Requirement of transcription-coupled nucleotide excision repair for the removal of a specific type of oxidatively induced DNA damage

Leen Sarmini¹, Mohammed Meabed¹, Eirini Emmanouil², George Atsaves², Elena Robeska^{3,4}, Bolesław T. Karwowski⁵, Anna Campalans ^{10,4}, Thanasis Gimisis² and Andriy Khobta ^{1,*}

¹Institute of Nutritional Sciences, Friedrich Schiller University Jena, Jena 07743, Germany, ²Department of Chemistry, National and Kapodistrian University of Athens, Athens 15771, Greece, ³Université Paris-Saclay, CEA/IBFJ/IRCM. UMR Stabilité Génétique Cellules Souches et Radiations, Fontenay-aux-Roses, F-92265, France, ⁴Université de Paris Cité, CEA/IBFJ/IRCM. UMR Stabilité Génétique Cellules Souches et Radiations, Fontenay-aux-Roses, F-92265, France and ⁵DNA Damage Laboratory of Food Science Department, Faculty of Pharmacy, Medical University of Lodz, Lodz 90-151, Poland

Received July 16, 2022; Revised March 06, 2023; Editorial Decision March 27, 2023; Accepted March 29, 2023

ABSTRACT

Accumulation of DNA damage resulting from reactive oxygen species was proposed to cause neurological and degenerative disease in patients, deficient in nucleotide excision repair (NER) or its transcriptioncoupled subpathway (TC-NER). Here, we assessed the requirement of TC-NER for the repair of specific types of oxidatively generated DNA modifications. We incorporated synthetic 5',8-cyclo-2'-deoxypurine nucleotides (cyclo-dA, cyclo-dG) and thymine glycol (Tg) into an EGFP reporter gene to measure transcription-blocking potentials of these modifications in human cells. Using null mutants, we further identified the relevant DNA repair components by a host cell reactivation approach. The results indicated that NTHL1-initiated base excision repair is by far the most efficient pathway for Tq. Moreover, Tq was efficiently bypassed during transcription, which effectively rules out TC-NER as an alternative repair mechanism. In a sharp contrast, both cyclopurine lesions robustly blocked transcription and were repaired by NER, wherein the specific TC-NER components CSB/ERCC6 and CSA/ERCC8 were as essential as XPA. Instead, repair of classical NER substrates, cyclobutane pyrimidine dimer and N-(deoxyguanosin-8-yl)-2-acetylaminofluorene, occurred even when TC-NER was disrupted. The strict requirement of TC-NER highlights cyclo-dA and cyclo-dG as candidate damage types, accountable for cytotoxic and degenerative responses in individuals affected by genetic defects in this pathway.

INTRODUCTION

The nucleotide excision repair (NER) pathway is of a paramount importance for protection from potentially mutagenic DNA damage inflicted by environmental carcinogens. Hereditary defects in NER genes typically translate to sensitivity to UV irradiation and increased tumour incidence already at young age, which has established their causal role in the cancer predisposition syndrome xeroderma pigmentosum (XP) (1). Two subpathways-the global genome (GG-NER) and transcription-coupled (TC-NER)-have been recognised within the NER system. The discovery of their molecular details led to understanding that genetic defects in the components of GG-NER and TC-NER cause different disease phenotypes, as extensively reviewed elsewhere (2-4). Thus, defects in GG-NER genes (such as XPC and XPE/DDB2) confer cancer predisposition, since accumulating DNA damage causes increased mutation rates in replicating cells. In contrast, loss-of-function mutations in TC-NER components (most prominently CSA/ERCC8 and CSB/ERCC6) do not lead to cancer but rather to diseases characterised by pronounced neurological features and segmental degeneration, such as Cockayne syndrome (CS) and an even more severe cerebro-oculo-facio-skeletal syndrome (COFS). The molecular defect beneath the pathogenesis of CS is commonly defined as incapacity of cells to recover transcription after exposure to genotoxic agents, how-

 $\ensuremath{\mathbb{C}}$ The Author(s) 2023. Published by Oxford University Press on behalf of Nucleic Acids Research.

^{*}To whom correspondence should be addressed. Tel: +49 3641 9 49771; Email: a.khobta@uni-jena.de

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

ever the nature of the responsible DNA damage remains unknown.

Blockage of transcribing RNA polymerase complexes at unrepaired damage sites is highly cytotoxic (5,6). In healthy cells, the capacity of DNA modifications to block elongating RNA polymerase II provides a versatile substrate recognition mechanism for TC-NER, however this very mechanism is impaired in CS (7–13). In contrast, GG-NER components primarily recognise structural properties of damaged DNA duplex itself (14-21). Since GG-NER and TC-NER use fundamentally different damage recognition principles, their substrate spectra may not fully overlap. Accordingly, some types of DNA modifications strictly require TC-NER for their removal (22). Others are efficiently, or even preferentially, recognised by GG-NER (23). Based on these considerations, the range of candidate DNA lesions causing neuropathology in CS can be narrowed down to transcription blocking NER substrates, which cannot undergo repair by other mechanisms, including GG-NER. Moreover, as central nervous system is protected from exogenous genotoxic substances by the blood brain barrier, the damage in question should likely arise from an endogenous source.

Since neuronal function heavily relies on aerobic metabolism, it has been proposed that reactive oxygen species arising as by-products during respiration could be responsible for the critical DNA damage (24). Several types of oxidatively generated DNA lesions were suggested as candidates. In the first line, these include intrastrand crosslinks, such as 5',8-cyclopurine-2'-deoxynucleosides (cyclo-dA, cyclo-dG), whose second covalent linkage to deoxyribose makes them inherently resistant to base excision repair (BER), unlike other common DNA oxidation products (24,25). Biochemical evidence has characterised both cyclo-dA cyclo-dG as fairly good NER substrates (26–28), and their transcription blocking potentials towards human RNA polymerase II have been reported (29). Another candidate is thymine glycol (Tg), which is a reported NER substrate as well (30). Both types of cyclopurine lesions and thymine glycol can be generated in DNA by reactions with hydroxyl radicals, which arise through radiolytic decay of water or by Fenton-like reactions of hydrogen peroxide (31-34). The latter is ubiquitously generated in cells as an intermediate detoxification product of superoxide radical anions released during various redox reactions, including the mitochondrial electron transport (35).

We previously developed procedures for efficient incorporation of various types of synthetic DNA modifications into plasmid-borne reporter genes (36,37) and proposed a reliable methodology for quantification of the reporter expression in transfected human cells (38). The combination of these techniques provided a means to both determine the impairment of transcription by structurally defined types of DNA damage and, based on cells' capacities to recuperate the expression, measure the specific repair efficiencies (22,23). In the present study we used patientderived cell lines, carrying genetic defects in the specific NER genes, as well as newly generated knockout cell models, to define specific contributions of GG-NER and TC-NER to the repair of synthetic cyclo-dA, cyclo-dG and Tg lesions.

MATERIALS AND METHODS

Synthesis of the protected phosphoramidite of $2'\alpha$ -fluorothymine glycol

The synthesis of the protected phosphoramidite of $2'\alpha$ fluorothymine glycol is reported in detail in Supplementary Information (Supplementary Figure S1). In brief, the benzvl protected precursor $2'\alpha$ -F-thymidine was synthesised through a minor modification of the reported synthesis (39). 5'-O-(4,4'-Dimethoxytrityl)-5,6-dihydro-5,6di[(tert-butyl)dimethoxysilyl]-2'α-fluorothymidine was obtained as an 8:1 mixture of two inseparable diastereomers. After comparing with the bibliographic spectra (40), the major diastereomer was assigned to (5R,6S)and the minor to (5S, 6R). The last step of the synthesis was the formation of the (2-cyanoethyl)-N,Ndiisopropylphosphoramidite. Both the reaction and the chromatographic purification of the crude mixture took place under strictly anhydrous and degassed conditions. The desired product was obtained as an 8:1 mixture of the (5R.6S) and (5S.6R) diastereomers.

