nicked with Nb.*Bts*I (New England Biolabs) to generate a single-strand break in the transcribed strand. The nicked strand was then digested with ExoIII (New England Biolabs), and the remaining single-stranded circular DNA (ssDNA) was purified using a 1% agarose gel. The ssDNA was extracted from the agarose gel using a Gel Extraction Kit (Qiagen). Twenty micrograms of the ssDNA were combined with 9 μl of a 10 mM oligonucleotide sequence containing a thymine dimer localized at bases 614 and 615 of the plasmid [5′ TCAGGGCGGAXGGTTGC 3′, where X denotes a thymine dimer] (Trilink Biotech) and 1× *Pfu* Polymerase Buffer (Thermo Scientific). The mixture was heated to 85°C in a thermocycler for 6 min, and then allowed to anneal by cooling to 37°C at 1°C/min. Following annealing, 12 μl of 10 mM dNTPs and 30 U *Pfu* Polymerase AD were added to the mixture. The oligo was extended at 61°C for 1.5 h. Following extension, the plasmid was purified using a PCR Purification kit following manufacturer instructions (Qiagen). After elution in 105.4 μl EB Buffer (10 mM Tris–Cl, pH 8.5), 5 μl T4 DNA Ligase buffer (New England Biolabs), 2 μl 10 mM dNTPs, 2 μl 25 mM ATP, 0.5 μl 10 mg/ml BSA, 1.5 U T4 DNA Polymerase (New England Biolabs) and 8 U T4 DNA Ligase (New England Biolabs) were added, and the solution was incubated at 16°C for 1 h. The solution was subjected to electrophoresis in a 1% agarose gel to purify the ligated band. Ligated plasmid was extracted from the gel using a Gel Extraction kit (Qiagen). Gel extraction was performed according to manufacturer instructions except that plasmids were eluted in TE buffer (10 mM Tris–Cl, pH 8.0, 1 mM EDTA) (Invitrogen)

Site-specific εA (GFP-616-εA) and εC (GFP-615-εC) reporter. pMax:GFP WT plasmid was prepared using

a Giga Prep kit according to manufacturer instructions. Single stranded plasmid DNA was prepared as described above. Annealing of a site-specific oligo was performed as described above, using an oligonucleotide containing εA [5′GCTCAGGGCGGXCTGGGTGCTCAG GTAGTG, where X denotes εA] or εC [GCTCAGGG CGGAXTGGGTGCTCAGGTAGTG, where X denotes εC] (Life Technologies). Following annealing, 12 μl of 10 mM dNTPs and 30 U of *Pfu* Polymerase AD were added to the mixture, and the oligo was extended at 68°C for 1.5 h. After extension, the plasmid was purified using phenol/chloroform/isoamyl alcohol extraction and ethanol precipitation. The plasmid DNA was resuspended in a mixture containing 105.4 μl EB Buffer (10 mM Tris–Cl, pH 8.5), 5 μl T4 DNA Ligase buffer (New England Biolabs), 2 μl 10 mM dNTPs, 2 μl 25 mM ATP and 0.5 μl 100X BSA. Once the DNA was resuspended, 1.5 U T4 DNA Polymerase and 8 U T4 DNA Ligase (New England Biolabs) were added, and the solution was incubated at 16°C for 1 h. In order to remove excess salts from the solution, the buffer was changed following ligation using 30,000 NMWL Filtration columns (Millipore). The plasmid-containing solution was concentrated three times and finally resuspended in 150 μl TE for purification by gel electrophoresis. In order to purify the plasmid, a 1.5–2% NuSieve GTG low melting point agarose gel (Lonza) was used. The plasmid DNA was extracted from the gel according to the manufacturer's instructions. Briefly, the gel was melted and a series of phenol/chloroform extractions were performed, followed by ethanol precipitation and resuspension of the plasmid in TE buffer.

Primary cells and cell lines

Animals. *Aag*$^{-/-}$ and Triple Knockout (TKO, *Aag*$^{-/-}$,*Alkbh2*$^{-/-}$*Alkbh3*$^{-/-}$) mice were described previously (43). All mice used for experiments were on a mixed B6/129s background. Age- and gender-matched mice were 6–10 weeks old at the time of euthanasia. Mice were euthanized by CO_2 asphyxiation and their spleen was subsequently removed. Single-cell suspensions of splenocytes were prepared by pushing tissue pieces through a 70 μm pore size nylon mesh screen. Erythrocytes were lysed in ACK lysis buffer.

Mice were housed in an Association of Assessment and Accreditation of Laboratory Animal Care–accredited facility. All procedures were approved by the Massachusetts Institute of Technology Committee on Animal Care.

Human cells. B-lymphoblastoid cells (GM01712, CSB; GM01857, CSA; GM01953, healthy; GM02246, XPC; GM02345, XPA; GM02253, XPD; GM07752, healthy and GM21148, XPB) were maintained in RPMI 1640 medium (Gibco) supplemented with 15% FBS, 2 mM L-glutamine and 1% penicillin and streptomycin. SV40-transformed CSA-deficient CS3BE and CSB-deficient CS1AN cell lines (Coriell Institute) transfected with pcDNA-CSA or pcDNA-CSB, respectively (previously described (44)) were maintained in DMEM medium (Gibco) supplemented with 10% HI-FBS, 2 mM L-glutamine, and 1% penicillin and streptomycin. Adherent cell lines were sub-cultured by trypsinization. Cells were maintained at 37°C and 5% CO_2.

Rodent cells. Mouse embryonic fibroblasts (MEFs) were prepared from wild-type C57BL/6J and previously described *Aag*$^{-/-}$ mouse models (45). Primary MEFs were previously immortalized by transfecting MEFs with a pSV3-neo plasmid expressing large T-antigen. Immortalized MEFs were cultured in DMEM medium supplemented with 10% HI-FBS, 2 mM L-glutamine and 1% penicillin and streptomycin, and were subcultured by trypsinization. Cells were maintained at 37°C and 5% CO_2.

Splenic lymphocytes were cultured in supplemented RPMI Glutamax: 20% heat-inactivated FBS, 1% penicillin/streptomycin. B-lymphocytes were stimulated with 10 μg/ml Kdo2-Lipid A (Avanti Polar Lipids) for 48 h before transfection.

DNA repair assays transfections

Electroporation. Lymphoblastoid cell lines were transfected by electroporation. 2×10^6 lymphoblastoid cells in 100 μl of complete medium were combined with a reporter plasmid mixture (Table 1). Cells were electroporated using a 96-well electroporation plate and gene pulser (Bio-Rad) with an exponential waveform at 260 V and 950 μF. Following electroporation, 100 μl of complete medium were added to each well in the plate, and the 200 μl volume was transferred to a 12-well plate containing 800 μl of equilibrated medium.

Lipofection. MEFs and human fibroblasts were transfected by lipofection using Lipofectamine LTX (LifeTechnologies), following manufacturer's instructions. Briefly, plasmid reporter DNA mixture (Table 1) combined with Opti-MEM, Lipofectamine and Plus reagents, and 200 μl of the total mixture were added to each well of the 6-well plate.

Flow cytometry analysis of lesions

Eighteen hours after transfection, cells suspended in culture media were analyzed for fluorescence on a BD LSR II cytometer running FACSDiva software. Cell debris, doublets and aggregates were excluded based on their side- and forward-scatter properties. TO-PRO-3 was added to cells 5–10 min before analysis and was used to exclude dead cells from the analysis. The following fluorophores and their corresponding detectors (in parentheses) were used: tagBFP (Pacific Blue), EGFP (FITC), mOrange (phycoerythrin; PE) and TO-PRO-3 (allophycocyanin; APC). Compensation was set by using single-color controls. Regions corresponding to cells positive for each of the five fluorescent proteins were established by using single-color dropout controls. For reporters that required compensation in more than one detector channel, fluorescence in the reporter channel was plotted separately against each of the channels requiring compensation. Using these plots, both single-color controls and the dropout control (in which the reporter of interest was excluded from the transfection) were used to establish regions corresponding to positive cells.