Patient derived cell lines

All patient-derived cell lines were purchased from the NIGMS Human Genetic Cell Repository, Coriell Institute for Medical Research (Camden, New Jersey, USA). Immortalised skin fibroblasts of the following NER complementation groups were used: XP-A (GM04312), CS-B (GM16095), CS-A (GM16094), and XP-C (GM15983). The reference repair-proficient cell lines were MRC-5VA1 (AG10076) and the cell line derived from GM04312 by complementation with human XPA cDNA (GM15876) (23,41).

Targeted disruption of CSA/ERCC8, DDB2/XPE and NTHL1 genes in HeLa cells

Cells used for gene knockout were monoclonal descendants of HeLa cervical carcinoma cell line (German Collection of Microorganisms and Cell Cultures No. ACC 57). Each gene was targeted by a pair of sgRNAs expressed from the pX330-SpCas9-HF1 vector (Addgene, Watertown, MA, USA) (42), designed to cut on both sides of exon sequences critical for the protein functions. Potentially suitable sgRNA sequences were selected using the CHOPCHOP web tool (43). Clones isolated by single cell sorting of transfected cells were expanded in 96well plates, screened by PCR, and validated by Western blotting, as described previously (44). In the case of CSA knockout, the TC-NER deficiency was independently validated by the host-cell reactivation assay using synthetic 3-(deoxyguanosin- N^2 -yl)-2-acetylaminofluorene (dG(N^2)-AAF) as a specific TC-NER substrate (22). Comprehensive gene targeting results as well as sgRNA and PCR primer sequences are reported in Supplementary Information (Supplementary Figures S2–S4; Supplementary Table S1).

Generation of the reporter constructs carrying specific synthetic DNA modifications

The procedure for efficient site-specific incorporation of synthetic DNA modifications (Figure 1A) into the 5'-



Figure 1. Analysis of transcription blocking capacities of cyclo-dA, cyclo-dG and 2'F-Tg modifications in human XP-A (GM04312) cells. (A) Structures of synthetic DNA modifications incorporated into vector DNA. (B) Schemes indicating positions of the incorporated modifications (marked bold in the sequence) within the BmtI site. Cyclo-dA was incorporated opposite to T, cyclo-dG opposite to C, and 2'F- or deoxy-Tg opposite to A. A pair of vectors with inverted target sequence was used to selectively incorporate synthetic oligonucleotides with modifications into either the transcribed (TS) or non-transcribed DNA strand (NTS). (C) Verification of efficient incorporation of synthetic DNA modifications by BmtI cleavage inhibition. (D) Verification of resistance of 2'F-Tg incorporated into vector DNA to endonuclease III (*Nth*) compared to the 2'-deoxy-Tg substrate. The reference 'no adduct' constructs were obtained by ligation of the modification-free synthetic oligonucleotide. (E) Quantification of EGFP expression in XP-A (GM04312) cells by flow cytometry. EGFP fluorescence distribution plots show overlaid expression data for reporter constructs containing the specified modifications (amber) along with the modification-free constructs (blue). Bar charts on the right show quantification of the EGFP expression. Median fluorescence values were normalised relative to the modification-free constructs and mean relative expression calculated for three to six independent experiments (±SD).

untranslated region of the enhanced green fluorescent protein (EGFP) gene was described previously (22). Expression vectors pZAJ-5C of pZAJ-5W were cut at the available tandem sites with the Nb.BsrDI nicking endonuclease (NEB GmbH, Frankfurt am Main, Germany) in either the transcribed DNA strand (pZAJ-5C) or the opposite strand (pZAJ-5W) to generate 18-nt gaps. Synthetic oligonucleotides containing the specified modifications or unmodified control strands (all HPLC-purified and validated by mass spectrometry) were annealed and ligated to obtain fully double stranded covalently closed reporter constructs for transfection.

5'-Synthetic 18-mer oligonucleotides CATTGCTTCGCT[cyclo-dA]GCACG 5'and CATTGCTTC[cyclo-dG]CTAGCACG containing 5'Sisomers of the specified cyclopurine modifications were purchased from Kaneka Eurogentec S.A. (Seraing, Belgium). For specific incorporation of 5'R and 5'Scyclo-dA stereoisomers, phosphoramidites of the respective 2'-deoxyribonucleosides were synthesised according to the previously described protocol (45), and the oligonucleotides were produced in the Bioorganic Chemistry Department of the Polish Academy of Science (Lodz, Poland), as described previously (46). The 5'-CATTGCTTCGC[Tg]AGCACG oligonucleotides 5'-CATTGCTTCGC[2'F-Tg]AGCACG and containing 2'-deoxy thymine glycol (Tg) or its $2'\alpha$ -fluorinated analog (2'F-Tg) were produced at Metabion GmbH (Planegg, Germany). The oligonucleotide 5'-CATTGC[TT dimer]CGCTAGCACG containing a cis-svn cyclobutane pyrimidine dimer (TT dimer) was from TriLink BioTechnologies (San Diego, CA). Oligonucleotides containing $dG(N^2)$ -AAF and N-(deoxyguanosin-8-yl)-2-acetylaminofluorene (dG(C8)-AAF) at the same position as cyclo-dG were described previously (22). The reference unmodified deoxyribo-oligonucleotide 5'-CATTGCTTCGCTAGCACG was from Eurofins Genomics (Ebersberg, Germany).

Verification of lesion incorporation into reporter vectors

Correct incorporation of synthetic oligonucleotides was monitored by inhibition of ligation by unphosphorylated synthetic strands and, additionally, by formation of covalently closed circular DNA in the presence of polynucleotide kinase, as described previously (37). Incorporation of thymine dimer was further specifically verified by incision with T4 endonuclease V (NEB), as described previously (23). The presence of Tg lesion was verified by an analogous reaction with endonuclease III (NEB), as described previously (37). The presence of cyclo-dA and cyclo-dG adducts was confirmed by inhibition of cleavage by BmtI restriction endonuclease (NEB) at the specific 5'-GCTAGC sequence (Figure 1B and C). The presence of $dG(N^2)$ -AAF and dG(C8)-AAF was verified, as described previously (22). by the inhibition of NheI (NEB) cleavage (Supplementary Figure S5).

EGFP expression and host-cell reactivation assays

Transfection of cells with reporter constructs carrying the specified synthetic modifications were performed exactly as

described previously for other types of DNA lesions (23). Constructs prepared with oligonucleotides with and without modifications were always transfected and analysed in parallel. They were mixed with equal amounts of pDsRed-Monomer-N1 (Clontech, Saint-Germainen-Laye, France) as a marker allowing selective measurement of transfected cells during flow cytometry. EGFP and DsRed expression was analysed in cells fixed 24 h post-transfection using FAC-SCalibur[™] and the CellQuest[™] Pro software (Beckton Dickinson GmbH, Heidelberg, Germany), as described previously (23). Where indicated, CytoFLEX instrument with the CytExpert software (Beckman Coulter GmbH, Krefeld, Germany) was used for analysis with quantitatively identical results. After exclusion of fragments and aggregates, transfected cells were selectively gated by the same DsRed expression threshold to generate the EGFP fluorescence distribution plots shown in figures. Median EGFP fluorescence values (M) of the obtained distributions were used to calculate relative expression as M(construct with modification): M(reference construct without modification).

RESULTS

Cyclopurine lesions but not thymine glycol inhibit transcription by the RNA polymerase-mediated mechanism

Transcription-blocking potentials of nucleobase modifications, can be determined by their incorporation into the transcribed strand of a reporter gene and measurement of the gene expression in a NER-deficient cell model, such as cells derived from XP-A patients (22,27,29). Due to the presence of the covalent 5',8 bonds (Figure 1A), cyclo-dA and cyclo-dG adducts are intrinsically resistant to excision by DNA *N*-glycosylases of the BER pathway and, thereby, cannot be removed from vector DNA transfected to XP-A cells. This is different for Tg lesion, which can be removed from DNA by at least two different DNA *N*-glycosylases (47–51). Therefore, to assess transcription blocking capacity of Tg, we used its BER-resistant 2' α -fluorinated analog (Figure 1A), which was synthesised specifically for this purpose (Supplementary Figure S1).

We used a pair of EGFP expression vectors containing an inverted sequence in the 5'-untranslated region (22) to incorporate cyclo-dA, cyclo-dG or 2'F-Tg into either the transcribed DNA strand or the non-transcribed strand. An efficient and correct incorporation of synthetic oligonucleotides carrying the specified modifications was confirmed by analytical ligation reactions (Supplementary Figure S6). By design, all modifications were placed within the recognition sequence of the BmtI restriction endonuclease. Thereby, the presence of modifications could additionally be validated by prevention of BmtI cleavage (Figure 1B and C). Finally, we compared susceptibilities of synthetic Tg and 2'F-Tg to endonuclease III, which has both DNA N-glycosylase and AP-lyase activities. At the saturating concentration resulting in complete cleavage at 2'deoxyTg, only a non-specific nicking activity (at the levels equivalent to the respective dT controls) was detectable in the plasmid DNA containing the $2'\alpha$ -fluorinated analog (Figure 1D and Supplementary Figure S7). The results thus indicate that 2'F-Tg is BER-resistant, as intended by design.

To assess transcription blocking capacities of cyclo-dA, cyclo-dG, and Tg, the expression of constructs containing the specified modifications in the transcribed DNA strand was measured in NER-deficient XP-A cells. Based on the EGFP fluorescence levels, 2'F-Tg caused a very minor inhibition of the EGFP expression, which can hardly be regarded as biologically significant (Figure 1E). In a sharp contrast, the presence of a single cyclopurine modification led to a 10-fold (cyclo-dA, $P = 7.8 \times 10^{-12}$, heteroscedastic Student's *t*-test) or a 17-fold (cyclo-dG, $P = 2.4 \times 10^{-12}$) reduction of the gene expression, relative to the modificationfree control (Figure 1E, 'TS'). These values reflect exceptionally high transcription-blocking potencies, alike those of cyclobutane pyrimidine dimers or acetylaminofluorene adducts (22). It is important to note that EGFP expression remained unaffected when the same cyclopurine adducts were placed in the opposite DNA strand (Figure 1E, 'NTS'). The strict confinement of the effects of cyclo-dA and cyclo-dG to the transcribed strand provides a proof that impaired gene expression results from interaction of the adducts with components of the transcribing RNA polymerase complex as the only strand-specific biological component.

Removal of transcription-blocking cyclopurine lesions requires XP-A

To investigate whether transcription-blocking cyclo-dA and cyclo-dG can be repaired by reconstitution of NER capacity, we performed the gene expression recovery analyses in XP-A cells complemented by expression of human XPA cDNA (Figure 2). We observed a very significant reactivation of both cyclo-dA (from $9.51 \pm 0.74\%$ to $35.40 \pm 5.47\%$) and cyclo-dG (from 5.50 \pm 0.61% to 55.88 \pm 3.72%) reporter constructs in the XP-A corrected cell lines, indicating that both adducts undergo repair, which requires a functional XPA gene. Even higher levels of reporter reactivation were documented in fully NER-proficient MRC-5 cells. Besides, we have noticed that the recovery of gene expression was more pronounced in the case of cyclodG (74.18 \pm 9.61%) compared to the cyclo-dA construct $(52.33 \pm 5.86\%)$, suggesting that cyclo-dG is repaired more efficiently (Figure 2).

TC-NER is strictly required for the repair of cyclopurine lesions

The requirement of XPA implied that repair of cyclo-dA and cyclo-dG takes place by the NER mechanism. Based on this, we further assessed the requirements of TC-NER and GG-NER for the removal of both types of cyclopurine lesions from transcribed DNA, using cell lines derived from patients with defects in the respective damage recognition components (Figure 3). We found that CS-A and CS-B cells were deficient in the reporter gene reactivation to the same degree as the XP-A cell line, indicating that GG-NER, although unaffected in CS cells, fails to recognise cyclo-dA and cyclo-dG. This notion was further supported by high repair capacities of XP-C cells, which indicated that both cyclo-dA and cyclo-dG in the transcribed DNA strand are recognised very efficiently by the remaining TC-NER.





Figure 2. Recovery of gene expression from the transcription block caused by synthetic cyclo-dA or cyclo-dG modifications after complementation with XPA. Expression of the same constructs was analysed in parallel in the XP-A cells (with and without XPA complementation) and MRC-5 as a NER-proficient control. Fluorescence distribution plots of a representative experiment and quantification of the EGFP expression relative to the modification-free construct (mean of $n \ge 4$ independent experiments, \pm SD; *P*-values calculated by Student's two-tailed *t*-test).



Figure 3. Repair capacities of CS-B (GM16095), CS-A (GM16094), and XP-C (GM15983) cells towards transcription-blocking cyclo-dA or cyclo-dG in the transcribed DNA strand. XP-A cells were transfected with the same constructs as a reference for a total NER defect. Fluorescence distribution plots of a representative experiment and quantification of the EGFP expression relative to the modification-free construct (mean of $n \ge 3$ independent experiments, \pm SD).

Since differences observed between the patients-derived cell lines could potentially be influenced by the individual genetic backgrounds, we also generated isogenic cell lines with specific defects in either TC-NER or GG-NER pathways by disruption of CSA/ERCC8 or DDB2/XPE genes in HeLa cells (Supplementary Figures S2 and S3). Reporter reactivation analyses in these cell lines revealed a severe repair defect in the HeLa-derived TC-NER deficient CSA knockout cell line (Figure 4). Indeed, CSA knockout cells showed relative expression values of $13.96 \pm 0.73\%$ (compared to $47.43 \pm 2.67\%$ in the parental HeLa cell line) for cyclo-dA and $5.21 \pm 0.07\%$ for cyclo-dG (compared to $70.78 \pm 2.63\%$ in HeLa). Since the expression of cyclo-dA constructs in the CSA knockout cell line was barely above the levels attributable to transcriptional bypass in totally NER-deficient XP-A cells (9.51 \pm 0.74%) (Figure 3), we conclude that CSA is of a major importance for cyclo-dA repair. We further conclude that CSA is essential for cyclodG repair, since the reporter expression was negligible in the absence of CSA and did not exceed the level measured in the XP-A cells $(5.5 \pm 0.62\%)$.

In contrast, *DDB2* knockout cells retained essentially full repair capacities for both cyclo-dA (44.47 \pm 3.62% relative EGFP expression compared to 47.43 \pm 2.67% in the parental cell line) and cyclo-dG (66.03 \pm 2.21% compared to 70.78 \pm 2.63% in the parental cell line). Collectively, the results in the isogenic *DDB2* and *CSA* knockout cell lines exactly recapitulated the TC-NER and GG-NER phenotypes of the patients-derived cell lines (Figure 3), thereby corroborating the conclusions that repair of cyclopurine nucleotides in the transcribed strand of the reporter gene is dominated by TC-NER whereas GG-NER contribution to total repair capacity is negligible.

GG-NER of non-cyclopurine lesions is excellently detectable in the transcribed DNA strand

Considering a peculiar chromatin structure of the transfected plasmid vectors (38) and, especially, the fact that the reporter-based repair assay is confined to the transcribed



Figure 4. Repair capacities of HeLa ('WT') and the derived knockout cell lines ('CSA ko' and 'DDB2 ko') towards transcription-blocking cyclo-dA or cyclo-dG in the transcribed DNA strand, analysed by the CytoFLEX flow cytometer. Host-cell reactivation of reporter constructs containing synthetic dG(C8)-AAF and TT dimer (T \Box T), as NER substrates processed by both TC-NER and GG-NER, is shown for comparison. Fluorescence distribution plots of a representative experiment and quantification of the EGFP expression relative to the modification-free construct (mean of *n* = 4 independent experiments, ±SD).