Table 1. Combinations of reporter plasmids and types of DNA damage used in each experiment

Comb.	DNA plasmid reporters			
	BFP	*GFP*	*Carrier*	*mOrange*
1	***WT*** 90 ng	***WT*** 30 ng		
2	***WT*** 90 ng	***616-εA*** 30 ng		
3	***WT*** 90 ng	***615-εC*** 30 ng		
4	***WT*** 400 ng	***WT*** 100 ng		
5	***WT*** 400 ng	***616-εA*** 100 ng		
6	***WT*** 400 ng	***615-εC*** 100 ng		
7	***WT*** 100 ng	***WT*** 100 ng		
8	***WT*** 100 ng	***UV*** 100 ng		
9	***WT*** 100 ng	***614-T^T*** 100 ng		
10	***WT*** 100 ng	***616-εA*** 100 ng		
11	***WT*** 100 ng	***615-εC*** 100 ng		
12	***WT*** 25 ng	***WT*** 25 ng	450 ng	
13	***WT*** 25 ng	***UV*** 25 ng	450 ng	
14	***WT*** 25 ng	***614-T^T*** 25 ng	450 ng	
15	***WT*** 25 ng	***616-εA*** 25 ng	450 ng	
16	***WT*** 25 ng	***615-εC*** 25 ng	450 ng	
17	***WT*** 25 ng	***WT*** 25 ng	450 ng	***WT*** 25 ng
18	***WT*** 25 ng	***615-εC*** 25 ng	375 ng	***UV*** 100 ng

Each column corresponds to a different plasmid reporter expressing a fluorescent protein, with the exception of 'carrier' which corresponds to an empty vector. WT = wild-type; εA = 1,N^6-ethenoadenine; εC = 3,N^4-ethenocytosine; T^T = thymine dimer. Numbers before εA, εC and T^T correspond to the site of the lesion in the plasmid. UV corresponds to plasmids treated with ultraviolet light at 800 J/m^2.

Calculation of percent fluorescent reporter expression. Every experimental setup consisted of two sets of transfections: a control transfection (CT) and a sample transfection (ST) containing one or more reporters with DNA lesions. Both transfections included the same color combination with the same undamaged reporter to normalize each set for transfection efficiency.

Fluorescence index (FI) for a given reporter within one transfection was calculated as follows:

$$FI = \frac{C_F \times MFI}{C_L}$$

where C_F is the number of positive fluorescent cells for that given fluorophore, *MFI* is the mean fluorescence intensity of the C_F, and C_L is the total number of live cells.

The normalized fluorescence index for a given reporter FI^O was calculated as follows:

$$FI^O = \frac{FI^n}{FI^E}$$

where FI^n corresponds to the *FI* of a reporter normalized to the *FI* of the transfection efficiency normalization plasmid, FI^E.

Normalized reporter expression from a sample transfection, $FI^O{}_{ST}$, and that from the same reporter plasmid in control transfection, $FI^\circ{}_{CT}$, were used to compute the percent reporter expression (%*R.E.*) as follows:

$$\%R.E. = \frac{FI^O{}_{ST}}{FI^O{}_{CT}} \times 100$$

RNA isolation for RNA-Seq

At 18 h, a fraction of transfected cells were harvested by centrifugation, washed three times with PBS, and resuspended in 1 ml TRIzol reagent. The suspension was extracted with 200 μl chloroform. The aqueous phase was removed, combined with one volume of absolute ethanol, and applied to a Qiagen RNeasy miniprep spin column. The column was washed two times with 500 μl Buffer RPE (Qiagen) and finally eluted in 40 μl diethylpyrocarbonate (DEPC)-treated water. The quality of the RNA preparation was determined by using a bioanalyzer to confirm an RNA integrity number (RIN) of at least 9.0. At least 190 ng of total RNA were stored in TE buffer at −80°C until submission for mRNA sequencing (RNA-Seq).

Total RNA samples were submitted to the Massachusetts Institute of Technology BioMicroCenter for preparation and sequencing. Briefly, total RNA was poly-A purified, fragmented, and converted to cDNA by using the Clontech SMARTseq protocol. Library construction from cDNA was performed using the Beckman Coulter SPRIworks system. During library amplification, a unique bar code was introduced for each of the 12 samples corresponding to the four transfections performed in triplicate (Supplementary Table S1) and from which total RNA was generated. All samples were combined and ran on a NextSeq instrument.

Next generation sequencing data analysis

RSEM (46) (version 1.2.15) was used to estimate gene expression based on hg19 UCSC known gene annotations. The count table was imported into DESeq2 (47) (version 1.10.1) for differential gene expression analysis. BWA (48) (version 0.7.10) was used to map the raw reads to the plasmids. Then all properly paired, uniquely and perfectly mapped reads were counted to estimate expression of the reporters. Relative expression levels of the reporters were computed by normalizing their counts to BFP expression in the same sample. SAMtools (49) (version 1.3) mpileup (options: -d1000000 -B) was used to process the bam alignments, followed by VarScan (50) (version 2.3.6) to call mutations on the reporters.

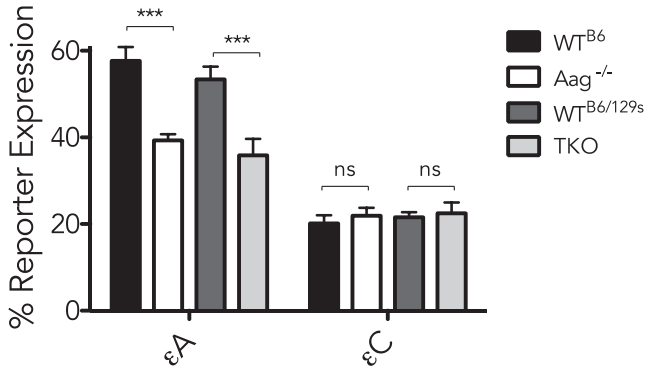

Figure 1. Repair of εA and εC in WT, $Aag^{-/-}$ and TKO MEFs. Cell lines were transfected with site-specific εA or εC GFP repair reporters. WT^{B6} and $Aag^{-/-}$ cell lines are derived from C57BL/6 background mice whereas $WT^{B6/129s}$ and TKO (triple knockout, $Aag^{-/-}/Alkbh2^{-/-}/Alkbh3^{-/-}$) cell lines are derived from mixed background mice (C57BL/6 and 129s). Error bars show standard deviation of at least three biological replicates. Two-tailed unpaired *t* test $^{***}P \leq 0.001$; ns, not significant.

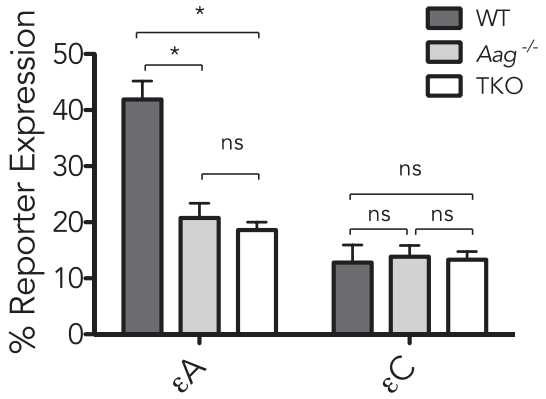

Figure 2. Repair of εA and εC in B-splenic-lymphocytes from WT, $Aag^{-/-}$ and TKO MEFs. Following spleen removal, lymphocytes were stimulated with 10 μM Kdo2-Lipid-A for 48 h before transfection. Cells were transfected with site-specific εA or εC GFP repair reporters. TKO = triple knockout, $Aag^{-/-}/Alkbh2^{-/-}/Alkbh3^{-/-}$). All mice were in a mixed background (C57BL/6 and 129s). Error bars show standard deviation of two technical replicates. Two-tailed unpaired *t* test $^{*}P \leq 0.05$; ns, not significant.

RESULTS

εA is repaired in vivo by AAG, but not ALKBH2 or ALKBH3; εC is not repaired by any of them

In order to test the *in vivo* role of known DNA repair proteins in the repair of εA and εC adducts, we performed an HCR assay by transfecting reporter constructs harboring site-specific lesions into repair proficient (WT^{B6} and $WT^{B6/129s}$) or repair deficient MEFs ($Aag^{-/-}$ or $Aag^{-/-}/Alkbh2^{-/-}/Alkbh3^{-/-}$, hereafter referred to as Triple KnockOut, TKO). Two different WT MEFs were used as controls, since the repair deficient cells were derived from mice with different genetic backgrounds ($Aag^{-/-}$ from C57BL/6 and the TKO from the mixed background B6/129s). The presence of either εA or εC resulted in a decreased % reporter expression (%R.E.) compared to undamaged controls. The decrease in %R.E. was greater for εC than for εA (Figure 1, plasmid combinations #1, 2 and 3 in Table 1). For the εA reporter, $Aag^{-/-}$ MEFs exhibited significantly lower %R.E. compared to WT, agreeing with previously published work demonstrating repair of εA by AAG-initiated BER (45). Interestingly, there appeared to be no difference in the repair of εC in WT versus $Aag^{-/-}$ cells. Moreover, TKO MEFs, while deficient in the repair of εA, did not appear more deficient than $Aag^{-/-}$ cells, although a direct comparison should be interpreted with caution given the varying genetic backgrounds. TKO MEFs also showed no difference in εC repair compared to WT cells (Figure 1, plasmid combinations #1, 2 and 3 in Table 1).

Importantly, the same repair relationships were observed following transfection of site-specific εA and εC plasmids into Kdo2-Lipid-A-stimulated mouse primary splenic-B-lymphocytes isolated from each mouse strain (Figure 2, plasmid combinations #4, 5 and 6 in Table 1). Specifically, in the absence of Aag, repair of εA was reduced, while additional deficiencies in Alkbh2 and Alkbh3 did not further reduce this repair. εC repair remained constant in all three genotypes.

TC-NER and downstream NER components are involved in the repair of εC

Given the strong decrease in %R.E. observed for site-specific εC lesions, we sought to determine if this decrease was a result of a transcription block that could be repaired by the NER pathway. To this end, we used a panel of B-lymphoblastoid cell lines derived from individuals with CS or XP, diseases characterized by deficiencies in the different sub-pathways of NER and sensitivity to UV light. Previously, we established a reporter assay system to evaluate DNA repair capacity for UV-induced DNA damage using plasmids treated with 800 J/m^2 UV that were subsequently transfected into cell lines deficient in NER (XPA, XPB, XPC or XPD) (42). In this study, deficiencies in XPA, XPB, XPC and XPD resulted in decreased reactivation of UV-treated plasmids compared to reactivation in healthy controls (Figure 3A, plasmid combinations #7 and 8 in Table 1). We also determined DNA repair capacity of UV damage in CS cell lines defective in TC-NER. Cells lacking CSA displayed significantly decreased reactivation; while plasmid reactivation appeared lower in CSB-deficient cells in comparison to cells from the healthy control, the difference did not quite reach significance ($P = 0.056$).

We had previously shown that XPA-deficient lymphoblastoid cells could not efficiently repair plasmids containing a site-specific thymine dimer (T^T), the most common lesion caused by UV light (42). Here, we tested whether the NER deficiencies of the other XP and CS cell lines could be phenotyped using a site-specific thymine dimer, a lesion known to block transcription. Following transfection of a GFP reporter bearing a site-specific T^T, we reproduced our previous XPA mutant results. Moreover, we observed similar repair deficits in cells defective for TC-NER (CSA and CSB) as well as the other downstream NER components that operate after the damage recognition step (XPB and XPD). Interestingly, deficiencies in the GG-NER subpathway (cells deficient in XPC) did not appear to affect the repair of this site-specific lesion (Figure 3B, plasmid com-