DNA strand, we questioned whether GG-NER would be as readily detected within the present experimental system as TC-NER. We therefore next determined repair efficiencies of two more synthetic adducts, a cis-syn cyclobutane pyrimidine dimer (TT dimer) and an N-(deoxyguanosin-8-yl)-2acetylaminofluorene (dG(C8)-AAF), both of which have very strong transcription-blocking potentials (22) and are well characterised substrates for both TC-NER (9,52,53) and GG-NER (14,54,55). In XP-A cells, both dG(C8)-AAF and TT dimer caused a nearly absolute abolition of transcription, very much alike cyclo-dA and cyclo-dG (Supplementary Figures S5 and S8). Remarkably, however, expression of the dG(C8)-AAF construct was reactivated to the level of the adduct-free control not only in HeLa cells but also in the DDB2 and CSA knockouts. Such a complete reactivation should indicate that either GG-NER or TC-NER alone is sufficient for the dG(C8)-AAF repair, and that GG-NER is well detectable in the transcribed DNA strand.

For the construct containing TT dimer, relative expression recovered to $38.97 \pm 5.60\%$ under the *CSA* knockout conditions (compared to $74.27 \pm 9.30\%$ in the parental cell line). Thus, despite the well recognised fact that TT dimers

are repaired by TC-NER far more efficiently than by GG-NER (14,54–56), the impact of *CSA* knockout on the capacity of cells to repair TT dimer was by far milder than observed for cyclopurine adducts (Figure 4). Combined, the results indicate that GG-NER provides a significant backup repair activity for the removal of TT-dimers and even more for dG(C8)-AAF. On the other hand, GG-NER is largely or entirely ineffective against cyclo-dA or cyclo-dG.

GG-NER makes a minimal or no contribution to the repair of both cyclo-dA diastereoisomers

According to the manufacturer's communication (Kaneka Eurogentec S.A.), oligonucleotide strands used for construction of the reporters described above were synthesised with 5'S diastereoisomers of the cyclopurine 2'-deoxyribonucleosides. Thereby our previous conclusions about TC-NER requirements for the repair were limited to 5'S-cyclo-dG and 5'S-cyclo-dA. It was reported in the literature that 5'R stereoisomers of cyclo-dA or cyclo-dG are slightly better NER substrates under cell-free conditions than their 5'S counterparts (28). These observations should

refer specifically to GG-NER activity because the assay used by the authors does not detect TC-NER. We therefore questioned whether GG-NER significantly contributes to repair of 5'R-cyclo-dA in the cellular context. Cyclo-dA was chosen, because data obtained with its 5'S stereoisomer did not entirely exclude a minor contribution of GG-NER to the repair. We performed stereoselective synthesis of 5'R- and 5'S-cyclo-dA phosphoramidites, synthesized the respective oligonucleotides, as specified under Materials and Methods, and incorporated them into the transcribed DNA strand of the reporter EGFP gene (Figure 5A).

Expression analyses in the XP-A cell line, showed that both cyclo-dA stereoisomers inflicted strong transcription blockage, 5'R-cyclo-dA being bypasses with a slightly higher efficiency (14.3 \pm 0.58% EGFP expression relative to the adduct-free construct) than the 5'S counterpart $(5.93 \pm 0.28\%)$ (Figure 5B). In fully NER-proficient HeLa cells the expression recovered to the levels of $44 \pm 0.94\%$ (5'R) and $51.81 \pm 4.45\%$ (5'S), suggesting that 5'S-cyclodA may be repaired with a slightly higher efficiency. The reporter reactivation rates were not significantly affected by the DDB2 knockout (Figure 5C). In contrast, CSA knockout led to strong decrease of the EGFP expression in comparison to the parental cell line, which was highly significant for both cyclo-dA stereoisomers ($P = 9.0 \times 10^{-7}$ for 5'R- and $P = 8.5 \times 10^{-7}$ for 5'S-cyclo-dA, heteroscedastic Student's *t*-test). The expression levels in the absence of CSA were as low as in the XP-A cell line (for 5'S isomer) or slightly above the XP-A value (for 5'R). The results thus indicate that TC-NER is by far more efficient than GG-NER for both cyclo-dA stereoisomers and, within the accuracy margins of the host-cell reactivation assay, the only NER subpathway contributing to the removal of 5'S-cyclo-dA from the transcribed DNA strand. Also for 5'R-cyclo-dA, TC-NER is clearly the primary repair mechanism, although a minor contribution of GG-NER cannot be entirely excluded for this stereoisomer.

Thymine glycol is predominantly processed by BER

Excision of DNA damage can lead to impairment of transcription via conflicts between its components and the repair proteins or the strand-cleaved intermediates (9,57). During NER, such conflicts are probably prevented by the multiprotein transcription factor IIH (TFIIH), which has essential functions in both DNA repair reactions and transcription. Accordingly, active NER leads to reactivation of the reporter gene expression by removing transcriptionblocking adducts, as reported previously (22) and seen in the experiments described above (Figure 2). It needs to be noted that BER, in contrast to NER, typically has a different outcome for the reporter gene expression. BER starts from recognition and release of a modified nucleobase by the specific DNA N-glycosylase and proceeds with cleavage of the abasic deoxyribose remnant by the endo-/exonuclease APE1. Such a mode of excision commonly leads to transcriptional silencing, as reported for several BER substrates (36,44,58–60). So, in many cell lines, including HeLa, BER activity can be deduced based on a diminished gene expression even if the primary DNA modification does not impede the RNA polymerase progression as such (36,59). Returning to the question about the primary repair mechanism of non-transcription-blocking Tg, we therefore further compared the expression of reporter constructs containing BER-resistant 2'F-Tg and its physiological 2'-deoxy counterpart in HeLa cells. Moreover, we examined the role of NTHL1—one of the DNA glycosylases known to recognise Tg—by disrupting the *NTHL1* gene (Figure 6A).

As expected from previous experiments in the XP-A cell line (Figure 1E), 2'F-Tg did not cause an appreciable decrease of EGFP expression also in HeLa cells (Figure 6B). In contrast, a strong decrease of gene expression was caused by 2'-deoxy-Tg in either the transcribed or the nontranscribed DNA strand. The absence of strand specificity indicated that the decrease of transcription did not require a direct interaction of elongating RNA polymerase with primary DNA modification. Moreover, the fact that such an effect was not caused by 2'F-Tg strongly suggested its dependence on BER. We therefore disrupted the NTHL1 gene in HeLa cells (Supplementary Figure S4), which fully reversed the inhibitory effect of 2'-deoxy-Tg (Figure 6). The expression levels of constructs carrying 2'-deoxy- or 2'F-Tg were no longer different in the NTHL1 knockout cells, which implies that ablation of NTHL1 is equivalent to total BER deficiency. Combined, our findings imply that Tg is primarily processed by BER. Although the fraction of Tg lesions excised by BER cannot be determined exactly, based on more than a 2-fold drop in the gene expression levels $(35.33 \pm 0.37\%$ residual expression for Tg in the transcribed strand, $P = 1.07 \times 10^{-5}$; $41.35 \pm 2.62\%$ in the opposite strand, $P = 6.66 \times 10^{-4}$) we conclude that BER dominates even in fully NER-proficient HeLa cells.