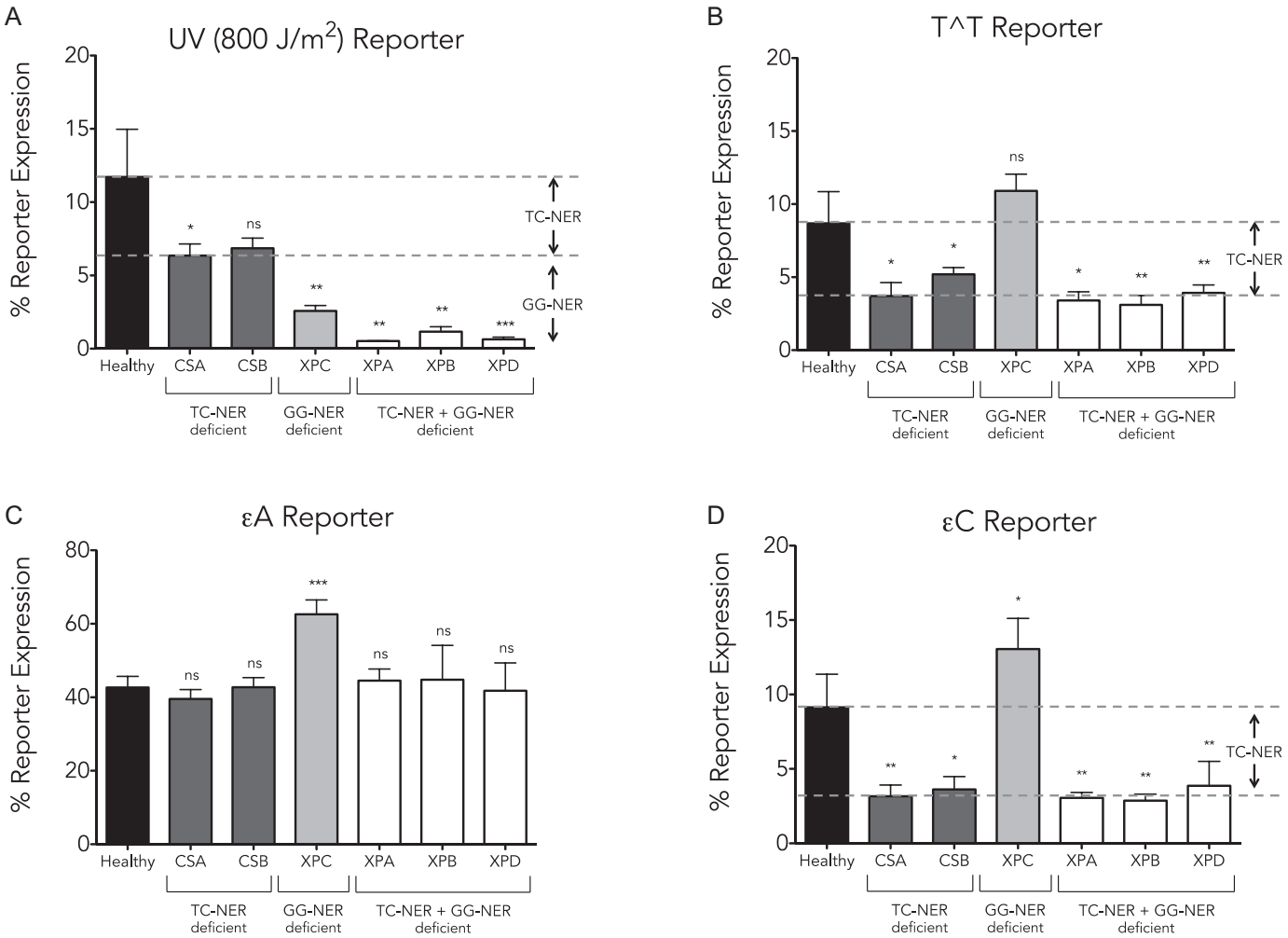

Figure 3. Repair of UV-irradiated and site-specific thymine dimer, εA and εC plasmids by a panel of B-lymphoblastoid cell lines with varying NER deficiencies. Cell lines were transfected with GFP repair reporters irradiated with 800 J/m^2 UV-C (**A**) or with GFP reporters containing a site-specific thymine dimer (T^T) (**B**), εA (**C**) or εC (**D**). Black boxes correspond to NER proficient cells; dark gray, TC-NER deficient; light grey, GG-NER deficient; white, downstream-NER (TC-NER and GG-NER) deficient. Dashed lines represent a threshold between repair assumed to occur through TC-NER (above line) and GG-NER (below line) based on the lowest repairing TC-NER deficient cell line tested. Error bars show standard deviation of at least three biological replicates. Two-tailed unpaired *t* test between healthy control and each deficient cell line: *$P \leq 0.05$; **$P \leq 0.01$; ***$P \leq 0.001$; ns, not significant.

binations #7 and 9 in Table 1), even though the repair of randomly induced UV damage was clearly deficient (Figure 3A), pointing out the power of our approach in identifying effective TC-NER substrates.

After validating that NER-deficient cell lines were deficient in the repair of a site-specific transcription blocking base lesion, we transfected the site-specific εA or εC plasmid into CS or XP cells. None of the cells displayed a deficiency in the repair of εA; curiously, the XPC-deficient cell line actually displayed higher εA repair capacity than the healthy control (Figure 3C, plasmid combinations #7 and 10 in Table 1). In contrast, repair of εC was deficient in five out of six NER-deficient cell lines, closely resembling the repair phenotypes observed for the site-specific T^T lesion (Figure 3D, plasmid combinations #7 and 11 in Table 1). Deficiencies in TC-NER, as well as the downstream NER factors, resulted in significantly diminished εC repair capacity. Again, the GG-NER-deficient XPC line appeared more repair-proficient than the healthy control. It should be noted that the 'healthy' cell line GM07752 happens to have a particularly low NER capacity compared with other 'healthy' lymphoblastoid cell lines (42).

To further investigate the role of TC-NER in the repair of εC, we utilized human SV40-transformed fibroblasts derived from CS patients deficient in CSA (CS3BE) or CSB (CS1AN) plus their respective isogenic repair-proficient cDNA-complemented counterparts. We determined the repair capacity of these four cell lines for UV treated plasmids and for plasmids with a site-specific T^T, εA or εC lesion. Consistent with our results in the non-isogenic lymphoblastoid cells, both CSA- and CSB-deficient cells showed decreased repair of UV treated plasmids as well as site-specific T^T and εC plasmids, in comparison to their complemented equivalents. Furthermore, no significant difference in the repair of εA was detected in either CSA- or CSB-deficient cells relative to the complemented controls (Figure 4A and

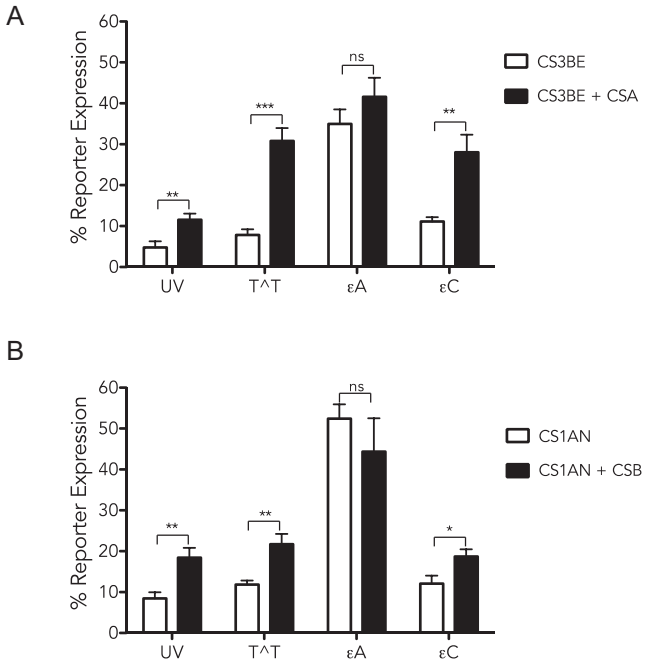

Figure 4. DNA repair capacity measurements of isogenic patient derived fibroblasts. DNA repair capacity of two pairs of isogenic patient derived fibroblasts (**A,** CSA, **B,** CSB) deficient (white boxes) or complemented with the respective deficient protein (dark boxes) was measured. They were transfected with GFP repair reporters that were (i) irradiated with 800 J/m^2 UV-C or that contained a site-specific, (ii) T^T, (iii) εA or (iv) εC. Error bars show standard deviation of at least three biological replicates. Two-tailed unpaired *t* test *$P \leq 0.05$; **$P \leq 0.01$; ns, not significant.

B, plasmid combinations #12–16 in Table 1). Taken together, these results indicate that TC-NER is involved in the repair of εC DNA lesions.

εC can be bypassed by RNAPII and is a source of transcriptional mutagenesis

Given this newly discovered *in vivo* blockage of transcription by εC and its repair through TC-NER, we sought to determine whether this lesion could sometimes be bypassed *in vivo* by RNAPII resulting in transcriptional mutagenesis. εC was positioned opposite G at GFP base-pair 615 of the coding strand (Figure 5A), at the wobble position of the codon for glutamine. Using conventional site-directed mutations, we established that a T at this wobble position also results in glutamine and fluorescence, whereas either G or A results in histidine and a non-fluorescent variant (data not shown). As such, given our fluorescent experimental setup, we cannot discern if the production of fluorescent GFP is due to removal of the transcription block by εC repair, or due to the misincorporation of A opposite εC during transcription. Similarly, reduced fluorescence could be the result of either transcriptional blockage or transcriptional mutagenesis to C or U. To address all of these possibilities, we turned to high-throughput sequencing of the reporter transcripts.

We performed a multiplexed HCR assay that included the GFP site-specific εC plasmid and a UV treated (800 J/m^2) mOrange vector (Figure 5A, plasmid combinations #17 and 18 in Table 1), which were transfected into CSA-deficient and CSA-complemented fibroblasts. Eighteen hours after transfection, the samples were split in two; one half was assayed through flow cytometry, while the other half was used for RNA isolation and subjected to RNA-Seq (Supplementary Tables S1 and S2). Both analyses yielded remarkably similar results (Figure 5B), showing that the differences between proficient and deficient cells in the repair of both UV-irradiated and εC constructs were equivalent using flow cytometry and RNA-Seq methods. Relative transcript levels for both endogenous and plasmid reporter genes were determined in CSA-deficient and -proficient cell lines (Figure 5C and Supplementary Table S3). Genes expressed at the same level under both conditions appear on the diagonal, as is the case for the majority of the endogenous genes, revealing no major transcriptional changes in the presence of the damaged plasmids.

Transcriptional blockage was evident for the damaged plasmids due to the presence of either UV lesions (mOrange) or εC (GFP), since reporter gene expression falls under the diagonal. Nevertheless, DESeq2 expression analysis of the εC reporter did not reach statistical significance for the CSA-complemented samples. Furthermore, sequence analysis of the RNA-Seq data at and around where εC was positioned (site 615 of GFP) revealed that the lesion can be bypassed by RNAPII (Figure 5D). As expected, transcripts from control plasmids containing an undamaged C at site 615 had a very low transcriptional mutagenesis frequency (0.086% on average). In contrast, εC bypass in CSA-deficient cells resulted in the incorporation of an erroneous base across from the damage in 9.5% of transcripts (compared to 2.8% in CSA-complemented cells). Specifically A, T(U) and C were present in 5.6%, 3.3% and 0.6% of the transcripts, respectively.

As expected, the repair-proficient CSA-complemented cells showed a lower percentage of transcriptional mutagenesis events given that the lesion is being actively repaired more efficiently by TC-NER. Finally, sequence level analysis of the UV irradiated mOrange reporter revealed transcriptional mutagenesis of thymine dimers caused by UV irradiation of AA to AG in CSA-deficient cells (Supplementary Figure S1).

DISCUSSION

Using a multiplexed HCR assay, we were able to reveal a novel role for the NER pathway, specifically TC-NER, in the repair of the highly mutagenic and cytotoxic DNA lesion, εC. Given that this lesion is a product of endogenous LPO products, characteristic of cancer-prone inflammatory diseases, we propose that TC-NER may play an important role in protection against inflammation-associated disease. That both humans and mice deficient in TC-NER undergo premature aging could in part be due to a deficiency in the repair of inflammation associated DNA damage (51,52).

It is noteworthy that we did not observe a major contribution by Alkbh2 and Alkbh3 in the repair of εA, in either MEFs or mouse splenic-B-lymphocytes (Figures 1 and 2). Both enzymes have been shown *in vitro* to repair εA DNA lesions (26–28). Deletion of both enzymes together with Aag, in chronic-inflammation mouse models, has been

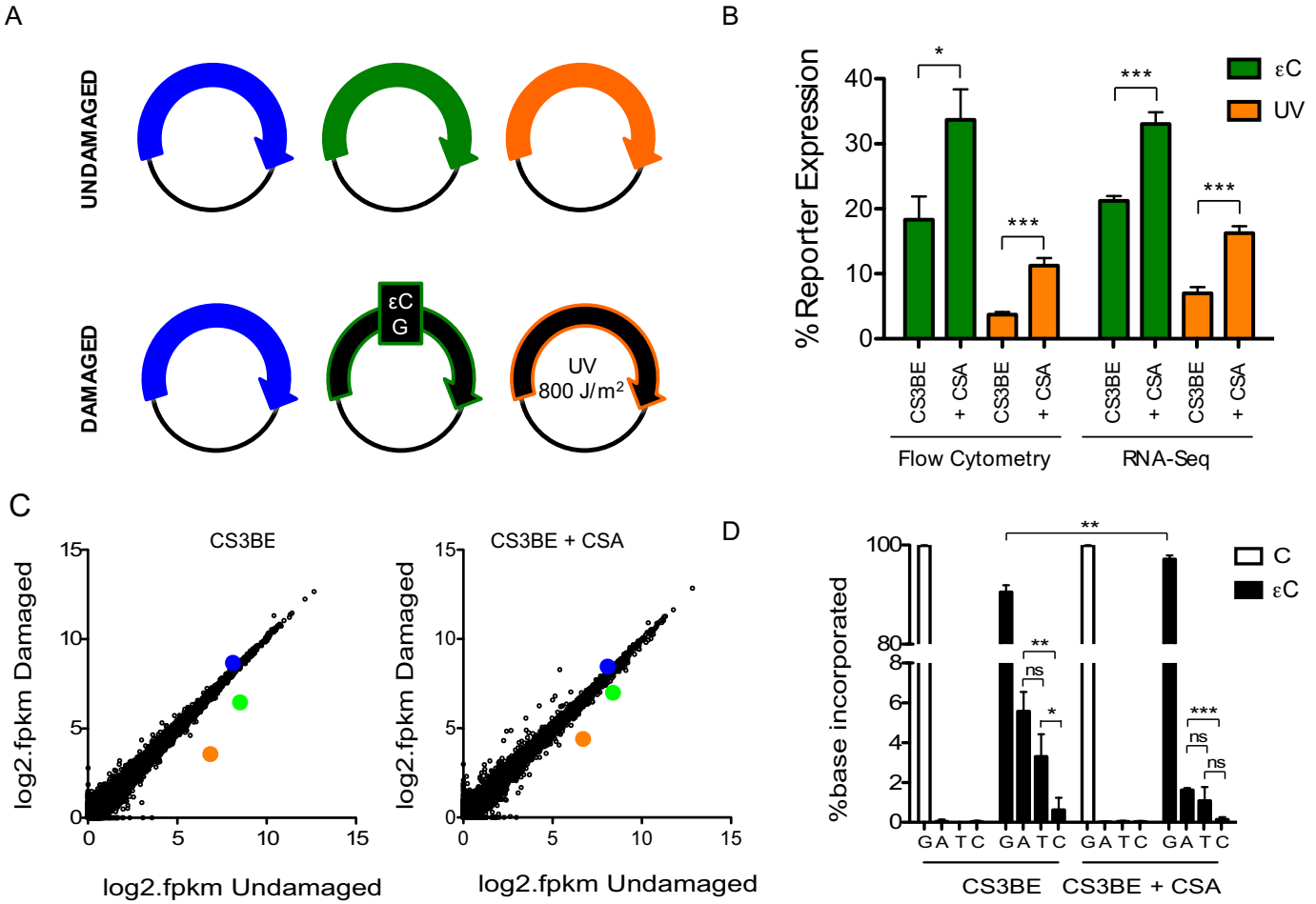

Figure 5. HCR readout by flow cytometry and RNA-Seq. DNA repair capacity of patient derived fibroblasts deficient in CSA (CS3BE) and their CSA complemented counterpart (CS3BE + CSA). (**A**) Cells were transfected with BFP as a transfection control reporter, GFP bearing a site-specific εC and mOrange irradiated with 800 J/m^2 UV-C. (**B**) %R.E. for flow cytometry data (filled bars) was calculated as explained in Materials and Methods. %R.E. RNA-Seq data (empty bars) was calculated directly from transcript counts and normalized to efficiency normalization plasmid counts and the respective control transfection. (**C**) Representative gene expression profile for undamaged versus damaged transfections of both cell lines. Endogenous transcripts are shown in black and reporter plasmids are colored in blue, green and orange. (**D**) RNA-Seq sequence variants detected in transcripts at position 615 in a GFP reporter containing either a C or an εC on the transcribed strand. Frequencies are reported for the expected variant (G) as well as for the reminding 3 possible variants (C, G, T/U). Error bars show standard deviation for three biological replicates. Two-tailed unpaired *t* test: $^*P \leq 0.05$; $^{**}P \leq 0.01$; $^{***}P \leq 0.001$; ns, not significant.

shown to lead to increased accumulation of ε-adducts in colon epithelium (specifically 1,N^2-εG, while εA did not reach significance) compared to Aag deletion alone (43). It is possible that the contribution of these enzymes in εA repair is cell-type specific and not expressed or active in MEFs or splenic-B-lymphocytes; although they have been shown to be expressed in the spleen of C57BL/6 mice (53). Alternatively, *in vivo*, they may simply not play a prominent role in the repair of εA.