DISCUSSION

Potential diversity of repair mechanisms of Tg in mammalian cells remained the matter of debate despite characterisation of NTHL1 and NEIL1 as the main specific DNA glycosylases (47–49). Due to its non-planar structure, Tg causes a significant destabilisation of DNA double helix (61,62) and, at least under specific circumstances, can also cause blockage or pausing of various RNA polymerases (63-66). Based on these properties, Tg is regarded by many as a potential TC-NER substrate. However, such a view is hard to reconcile with the observation that mammalian RNA polymerase II efficiently bypasses Tg in the template DNA strand under cell-free conditions (67). Several attempts to demonstrate TC-NER of Tg in cells were undertaken, but all the reports were subsequently disclosed as invalid, leaving the question open (https: //www.science.org/doi/full/10.1126/science.300.5626.1657b; https://aacrjournals.org/cancerres/article/63/13/3846/ 510149/Retraction; https://www.sciencedirect.com/science/ article/pii/S1568786402002367). In our opinion, discrepancies between the previous publications could be explained by uncontrolled BER activity, which would generate apyrimidinic lesion at the Tg site. This is reminiscent of another common oxidative product in DNA, 8-oxo-7,8dihydroguanine (8-oxoG). Unless cleaved by the respective

DNA glycosylase OGG1, 8-oxoG does not block transcription to any significant degree (36). Nonetheless, evidence suggests that TC-NER contributes to the repair, likely at



Figure 5. Transcription blockage by 5'*R* and 5'*S* cyclo-dA diastereoisomers and the impact of DDB2 and CSA knockouts on the repair. (A) Chemical structures of (5'*R*)- and (5'*S*)-cyclo-dA and verification of efficient incorporation of synthesised oligonucleotides containing the specific cyclo-dA diastereoisomers into the pZAJ-5C transcribed strand by the BmtI inhibition. Constructs obtained by ligation of unmodified oligonucleotide ('no adduct') or the commercial cyclo-dA/cyclo-dG oligonucleotides (as in Figure 1) are shown for comparison. (B) Transcription blockage in the XP-A (GM04312) cell line. Fluorescence distribution plots of a representative experiment and quantification of the EGFP expression relative to the modification-free construct (mean of n = 4 independent experiments, \pm SD). (C) Host-cell reactivation in HeLa ('WT') and the derived knockout cell lines ('CSA ko' and 'DDB2 ko'). (B and C) EGFP expression in cells transfected with constructs harbouring (5'*R*)- and (5'*S*)-cyclo-dA was analysed by the CytoFLEX flow cytometer. Fluorescence distribution plots of representative experiments and quantification of the EGFP expression relative to the modification-free construct are shown (mean of n = 4 independent experiments, \pm SD).

the level of a post-excision BER intermediate (68,69). With the BER-resistant Tg analog, we now could ascertain that the lesion is efficiently bypassed during transcription, since there was no strand-specific effect of 2'F-Tg on the reporter gene expression levels in NER-deficient cells (Figure 1). Since initiation of TC-NER requires transcription blockage (70), our results rule this pathway out as a conceivable repair mechanism for Tg. Still, the results do not formally exclude a possibility of GG-NER. A strand cleavage pattern characteristic for NER was previously reported in biochemically reconstituted repair reactions of DNA containing synthetic Tg (30). However, biological relevance of either GG-NER or a NEIL1-mediated BER appears to be low in the cellular context, as demonstrated by different outcomes of 2'-deoxy-Tg in the isogenic NTHL1-proficient versus knockout cell lines (Figure 6). Our results imply that



Figure 6. Impact of Tg (2'-deoxy and 2'F) lesions on the EFGP gene expression in HeLa ('WT') and the derived *NTHL1* knockout cell line ('NTHL1 ko'). (**A**) Validation of the NTHL1 knockout in selected clones by Western blotting with EPR15930 antibody (Abcam, ab191413). The last two lanes show parental 'WT' cell line after transfection with short interfering RNA against NTHL1 (siNTHL1) or unspecific short RNA (siCtrl). (**B**) Representative EGFP fluorescence distribution plots of cells transfected with constructs containing 2'-deoxy-Tg or 2'F-Tg in the specified DNA strands (TS or NTS) and the relative EGFP expression values (mean of n = 3 independent experiments, \pm SD).

NTHL1 recognises Tg in transcribed DNA in a strandindependent manner and processes it far more efficiently than any potentially concurrent mechanism.

Cyclo-dA and cyclo-dG were previously characterised as NER substrates (26,28). Elevated levels of cyclopurine lesions were reported in primary XP-C keratinocytes and fibroblasts (71) as well as in CS-A and CS-B cells (72), however neither the causality nor the relative contributions of GG-NER and TC-NER to the repair in cells was established in these studies. Biochemical reconstitution of TC-NER is exceedingly difficult; however, poor transcriptional bypass of cyclopurine lesions under cell-free conditions suggested a possibility of their recognition by TC-NER (29). Accordingly, analyses in NER-deficient rodent and human cells confirmed that both lesions strongly decreased the reporter gene expression rates (27,29), which was attributed partly to blockage and partly to erroneous bypass during transcription (29,73). In NER-proficient 293T cells, reactivation rates of the reporter constructs were approximately 30% and about one third of the reactivation capacity was lost after treatment with short interfering RNA against CSB, suggesting that the repair is partly CSB-dependent (29).

To obtain more conclusive information about repair pathways of cyclopurine nucleotides, we enhanced several parameters compared to the experimental procedures used in previous studies. Firstly, we used an improved design of reporter constructs carrying synthetic cyclo-dA and cyclodG. Both modifications were incorporated selectively into either the template or the non-template DNA strand. By this, the impairment of gene expression could be assigned to direct interactions of cyclo-dA and cyclo-dG with transcribing RNA polymerase complexes (Figure 1). Of note, the incorporation sites were chosen within the untranslated gene region to eliminate potential effects of transcriptional mutagenesis on the EGFP expression levels. Secondly, we refined the procedure for quantitative measurement of transcription blockage and recovery by recording the distribution of the EGFP expression over the whole population of individual cells, visualised by co-expression of a transfection marker. Finally, to strictly assign the repair of cyclodA and cyclo-dG to the specific NER subpathway, we applied cell lines with genetically defined and phenotypically confirmed total defects in the specific GG-NER and TC-NER components, including isogenic cell pairs generated by gene complementation and knockout. Our results provided unequivocal evidence that disruption of the central GG-NER genes *XPC* and *XPE/DDB2* does not affect the repair capacities of cells, whereas deletions of either *CSA* or *CSB* genes have equally severe consequences, manifested as total repair defects (Figures 3 and 4). These two lines of evidence independently indicate that TC-NER is the only relevant pathway for cyclo-dA and cyclo-dG repair.

As a potential limitation of the used methodology and of transfection-based DNA repair assays in general, it is necessary to consider that chromatin folding of vector DNA is not identical to nuclear chromatin, where the access of DNA repair factors could be limited by nucleosomes or the higher-order structures. However, in our opinion, it is unlikely that apparent absence of GG-NER of cyclopurine lesions in the transiently transfected gene is caused by a restricted access of the repair components to damage sites in DNA, since a less condensed chromatin structure should rather have an opposite effect. It is remarkable that two other well-characterised NER substrates-cyclobutane TT dimer and, especially, the dG(C8)-AAF adduct—display different balance between GG-NER and TC-NER than cyclo-dA and cyclo-dG. Thus GG-NER alone was sufficient for full repair of dG(C8)-AAF, as measured by the host-cell reactivation assay. TT dimer is a relatively poor GG-NER substrate and has long served as paradigmatic type of DNA damage for TC-NER studies (14,54-56). Nonetheless, we observed significant levels of its repair in the absence of CSA (Figure 4) or CSB (22), which indicates that even the lowest GG-NER levels were readily detectable at our assay settings. Consequently, the absent repair ac-



Figure 7. Model predicting dynamics of DNA lesions over the lifetime of a cell under constant exposure to endogenously arising DNA damaging agents (represented with a lightning bolt). Different shapes represent structurally different modification types, which are recognised by different DNA repair pathways. The levels in transcribed DNA strands of active genes (pink strings with broken arrows) are compared to the rest of the genome DNA. (A) In fully repair proficient cells, BER and GG-NER substrates (squares and circles) remain at constant levels on average in the genome, since their generation and repair rates are equal under physiological conditions (when repair capacities are not oversaturated). In contrast, TC-NER substrates (yellow triangles) remain at a steady level only in transcribed DNA strands of active genes, where the repair is efficient. The very same lesion types will accumulate all over the non-transcribed DNA sequences (red triangles on the black strings), but their burden can be tolerated, particularly in non-replicating cell types, as far as functionally critical genome regions remain unaffected. This is different in Cockayne syndrome, where TC-NER deficiency will cause the damage to accumulate also in transcribed segments of the genome, where the tolerance threshold is low. (**B**) Expected dynamics of accumulation of cyclopurine DNA lesions (whose repair explicitly requires TC-NER) in transcribed strands of the actively expressed genes (hot pink profiles) and in the rest of the genome (a lighter pink) in the repair proficient versus Cockayne syndrome cells.