Interestingly, εC can be bound by AAG (25), yet it is not excised by this DNA glycosylase (Figures 1 and 2). *In vitro*, the interaction between AAG and εC blocks the repair of this lesion by ALKBH2 (29). Surprisingly, *Aag*$^{-/-}$ mice show increased accumulation of εC (in the absence of Aag) in the colonic mucosa of inflammatory bowel disease mouse models (54). We hypothesized that if Aag binds to εC and recruits downstream repair enzymes, we would observe decreased εC plasmid reactivation in Aag$^{-/-}$ cells. However, one could also argue that there would be increased εC repair in *Aag*$^{-/-}$ cells, if the absence of Aag allowed Alkbh2 to access the lesion. In fact, neither turned out to be the case in MEFs or primary splenic-B-lymphocytes as no difference in the repair of εC was observed in the absence or presence of Aag. Given (i) the lack of a clear role for Aag and the Alkbh enzymes in εC repair and (ii) the strong decrease in reporter expression observed in the presence of this lesion, we propose that εC causes a transcription blockage similar to that seen for the site-specific TˆT reporter (42). We further propose that εC is repaired by the TC-NER machinery. Nevertheless, we cannot rule out the possibility that 'naked' εC and the Aag-εC complex have the same effect on transcription, namely transcriptional blockage. Finally, it is striking that the closely related lesions εA and εC can have such diverse effects on transcription. Through crystallographic analysis of the Aag–εC complex we have previously shown that the inability to protonate εC at C5

in comparison to the N7 position of εA renders the lesion incapable of excision by Aag (55). While this does not directly explain why εC is repaired by TC-NER and εA is not, it does provide a paradigm for why seemingly closely related DNA base lesions may be perceived very differently by DNA polymerases, RNA polymerases and DNA repair proteins.

It is notable that XPC-deficient cells exhibited reduced repair capacity for plasmids containing high levels of UV-damage located randomly in the reporter plasmid, indicating that this HCR-based assay is capable of monitoring GG-NER (Figure 3). However, it is also notable that GG-NER plays no significant role in repairing a single transcription-blocking lesion located in an actively transcribed region, probably because RNAPII lesion detection outcompetes XPC-mediated GG-NER lesion detection. As such, our method cannot differentiate between cells being deficient in GG-NER for the repair of site-specific transcription-blocking lesions or cells simply not utilizing GG-NER for the repair of a single DNA lesion located in a highly transcribed gene. Given that the plasmid used is non-replicative and that transcription from the CMV promoter is vigorous, it is possible that the observed roles for TC- and GG-NER would change if the plasmid were either replicative or transcribed less efficiently. Another indication that repair of site-specific transcription-blocking lesions is occurring through TC-NER and not GG-NER is that TC-NER deficient cells showed similar repair phenotypes for site-specific TˆT repair as cells deficient in downstream-NER steps, even though this was not the case with UV-irradiated plasmids.

The fact that the repair phenotype of a site-specific εC by human lymphoblastoid cell lines deficient in NER so closely recapitulated that of a site-specific TˆT strongly supports our conclusion that the repair of εC occurs via the TC-NER pathway (Figure 3). This observation was further validated in a different cell type, namely human fibroblasts, deficient or complemented for either CSA or CSB (Figure 4). Importantly, each pair of cell lines was isogenic, minimizing the contribution of any inter-individual variations in NER capacity.

TC-NER proteins have been postulated to play a role in the repair of BER substrates. In particular, the contribution of CSB to the repair of 8oxoG and uracil has been shown to be DNA glycosylase dependent (OGG1 and UNG, respectively) in yeast (56), mouse cells (57,58) and human fibroblasts (59). Similarly, CSB has been implicated in the repair of single-strand breaks (SSBs) in actively transcribed genes (60). These SSBs could also be produced through the BER pathway. Nevertheless, we did not observe at any point the participation of Aag in TC-NER of εA adducts. We cannot rule out the possibility that εC repair by TC-NER involves, to some degree, formation of BER intermediates created by the action of other DNA glycosylases (such as TDG, MBD4 or SMUG1) on εC. Nevertheless, Aag does not seem to play a role in modulating TC-NER for the repair of εC under our experimental conditions.

RNA-Seq analysis of our reporter transcripts allowed us to corroborate that, in fact, fewer transcripts are produced in the presence of εC, validating the interpretation of our flow cytometry results as indicating transcription blockage by εC (Figure 5). Moreover, we were able to describe the transcriptional mutagenesis properties of εC *in vivo*. The only studies on the effect of εC on transcription were done *in vitro* and date back to the 1980s; CAA treated poly(dC) oligos were shown to block transcription and induce ribonucleotide misincorporation by calf thymus RNAPII (61). To our knowledge, the transcription blockage and mutagenic properties of εC described here represent novel *in vivo* features for this lesion. RNAPII transcriptional mutagenesis tends to resemble the misincorporation profile exhibited by mammalian replicative DNA polymerases during synthesis opposite DNA lesions (41). In the case of εC, misincorporation during DNA replication has the preference T>A>>C (17). We observed a similar trend with A and T(U) misincorporation being favored and incorporation of C being the least likely event.

In the broader context of LPO-induced DNA damage in mammals, while bulky HNE–DNA and malondialdehyde-derived adducts are widely known to block transcription and known to be repaired by TC-NER (4,35,36), for the non-bulky ε-adducts, only 1,N^2-εG had been reported to block transcription *in vitro* (62). Altogether, our *in vivo* results showing transcriptional blockage and transcriptional mutagenesis by εC complement the notion that DNA damage and repair play a pivotal role in the etiology of disease; in particular, cancer-prone inflammatory diseases, degenerative diseases and other diseases of old age, which are characterized by increased oxidative stress that can lead to the formation of LPO products. In these situations, repair of εC becomes highly relevant in both a replicative and non-replicative context. On the one hand, it can contribute to the accumulation of potentially deleterious mutations during replication; and on the other hand, it can perturb cellular homeostasis via RNAPII blockage or production of mutated transcripts via transcriptional mutagenesis.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

FUNDING

National Institutes of Health (NIH) Directors Pioneer Award [DPI-ES022576]; NIH [R01-CA075576, R01-CA55042, R01-CA149261, P30-ES02109]; Intramural Research Program at the NIH, National Institute on Aging. L.D.S. is an American Cancer Society Research Professor. Funding for open access charge: LDS discretionary funds.
Conflict of interest statement. None declared.

REFERENCES

1. El Ghissassi,F., Barbin,A., Nair,J. and Bartsch,H. (1995) Formation of 1,N^6-ethenoadenine and 3,N^4-ethenocytosine by lipid peroxidation products and nucleic acid bases. *Chem. Res. Toxicol.*, **8**, 278–283.
2. Sodum,R.S. and Chung,F.L. (1988) 1,N^2-ethenodeoxyguanosine as a potential marker for DNA adduct formation by trans-4-hydroxy-2-nonenal. *Cancer Res.*, **48**, 320–323.
3. Sodum,R.S. and Chung,F.L. (1991) Stereoselective formation of in vitro nucleic acid adducts by 2,3-epoxy-4-hydroxynonanal. *Cancer Res.*, **51**, 137–143.
4. Winczura,A., Zdzalik,D. and Tudek,B. (2012) Damage of DNA and proteins by major lipid peroxidation products in genome stability. *Free Radic. Res.*, **46**, 442–459.

5. Kochetkov,N., Shibaev,V. and Kost,A. (1971) New reaction of adenine and cytosine derivatives, potentially useful for nucleic acids modification. *Tetrahedron Lett.*, **12**, 1993–1996.
6. El Ghissassi,F., Barbin,A. and Bartsch,H. (1998) Metabolic activation of vinyl chloride by rat liver microsomes: low-dose kinetics and involvement of cytochrome P450 2E1. *Biochem. Pharmacol.*, **55**, 1445–1452.
7. Block,J.B. (1974) Angiosarcoma of the liver following vinyl chloride exposure. *J. Am. Med. Assoc.*, **229**, 53–54.
8. Creech,J. Jr and Johnson,M. (1974) Angiosarcoma of liver in the manufacture of polyvinyl chloride. *J. Occup. Med.*, **16**, 150–152.
9. Lee,F. and Harry,D. (1974) Angiosarcoma of the liver in a vinyl-chloride worker. *Lancet*, **303**, 1316–1318.
10. Maltoni,C. and Cotti,G. (1988) Carcinogenicity of vinyl chloride in Sprague-Dawley rats after prenatal and postnatal exposure. *Ann. N. Y. Acad. Sci.*, **534**, 145–159.
11. Zimmerli,B. and Schlatter,J. (1991) Ethyl carbamate: analytical methodology, occurrence, formation, biological activity and risk assessment. *Mut. Res./Genet. Toxicol.*, **259**, 325–350.
12. Basu,A.K., Wood,M.L., Niedernhofer,L.J., Ramos,L.A. and Essigmann,J.M. (1993) Mutagenic and genotoxic effects of three vinyl chloride-induced DNA lesions: 1,N^6-ethenoadenine, 3,N^4-ethenocytosine, and 4-amino-5-(imidazol-2-yl) imidazole. *Biochemistry*, **32**, 12793–12801.
13. Pandya,G.A. and Moriya,M. (1996) 1,N^6-Ethenodeoxyadenosine, a DNA adduct highly mutagenic in mammalian cells. *Biochemistry*, **35**, 11487–11492.
14. Palejwala,V.A., Rzepka,R.W., Simha,D. and Humayun,M.Z. (1993) Quantitative multiplex sequence analysis of mutational hot spots. Frequency and specificity of mutations induced by a site-specific ethenocytosine in M13 viral DNA. *Biochemistry*, **32**, 4105–4111.
15. Cheng,K., Preston,B., Cahill,D., Dosanjh,M., Singer,B. and Loeb,L. (1991) The vinyl chloride DNA derivative N^2,3-ethenoguanine produces G-A transitions in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.*, **88**, 9974–9978.
16. Langouët,S., Mican,A.N., Müller,M., Fink,S.P., Marnett,L.J., Muhle,S.A. and Guengerich,F.P. (1998) Misincorporation of nucleotides opposite five-membered exocyclic ring guanine derivatives by *Escherichia coli* polymerases in vitro and in vivo: 1,N^2-ethenoguanine, 5,6,7,9-tetrahydro-9-oxoimidazo [1,2-a] purine, and 5,6,7,9-tetrahydro-7-hydroxy-9-oxoimidazo [1,2-a] purine. *Biochemistry*, **37**, 5184–5193.
17. Moriya,M., Zhang,W., Johnson,F. and Grollman,A. (1994) Mutagenic potency of exocyclic DNA adducts: marked differences between *Escherichia coli* and simian kidney cells. *Proc. Natl. Acad. Sci. U.S.A.*, **91**, 11899–11903.
18. Akasaka,S. and Guengerich,F.P. (1999) Mutagenicity of site-specifically located 1,N^2-ethenoguanine in Chinese hamster ovary cell chromosomal DNA. *Chem. Res. Toxicol.*, **12**, 501–507.
19. Barbin,A. (1999) Role of etheno DNA adducts in carcinogenesis induced by vinyl chloride in rats. *IARC Scientific Publ.*, 303–313.
20. Bartsch,H., Barbin,A., Marion,M., Nair,J. and Guichard,Y. (1994) Formation, detection, and role in carcinogenesis of ethenobases in DNA. *Drug Metab. Rev.*, **26**, 349–371.
21. Nair,U., Bartsch,H. and Nair,J. (2007) Lipid peroxidation-induced DNA damage in cancer-prone inflammatory diseases: a review of published adduct types and levels in humans. *Free Radic. Biol. Med.*, **43**, 1109–1120.
22. Fu,D., Calvo,J.A. and Samson,L.D. (2012) Balancing repair and tolerance of DNA damage caused by alkylating agents. *Nat. Rev. Cancer*, **12**, 104–120.
23. Yi,C., Jia,G., Hou,G., Dai,Q., Zhang,W., Zheng,G., Jian,X., Yang,C.G., Cui,Q. and He,C. (2010) Iron-catalysed oxidation intermediates captured in a DNA repair dioxygenase. *Nature*, **468**, 330–333.
24. Saparbaev,M., Kleibl,K. and Laval,J. (1995) *Escherichia coli*, Saccharomyces cerevisiae, rat and human 3-methyladenine DNA glycosylases repair 1,N^6-ethenoadenine when present in DNA. *Nucleic Acids Res.*, **23**, 3750–3755.
25. Lee,C.Y.I., Delaney,J.C., Kartalou,M., Lingaraju,G.M., Maor-Shoshani,A., Essigmann,J.M. and Samson,L.D. (2009) Recognition and processing of a new repertoire of DNA substrates by human 3-methyladenine DNA glycosylase (AAG). *Biochemistry*, **48**, 1850–1861.
26. Ringvoll,J., Moen,M.N., Nordstrand,L.M., Meira,L.B., Pang,B., Bekkelund,A., Dedon,P.C., Bjelland,S., Samson,L.D., Falnes,P.Ø. and Klungland,A. (2008) AlkB homologue 2-mediated repair of ethenoadenine lesions in mammalian DNA. *Cancer Res.*, **68**, 4142–4149.
27. Aas,P.A., Otterlei,M., Falnes,P., Vagbo,C.B., Skorpen,F., Akbari,M., Sundheim,O., Bjoras,M., Slupphaug,G., Seeberg,E. and Krokan,H.E. (2003) Human and bacterial oxidative demethylases repair alkylation damage in both RNA and DNA. *Nature*, **421**, 859–863.
28. Mishina,Y., Yang,C.G. and He,C. (2005) Direct repair of the exocyclic DNA adduct 1,N^6-ethenoadenine by the DNA repair AlkB proteins. *J. Am. Chem. Soc.*, **127**, 14594–14595.
29. Fu,D. and Samson,L.D. (2012) Direct repair of 3,N^4-ethenocytosine by the human ALKBH2 dioxygenase is blocked by the AAG/MPG glycosylase. *DNA Repair*, **11**, 46–52.
30. Meira,L.B., Moroski-Erkul,C.A., Green,S.L., Calvo,J.A., Bronson,R.T., Shah,D. and Samson,L.D. (2009) Aag-initiated base excision repair drives alkylation-induced retinal degeneration in mice. *Proc. Natl. Acad. Sci. U.S.A.*, **106**, 888–893.
31. Kavli,B., Sundheim,O., Akbari,M., Otterlei,M., Nilsen,H., Skorpen,F., Aas,P.A., Hagen,L., Krokan,H.E. and Slupphaug,G. (2002) hUNG2 is the major repair enzyme for removal of uracil from U: A matches, U: G mismatches, and U in single-stranded DNA, with hSMUG1 as a broad specificity backup. *J. Biol. Chem.*, **277**, 39926–39936.
32. Saparbaev,M. and Laval,J. (1998) 3,N^4-ethenocytosine, a highly mutagenic adduct, is a primary substrate for *Escherichia coli* double-stranded uracil-DNA glycosylase and human mismatch-specific thymine-DNA glycosylase. *Proc. Natl. Acad. Sci. U.S.A.*, **95**, 8508–8513.
33. Petronzelli,F., Riccio,A., Markham,G.D., Seeholzer,S.H., Genuardi,M., Karbowski,M., Yeung,A.T., Matsumoto,Y. and Bellacosa,A. (2000) Investigation of the substrate spectrum of the human mismatch specific DNA glycosylase MED1 (MBD4): Fundamental role of the catalytic domain. *J. Cell. Physiol.*, **185**, 473–480.
34. Zdżalik,D., Domańska,A., Prorok,P., Kosicki,K., van den Born,E., Falnes,P.Ø., Rizzo,C.J., Guengerich,F.P. and Tudek,B. (2015) Differential repair of etheno-DNA adducts by bacterial and human AlkB proteins. *DNA Repair*, **30**, 1–10.
35. Maddukuri,L., Speina,E., Christiansen,M., Dudziń ska,D., Zaim,J., Obtuł owicz,T., Kabaczyk,S., Komisarski,M., Bukowy,Z., Szczegielniak,J. *et al.* (2009) Cockayne syndrome group B protein is engaged in processing of DNA adducts of lipid peroxidation product trans-4-hydroxy-2-nonenal. *Mut. Res./Fundam. Mol. Mech. Mutagen.*, **666**, 23–31.
36. Cline,S.D., Riggins,J.N., Tornaletti,S., Marnett,L.J. and Hanawalt,P.C. (2004) Malondialdehyde adducts in DNA arrest transcription by T7 RNA polymerase and mammalian RNA polymerase II. *Proc. Natl. Acad. Sci. U.S.A.*, **101**, 7275–7280.
37. Hanawalt,P.C. and Spivak,G. (2008) Transcription-coupled DNA repair: two decades of progress and surprises. *Nat. Rev. Mol. Cell Biol.*, **9**, 958–970.
38. Marteijn,J.A., Lans,H., Vermeulen,W. and Hoeijmakers,J.H. (2014) Understanding nucleotide excision repair and its roles in cancer and ageing. *Nat. Rev. Mol. Cell Biol.*, **15**, 465–481.
39. Li,C., Wang,L.E. and Wei,Q. (2009) DNA repair phenotype and cancer susceptibility—a mini review. *Int. J. Cancer*, **124**, 999–1007.
40. Bradford,P.T., Goldstein,A.M., Tamura,D., Khan,S.G., Ueda,T., Boyle,J., Oh,K.-S., Imoto,K., Inui,H., Moriwaki,S.-I. *et al.* (2011) Cancer and neurologic degeneration in xeroderma pigmentosum: long term follow-up characterises the role of DNA repair. *J. Med. Genet.*, **48**, 168–176.
41. Brégeon,D. and Doetsch,P.W. (2011) Transcriptional mutagenesis: causes and involvement in tumour development. *Nat. Rev. Cancer*, **11**, 218–227.
42. Nagel,Z.D., Margulies,C.M., Chaim,I.A., McRee,S.K., Mazzucato,P., Ahmad,A., Abo,R.P., Butty,V.L., Forget,A.L. and Samson,L.D. (2014) Multiplexed DNA repair assays for multiple lesions and multiple doses via transcription inhibition and transcriptional mutagenesis. *Proc. Natl. Acad. Sci. U.S.A.*, **111**, E1823–E1832.
43. Calvo,J.A., Meira,L.B., Lee,C.-Y.I., Moroski-Erkul,C.A., Abolhassani,N., Taghizadeh,K., Eichinger,L.W., Muthupalani,S.,