tivity in CS-B, CS-A (Figure 3), and HeLa-derived CSA knockout (Figures 4 and 5) cell models means that the remaining GG-NER pathway does not process 5'S stereoisomers of cyclopurine nucleotides at all. This conclusion is reinforced by re-evaluation of published quantitative data from biochemical reconstitution of human GG-NER. Indeed, although cyclopurine nucleotides are processed by NER, their excision efficiencies are 40 to 150 times lower compared to a cis-platin adduct as a canonical reference substrate (26). Moreover, 5'S stereoisomers of both cyclodA and cyclo-dG lesions destabilise the DNA double helix to lesser degrees than the 5'R counterparts and, accordingly, are even poorer GG-NER substrates (28). The cumulative evidence thus corroborates the conclusion that cellular GG-NER capacity is insufficient to warrant a biologically meaningful repair of 5'S stereoisomers of cyclopurine nucleotides.

Mutations in genes involved in TC-NER commonly result in developmental phenotypes, including early cessation of growth, microcephaly, multiple neurodegenerative features, and a greatly reduced life expectancy. Most affected in CS are neural tissues in the brain that are characterised by high levels of oxidative metabolism, high transcription rates and a slow or absent cell division (4,74). Even under physiologically low rates of damage generation, the deficit of a suitable repair mechanism in the absence of replication should lead to irreversible accumulation of modifications in DNA of such cells. If both GG-NER and TC-NER are impaired, multiple types of lesions would accumulate, leading to XP neurological disease (2). If only TC-NER is absent, only DNA modifications lacking structural determinants for efficient recognition by GG-NER would accumulate, leading to CS (Figure 7). Being refractory to GG-NER and strongly transcription blocking, cyclopurine lesions match the criteria of such cytotoxic DNA lesions.

DATA AVAILABILITY

CHOPCHOP is a web tool for selecting target sites for CRISPR/Cas9, CRISPR/Cpf1, CRISPR/Cas13 or NICKASE/TALEN-directed mutagenesis (https: //chopchop.cbu.uib.no/).

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank IMB Flow Cytometry Core Facility (Institute of Molecular Biology Mainz) for help with cell sorting and Service Project Z01 of the Collaborative Research Center SFB1361 for efforts towards immunological detection of thymine glycol. We are thankful to Antonio Manetto for advising us on the phosphoramidite and oligonucleotide synthesis; to Michael Musheev and Christoph Niehrs (Institute of Molecular Biology Mainz) for helpful discussions and for sharing unpublished data; to Wynand Roos and Anna Frumkina (Institute of Toxicology, University Medical Center Mainz) for help with optimisation of experimental techniques; and to former Khobta lab members, Nadine Müller, Aalaa Farag and Marta Rodriguez-Alvarez, who contributed at various stages to the *NTHL1* knockout.

FUNDING

Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) [KH263/4, Project-ID 393547839 -SFB 1361 INST 247/926 to A.K]; Deutsche Akademische Austausch Dienst (DAAD, German Academic Exchange Service) [57388935, 57547141, 57623644 to A.K.]: Thüringer Aufbaubank co-funded by the European Union (EFRE-Programm 2021-2027 Thüringen) [2022 FGI 0006 to A.K.]; Special Account for Research Grants of the National and Kapodistrian University of Athens [Research Code 17906 to T.G.]; Greek State Scholarship Foundation (IKY) and the German Society for Academic Exchanges (Deutscher Akademischer Austauschdienst - DAAD) IKY-DAAD 2020 research scholarship program [282/22.09.2020 to T.G.]; Agence nationale de la recherche (ANR, French National Research Agency) [ANR-18-CE92-0013 'TG-TOX' to A.C.]; CEA Radiobiology program [to A.C.]; Campus France [the collaborative program 'Partenariats Hubert Curien (PHC)-PROCOPE' grant to A.C.]. Funding for open access charge: DFG, German Research Foundation.

Conflict of interest statement. None declared.

REFERENCES

- Cleaver, J.E. (1968) Defective repair replication of DNA in xeroderma pigmentosum. *Nature*, 218, 652–656.
- Cleaver, J.E., Lam, E.T. and Revet, I. (2009) Disorders of nucleotide excision repair: the genetic and molecular basis of heterogeneity. *Nat. Rev. Genet.*, 10, 756–768.
- 3. Lehmann, A.R., McGibbon, D. and Stefanini, M. (2011) Xeroderma pigmentosum. *Orphanet. J. Rare. Dis.*, **6**, 70.
- 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.
- Ljungman,M. and Zhang,F. (1996) Blockage of RNA polymerase as a possible trigger for u.v. light-induced apoptosis. *Oncogene*, 13, 823–831.
- Derheimer, F.A., O'Hagan, H.M., Krueger, H.M., Hanasoge, S., Paulsen, M.T. and Ljungman, M. (2007) RPA and ATR link transcriptional stress to p53. *Proc. Natl. Acad. Sci. U.S.A.*, 104, 12778–12783.
- 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.
- Wang, W., Xu, J., Chong, J. and Wang, D. (2018) Structural basis of DNA lesion recognition for eukaryotic transcription-coupled nucleotide excision repair. *DNA Repair (Amst.)*, **71**, 43–55.
- Lans, H., Hoeijmakers, J.H.J., Vermeulen, W. and Marteijn, J.A. (2019) The DNA damage response to transcription stress. *Nat. Rev. Mol. Cell Biol.*, 20, 766–784.
- Gaul, L. and Svejstrup, J.Q. (2021) Transcription-coupled repair and the transcriptional response to UV-irradiation. *DNA Repair (Amst.)*, 107, 103208.

- Gonzalez, N., M., B. and Svejstrup, J.Q. (2021) Causes and consequences of RNA polymerase II stalling during transcript elongation. *Nat. Rev. Mol. Cell Biol.*, 22, 3–21.
- Oh, J., Xu, J., Chong, J. and Wang, D. (2021) Molecular basis of transcriptional pausing, stalling, and transcription-coupled repair initiation. *Biochim. Biophys. Acta Gene Regul. Mech.*, 1864, 194659.

Nucleic Acids Research, 2023 11

- van den Heuvel, D., van der Weegen, Y., Boer, D.E.C., Ogi, T. and Luijsterburg, M.S. (2021) Transcription-coupled DNA repair: from mechanism to Human disorder. *Trends Cell Biol.*, 31, 359–371.
- Sugasawa,K., Okamoto, T., Shimizu,Y., Masutani,C., Iwai,S. and Hanaoka,F. (2001) A multistep damage recognition mechanism for global genomic nucleotide excision repair. *Genes Dev.*, 15, 507–521.
- Isaacs, R.J. and Spielmann, H.P. (2004) A model for initial DNA lesion recognition by NER and MMR based on local conformational flexibility. *DNA Repair (Amst.)*, 3, 455–464.
- Wittschieben, B.O., Iwai, S. and Wood, R.D. (2005) DDB1-DDB2 (xeroderma pigmentosum group E) protein complex recognizes a cyclobutane pyrimidine dimer, mismatches, apurinic/apyrimidinic sites, and compound lesions in DNA. J. Biol. Chem., 280, 39982–39989.
- Camenisch, U., Dip, R., Schumacher, S.B., Schuler, B. and Naegeli, H. (2006) Recognition of helical kinks by xeroderma pigmentosum group A protein triggers DNA excision repair. *Nat. Struct. Mol. Biol.*, 13, 278–284.
- Min, J.H. and Pavletich, N.P. (2007) Recognition of DNA damage by the Rad4 nucleotide excision repair protein. *Nature*, 449, 570–575.
- Scrima,A., Konickova,R., Czyzewski,B.K., Kawasaki,Y., Jeffrey,P.D., Groisman,R., Nakatani,Y., Iwai,S., Pavletich,N.P. and Thoma,N.H. (2008) Structural basis of UV DNA-damage recognition by the DDB1-DDB2 complex. *Cell*, **135**, 1213–1223.
- Chen,X., Velmurugu,Y., Zheng,G., Park,B., Shim,Y., Kim,Y., Liu,L., Van Houten,B., He,C., Ansari,A. *et al.* (2015) Kinetic gating mechanism of DNA damage recognition by Rad4/XPC. *Nat. Commun.*, 6, 5849.
- Mu,H., Geacintov,N.E., Broyde,S., Yeo,J.E. and Scharer,O.D. (2018) Molecular basis for damage recognition and verification by XPC-RAD23B and TFIIH in nucleotide excision repair. *DNA Repair* (*Amst.*), **71**, 33–42.
- 22. Kitsera, N., Gasteiger, K., Luhnsdorf, B., Allgayer, J., Epe, B., Carell, T. and Khobta, A. (2014) Cockayne syndrome: varied requirement of transcription-coupled nucleotide excision repair for the removal of three structurally different adducts from transcribed DNA. *PLoS One*, 9, e94405.
- Kitsera, N., Rodriguez-Alvarez, M., Emmert, S., Carell, T. and Khobta, A. (2019) Nucleotide excision repair of abasic DNA lesions. *Nucleic Acids Res.*, 47, 8537–8547.
- Brooks, P.J. (2007) The case for 8,5'-cyclopurine-2'-deoxynucleosides as endogenous DNA lesions that cause neurodegeneration in xeroderma pigmentosum. *Neuroscience*, 145, 1407–1417.
- Wang, Y. (2008) Bulky DNA lesions induced by reactive oxygen species. *Chem. Res. Toxicol.*, 21, 276–281.
- Kuraoka, I., Bender, C., Romieu, A., Cadet, J., Wood, R.D. and Lindahl, T. (2000) Removal of oxygen free-radical-induced 5',8-purine cyclodeoxynucleosides from DNA by the nucleotide excision-repair pathway in human cells. *Proc. Natl. Acad. Sci. U.S.A.*, 97, 3832–3837.
- Brooks, P.J., Wise, D.S., Berry, D.A., Kosmoski, J.V., Smerdon, M.J., Somers, R.L., Mackie, H., Spoonde, A.Y., Ackerman, E.J., Coleman, K. *et al.* (2000) The oxidative DNA lesion 8,5'-(S)-cyclo-2'-deoxyadenosine is repaired by the nucleotide excision repair pathway and blocks gene expression in mammalian cells. *J. Biol. Chem.*, **275**, 22355–22362.
- Kropachev,K., Ding,S., Terzidis,M.A., Masi,A., Liu,Z., Cai,Y., Kolbanovskiy,M., Chatgilialoglu,C., Broyde,S., Geacintov,N.E. *et al.* (2014) Structural basis for the recognition of diastereomeric 5',8-cyclo-2'-deoxypurine lesions by the human nucleotide excision repair system. *Nucleic Acids Res.*, 42, 5020–5032.
- You, C., Dai, X., Yuan, B., Wang, J., Wang, J., Brooks, P.J., Niedernhofer, L.J. and Wang, Y. (2012) A quantitative assay for assessing the effects of DNA lesions on transcription. *Nat. Chem. Biol.*, 8, 817–822.
- Reardon, J.T., Bessho, T., Kung, H.C., Bolton, P.H. and Sancar, A. (1997) In vitro repair of oxidative DNA damage by human nucleotide excision repair system: possible explanation for neurodegeneration in

xeroderma pigmentosum patients. Proc. Natl. Acad. Sci. U.S.A., 94, 9463–9468.

- Chatgilialoglu, C., Bazzanini, R., Jimenez, L.B. and Miranda, M.A. (2007) (5'S)- and (5'R)-5',8-cyclo-2'-deoxyguanosine: mechanistic insights on the 2'-deoxyguanosin-5'-yl radical cyclization. *Chem. Res. Toxicol.*, 20, 1820–1824.
- Boussicault, F., Kaloudis, P., Caminal, C., Mulazzani, Q.G. and Chatgilialoglu, C. (2008) The fate of C5' radicals of purine nucleosides under oxidative conditions. J. Am. Chem. Soc., 130, 8377–8385.
- Belmadoui, N., Boussicault, F., Guerra, M., Ravanat, J.L., Chatgilialoglu, C. and Cadet, J. (2010) Radiation-induced formation of purine 5', 8-cyclonucleosides in isolated and cellular DNA: high stereospecificity and modulating effect of oxygen. *Org. Biomol. Chem.*, 8, 3211–3219.
- 34. Guerrero, C.R., Wang, J. and Wang, Y. (2013) Induction of 8,5'-cyclo-2'-deoxyadenosine and 8,5'-cyclo-2'-deoxyguanosine in isolated DNA by Fenton-type reagents. *Chem. Res. Toxicol.*, 26, 1361–1366.
- Cadet, J., Davies, K.J.A., Medeiros, M.H., Di Mascio, P. and Wagner, J.R. (2017) Formation and repair of oxidatively generated damage in cellular DNA. *Free Radic. Biol. Med.*, **107**, 13–34.
- 36. Kitsera, N., Stathis, D., Luhnsdorf, B., Muller, 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.
- Luhnsdorf, B., Kitsera, N., Warken, D., Lingg, T., Epe, B. and Khobta, A. (2012) Generation of reporter plasmids containing defined base modifications in the DNA strand of choice. *Anal. Biochem.*, 425, 47–53.
- Khobta, A., Anderhub, S., Kitsera, N. and Epe, B. (2010) Gene silencing induced by oxidative DNA base damage: association with local decrease of histone H4 acetylation in the promoter region. *Nucleic Acids Res.*, 38, 4285–4295.
- Ikeda,H., Fernandez,R., Wilk,A., Barchi,J.J. Jr., Huang,X. and Marquez,V.E. (1998) The effect of two antipodal fluorine-induced sugar puckers on the conformation and stability of the Dickerson-Drew dodecamer duplex [d(CGCGAATTCGCG)]2. *Nucleic Acids Res.*, 26, 2237–2244.
- 40. Doi,Y., Katafuchi,A., Fujiwara,Y., Hitomi,K., Tainer,J.A., Ide,H. and Iwai,S. (2006) Synthesis and characterization of oligonucleotides containing 2'-fluorinated thymidine glycol as inhibitors of the endonuclease III reaction. *Nucleic Acids Res.*, 34, 1540–1551.
- Levy, D.D., Saijo, M., Tanaka, K. and Kraemer, K.H. (1995) Expression of a transfected DNA repair gene (XPA) in xeroderma pigmentosum group A cells restores normal DNA repair and mutagenesis of UV-treated plasmids. *Carcinogenesis*, 16, 1557–1563.
- Bauer, D.E., Canver, M.C. and Orkin, S.H. (2015) Generation of genomic deletions in mammalian cell lines via CRISPR/Cas9. J. Vis. Exp., 95, e52118.
- Labun,K., Montague,T.G., Krause,M., Torres Cleuren,Y.N., Tjeldnes,H. and Valen,E. (2019) CHOPCHOP v3: expanding the CRISPR web toolbox beyond genome editing. *Nucleic Acids Res.*, 47, W171–W174.
- Muller, N., Ponkkonen, E., Carell, T. and Khobta, A. (2021) Direct and base excision repair-mediated regulation of a GC-rich cis-element in response to 5-formylcytosine and 5-carboxycytosine. *Int. J. Mol. Sci.*, 22, 11025.
- 45. Romieu,A., Gasparutto,D. and Cadet,J. (1999) Synthesis and characterization of oligonucleotides containing 5',8-cyclopurine 2'-deoxyribonucleosides: (5'R)-5',8-cyclo-2'-deoxyadenosine, (5'S)-5',8-cyclo-2'-deoxyguanosine, and (5'R)-5',8-cyclo-2'-deoxyguanosine. *Chem. Res. Toxicol.*, **12**, 412–421.
- K. J. S. Josefel Z. deckygalabolic. *Entrif.* 163, 103(10), 12, 112 444
 Kazmierczak-Baranska, J., Boguszewska, K., Szewczuk, M. and Karwowski, B.T. (2021) Effects of 5', 8'-cyclo-2'-deoxypurines on the base excision repair of clustered DNA lesions in nuclear extracts of the XPC cell line. *Cells*, 10, 3254.
- Takao, M., Kanno, S., Kobayashi, K., Zhang, Q.M., Yonei, S., van der Horst, G.T. and Yasui, A. (2002) A back-up glycosylase in Nth1 knock-out mice is a functional Nei (endonuclease VIII) homologue. *J. Biol. Chem.*, 277, 42205–42213.
- Rosenquist,T.A., Zaika,E., Fernandes,A.S., Zharkov,D.O., Miller,H. and Grollman,A.P. (2003) The novel DNA glycosylase, NEIL1,