Nordstrand,L.M., Klungland,A. and Samson,L.D. (2012) DNA repair is indispensable for survival after acute inflammation. *J. Clin. Invest.*, **122**, 2680–2689.
44. Iyama,T. and Wilson,D.M. III (2016) Elements that regulate the DNA damage response of proteins defective in Cockayne Syndrome. *J. Mol. Biol.*, **428**, 62–78.
45. Engelward,B.P., Weeda,G., Wyatt,M.D., Broekhof,J.L.M., De Wit,J., Donker,I., Allan,J.M., Gold,B., Hoeijmakers,J.H.J. and Samson,L.D. (1997) Base excision repair deficient mice lacking the Aag alkyladenine DNA glycosylase. *Proc. Natl. Acad. Sci. U.S.A.*, **94**, 13087–13092.
46. Li,B. and Dewey,C.N. (2011) RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics*, **12**, 323–339.
47. Anders,S. and Huber,W. (2010) Differential expression analysis for sequence count data. *Genome Biol.*, **11**, R106–R118.
48. Li,H. and Durbin,R. (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, **25**, 1754–1760.
49. Li,H., Handsaker,B., Wysoker,A., Fennell,T., Ruan,J., Homer,N., Marth,G., Abecasis,G. and Durbin,R. (2009) The sequence alignment/map format and SAMtools. *Bioinformatics*, **25**, 2078–2079.
50. Koboldt,D.C., Zhang,Q., Larson,D.E., Shen,D., McLellan,M.D., Lin,L., Miller,C.A., Mardis,E.R., Ding,L. and Wilson,R.K. (2012) VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res.*, **22**, 568–576.
51. Edifizi,D. and Schumacher,B. (2015) Genome instability in development and aging: insights from nucleotide excision repair in humans, mice, and worms. *Biomolecules*, **5**, 1855–1869.
52. Karikkineth,A.C., Scheibye-Knudsen,M., Fivenson,E., Croteau,D.L. and Bohr,V.A. (2017) Cockayne syndrome: clinical features, model systems and pathways. *Ageing Res. Rev.*, **33**, 3–17.
53. Petryszak,R., Keays,M., Tang,Y.A., Fonseca,N.A., Barrera,E., Burdett,T., Füllgrabe,A., Fuentes,A.M.-P., Jupp,S., Koskinen,S. *et al.* (2016) Expression Atlas update—an integrated database of gene and protein expression in humans, animals and plants. *Nucleic Acids Res.*, **44.1**, D746–D752.
54. Meira,L.B., Bugni,J.M., Green,S.L., Lee,C.W., Pang,B., Borenshtein,D., Rickman,B.H., Rogers,A.B., Moroski-Erkul,C.A., McFaline,J.L., Schauer,D.B., Dedon,P.C., Fox,J.G. and Samson,L.D. (2008) DNA damage induced by chronic inflammation contributes to colon carcinogenesis in mice. *J. Clin. Invest.*, **118**, 2516–2525.
55. Lingaraju,G.M., Davis,C.A., Setser,J.W., Samson,L.D. and Drennan,C.L. (2011) Structural basis for the inhibition of human alkyladenine DNA glycosylase (AAG) by 3,N^4-ethenocytosine-containing DNA. *J. Biol. Chem.*, **286**, 13205–13213.
56. Kim,N. and Jinks-Robertson,S. (2010) Abasic sites in the transcribed strand of yeast DNA are removed by transcription-coupled nucleotide excision repair. *Mol. Cell. Biol.*, **30**, 3206–3215.
57. Khobta,A., Kitsera,N., Speckmann,B. and Epe,B. (2009) 8-Oxoguanine DNA glycosylase (Ogg1) causes a transcriptional inactivation of damaged DNA in the absence of functional Cockayne syndrome B (Csb) protein. *DNA Repair*, **8**, 309–317.
58. Kitsera,N., Stathis,D., Lühnsdorf,B., Müller,H., Carell,T., Epe,B. and Khobta,A. (2011) 8-Oxo-7,8-dihydroguanine in DNA does not constitute a barrier to transcription, but is converted into transcription-blocking damage by OGG1. *Nucleic Acids Res.*, **39**, 5926–5934.
59. Guo,J., Hanawalt,P.C. and Spivak,G. (2013) Comet-FISH with strand-specific probes reveals transcription-coupled repair of 8-oxoGuanine in human cells. *Nucleic Acids Res.*, **41**, 7700–7712.
60. Khobta,A., Lingg,T., Schulz,I., Warken,D., Kitsera,N. and Epe,B. (2010) Mouse CSB protein is important for gene expression in the presence of a single-strand break in the non-transcribed DNA strand. *DNA Repair*, **9**, 985–993.
61. Kusmierek,J. and Singer,B. (1982) Chloroacetaldehyde-treated ribo-and deoxyribopolynucleotides. 2. Errors in transcription by different polymerases resulting from ethenocytosine and its hydrated intermediate. *Biochemistry*, **21**, 5723–5728.
62. Dimitri,A., Goodenough,A.K., Guengerich,F.P., Broyde,S. and Scicchitano,D.A. (2008) Transcription processing at 1,N^2-ethenoguanine by human RNA polymerase II and bacteriophage T7 RNA polymerase. *J. Mol. Biol.*, **375**, 353–366.