protects mammalian cells from radiation-mediated cell death. DNA Repair (Amst.), 2, 581–591.

- Katafuchi, A., Nakano, T., Masaoka, A., Terato, H., Iwai, S., Hanaoka, F. and Ide, H. (2004) Differential specificity of human and Escherichia coli endonuclease III and VIII homologues for oxidative base lesions. J. Biol. Chem., 279, 14464–14471.
- Miller,H., Fernandes,A.S., Zaika,E., McTigue,M.M., Torres,M.C., Wente,M., Iden,C.R. and Grollman,A.P. (2004) Stereoselective excision of thymine glycol from oxidatively damaged DNA. *Nucleic Acids Res.*, **32**, 338–345.
- 51. Ocampo-Hafalla,M.T., Altamirano,A., Basu,A.K., Chan,M.K., Ocampo,J.E., Cummings,A. Jr., Boorstein,R.J., Cunningham,R.P. and Teebor,G.W. (2006) Repair of thymine glycol by hNth1 and hNeil1 is modulated by base pairing and cis-trans epimerization. *DNA Repair (Amst.)*, 5, 444–454.
- Tornaletti,S., Reines,D. and Hanawalt,P.C. (1999) Structural characterization of RNA polymerase II complexes arrested by a cyclobutane pyrimidine dimer in the transcribed strand of template DNA. J. Biol. Chem., 274, 24124–24130.
- Brueckner, F., Hennecke, U., Carell, T. and Cramer, P. (2007) CPD damage recognition by transcribing RNA polymerase II. *Science*, 315, 859–862.
- Fitch, M.E., Nakajima, S., Yasui, A. and Ford, J.M. (2003) In vivo recruitment of XPC to UV-induced cyclobutane pyrimidine dimers by the DDB2 gene product. *J. Biol. Chem.*, 278, 46906–46910.
- 55. Reardon, J.T. and Sancar, A. (2003) Recognition and repair of the cyclobutane thymine dimer, a major cause of skin cancers, by the human excision nuclease. *Genes Dev.*, **17**, 2539–2551.
- Mellon, I., Bohr, V.A., Smith, C.A. and Hanawalt, P.C. (1986) Preferential DNA repair of an active gene in human cells. *Proc. Natl. Acad. Sci. U.S.A.*, 83, 8878–8882.
- 57. Khobta, A. and Epe, B. (2012) Interactions between DNA damage, repair, and transcription. *Mutat. Res.*, **736**, 5–14.
- Luhnsdorf, B., Epe, B. and Khobta, A. (2014) Excision of uracil from transcribed DNA negatively affects gene expression. J. Biol. Chem., 289, 22008–22018.
- Allgayer, J., Kitsera, N., Bartelt, S., Epe, B. and Khobta, A. (2016) Widespread transcriptional gene inactivation initiated by a repair intermediate of 8-oxoguanine. *Nucleic Acids Res.*, 44, 7267–7280.
- Kitsera, N., Allgayer, J., Parsa, E., Geier, N., Rossa, M., Carell, T. and Khobta, A. (2017) Functional impacts of 5-hydroxymethylcytosine, 5-formylcytosine, and 5-carboxycytosine at a single hemi-modified CpG dinucleotide in a gene promoter. *Nucleic Acids Res.*, 45, 11033–11042.
- Kung,H.C. and Bolton,P.H. (1997) Structure of a duplex DNA containing a thymine glycol residue in solution. *J. Biol. Chem.*, 272, 9227–9236.
- 62. Iwai,S. (2001) Synthesis and thermodynamic studies of oligonucleotides containing the two isomers of thymine glycol. *Chemistry*, **7**, 4343–4351.
- Hatahet, Z., Purmal, A.A. and Wallace, S.S. (1994) Oxidative DNA lesions as blocks to in vitro transcription by phage T7 RNA polymerase. *Ann. N. Y. Acad. Sci.*, **726**, 346–348.
- 64. Kathe,S.D., Shen,G.P. and Wallace,S.S. (2004) Single-stranded breaks in DNA but not oxidative DNA base damages block transcriptional elongation by RNA polymerase II in HeLa cell nuclear extracts. *J. Biol. Chem.*, 279, 18511–18520.
- Charlet-Berguerand, N., Feuerhahn, S., Kong, S.E., Ziserman, H., Conaway, J.W., Conaway, R. and Egly, J.M. (2006) RNA polymerase II bypass of oxidative DNA damage is regulated by transcription elongation factors. *EMBO J.*, **25**, 5481–5491.
 Kuraoka, I., Suzuki, K., Ito, S., Hayashida, M., Kwei, J.S., Ikegami, T.,
- 66. Kuraoka, I., Suzuki, K., Ito, S., Hayashida, M., Kwei, J.S., Ikegami, T., Handa, H., Nakabeppu, Y. and Tanaka, K. (2007) RNA polymerase II bypasses 8-oxoguanine in the presence of transcription elongation factor TFIIS. *DNA Repair (Amst.)*, 6, 841–851.
- Tornaletti,S., Maeda,L.S., Lloyd,D.R., Reines,D. and Hanawalt,P.C. (2001) Effect of thymine glycol on transcription elongation by T7 RNA polymerase and mammalian RNA polymerase II. *J. Biol. Chem.*, 276, 45367–45371.
- Kumar, N., Theil, A.F., Roginskaya, V., Ali, Y., Calderon, M., Watkins, S.C., Barnes, R.P., Opresko, P.L., Pines, A., Lans, H. *et al.* (2022) Global and transcription-coupled repair of 8-oxoG is initiated by nucleotide excision repair proteins. *Nat. Commun.*, **13**, 974.

- 69. 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.
- 70. Laine, J.P. and Egly, J.M. (2006) Initiation of DNA repair mediated by a stalled RNA polymerase IIO. *EMBO J.*, **25**, 387–397.
- D'Errico, M., Parlanti, E., Teson, M., de Jesus, B.M., Degan, P., Calcagnile, A., Jaruga, P., Bjoras, M., Crescenzi, M., Pedrini, A.M. *et al.* (2006) New functions of XPC in the protection of human skin cells from oxidative damage. *EMBO J.*, 25, 4305–4315.
- 72. Krokidis, M.G., D'Errico, M., Pascucci, B., Parlanti, E., Masi, A., Ferreri, C. and Chatgilialoglu, C. (2020) Oxygen-dependent

accumulation of purine DNA lesions in Cockayne Syndrome cells. *Cells*, **9**, 1671.

- Marietta, C. and Brooks, P.J. (2007) Transcriptional bypass of bulky DNA lesions causes new mutant RNA transcripts in human cells. *EMBO Rep.*, 8, 388–393.
- Lagerwerf,S., Vrouwe,M.G., Overmeer,R.M., Fousteri,M.I. and Mullenders,L.H. (2011) DNA damage response and transcription. *DNA Repair (Amst.)*, 10, 743–750.