13706–13713 *Nucleic Acids Research, 2014, Vol. 42, No. 22* *Published online 17 November 2014*
doi: 10.1093/nar/gku1183

Transcriptional bypass of regioisomeric ethylated thymidine lesions by T7 RNA polymerase and human RNA polymerase II

Changjun You[1], Pengcheng Wang[2], Xiaoxia Dai[1] and Yinsheng Wang[1,2,*]

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

Received September 22, 2014; Revised October 20, 2014; Accepted November 03, 2014

ABSTRACT

Alkylative damage to DNA can be induced by environmental chemicals, endogenous metabolites and some commonly prescribed chemotherapeutic agents. The regioisomeric *N*3-, O^2- and O^4-ethylthymidine (*N*3-, O^2- and O^4-EtdT, respectively) represent an important class of ethylated DNA lesions. Using nonreplicative double-stranded vectors containing an *N*3-EtdT, O^2-EtdT or O^4-EtdT at a defined site in the template strand, herein we examined the effects of these lesions on DNA transcription mediated by single-subunit T7 RNA polymerase or multisubunit human RNA polymerase II *in vitro* and in human cells. We found that O^4-EtdT is highly mutagenic and exclusively induces the misincorporation of guanine opposite the lesion, whereas *N*3-EtdT and O^2-EtdT display promiscuous miscoding properties during transcription. In addition, *N*3-EtdT and O^2-EtdT were found to inhibit strongly DNA transcription *in vitro* and in certain human cells. Moreover, *N*3-EtdT, but not O^2-EtdT or O^4-EtdT, is an efficient substrate for transcription-coupled nucleotide excision repair. These findings provide new important insights into how these alkylated DNA lesions compromise the flow of genetic information, which may help to understand the risk of these lesions in living cells.

INTRODUCTION

Environmental and endogenous genotoxic factors can result in a variety of alkylated DNA lesions, which represent a major class of DNA lesions (1). These agents include exogenous compounds like tobacco-specific nitrosamines, *N*-ethyl-*N*-nitrosourea and polycyclic aromatic hydrocarbons, as well as endogenous biochemical agents such as *S*-adenosyl-L-methionine, nitrosated peptides and polyamines (1–6). Hence, alkylative damage to DNA is an unavoidable consequence of endogenous metabolism and environmental exposure (1). Major alkylation sites in DNA include phosphate groups, ring nitrogens and exocyclic oxygen atoms of DNA bases, such as *N*3 of adenine, O^6 and *N*7 of guanine, and *N*3, O^2 and O^4 of thymine (7–13).

Many alkylated DNA lesions have been readily detected in various human tissues and urine samples (8–10,13,14). In particular, the levels of certain ethylated DNA adducts, including *N*3-, O^2- and O^4-ethylthymidine (*N*3-EtdT, O^2-EtdT and O^4-EtdT) (Figure 1a), were shown to be significantly higher in leukocyte and saliva DNA of smokers than nonsmokers (9,14). In this vein, it was reported that the levels of *N*3-EtdT, O^2-EtdT and O^4-EtdT in smokers' leukocyte DNA were ∼41.1, ∼44.8 and ∼48.3 lesions per 10^8 nucleosides, respectively, whereas those in nonsmokers were ∼4.1, ∼0.2 and ∼1.0 lesions per 10^8 nucleosides, respectively (9). In addition, *N*3-EtdT, O^2-EtdT and O^4-EtdT were found to be present in smokers' saliva DNA at frequencies of ∼4.5, ∼5.3 and ∼4.2 lesions per 10^8 nucleosides, respectively, but none of these lesions were detectable in saliva DNA of nonsmokers (14). Therefore, these three EtdT lesions have been proposed as potential useful biomarkers for exposure to ethylating agents and for cancer risk assessment (9,14).

Understanding how DNA lesions perturb the flow of genetic information during DNA replication and transcription may help to understand the risk of these lesions in living cells (15–18). In this context, it has been shown that O^4-EtdT is highly mutagenic during DNA replication and it is associated with cancer development in animal studies (11,19–22). In addition, both *N*3-EtdT and O^2-EtdT are able to strongly block DNA replication and substantially induce mutations *in vitro* and in *Escherichia coli* cells (20,21,23,24). To date, no studies have been carried out to investigate the effects of the three regioisomeric EtdT lesions on DNA transcription. Herein, we constructed nonreplicative double-stranded vectors containing an *N*3-EtdT, O^2-EtdT or O^4-EtdT at a defined site and assessed how

*To whom correspondence should be addressed. Tel: +1 951 827 2700; Fax: +1 951 827 4713; Email: Yinsheng.Wang@ucr.edu

© The Author(s) 2014. Published by Oxford University Press on behalf of Nucleic Acids Research.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

Figure 1. Experimental outline. (**a**) Chemical structures of *N*3-EtdT, O^2-EtdT and O^4-EtdT. (**b**) Schematic diagrams showing the procedures for the construction of the plasmids harboring a site-specifically incorporated *N*3-EtdT, O^2-EtdT or O^4-EtdT. (**c**) Strategy for assessing the impact of the EtdT lesions on DNA transcription. 'X' indicates *N*3-EtdT, O^2-EtdT or O^4-EtdT, which was located on the transcribed strand of *TurboGFP* gene downstream of the CMV and T7 promoters. The +1 transcription start sites are indicated by arrowheads. Lesion-bearing or lesion-free control plasmids were mixed with the competitor vector as DNA templates for *in vitro* or *in vivo* transcription assay. Although truncated RNA may be produced when transcription arrests at or near a lesion site, only run-off RNA is shown and used for RT-PCR. Among the RT-PCR products, only the wild-type sequence arising from the lesion-containing vector is shown. The arrows indicate the cleavage sites of Nt.BstNBI, NcoI and SfaNI. The last two enzymes are used to digest the RT-PCR products for subsequent PAGE and LC-MS/MS analyses.

these EtdT lesions compromise the efficiency and fidelity of DNA transcription *in vitro* and in human cells.

MATERIALS AND METHODS

Materials

Unmodified oligodeoxyribonucleotides (ODNs), [γ-^{32}P]ATP, enzymes and chemicals unless otherwise specified were purchased from Integrated DNA Technologies, Perkin-Elmer, New England BioLabs and Sigma-Aldrich, respectively. ON-TARGETplus SMARTpool siRNA against human CSB (L-004888) or XPC (L-016040), and Non-Targeting control siRNA (D-001210) were from Thermo Scientific Dharmacon. The 293T human embryonic kidney epithelial cells were obtained from American Type Culture Collection (ATCC). XPA-complemented (GM15876A) and XPA-deficient (XP12RO) human fibroblast cell lines were kindly provided by Professor Karlene A Cimprich (25). Cells were cultured in Dulbecco's Modified Eagle's medium supplemented with 10% fetal bovine serum (Invitrogen), 100 μg/ml streptomycin (ATCC) and 100 U/ml penicillin, and incubated at 37°C in 5% CO_2 atmosphere.

Transcription template preparation

We prepared the DNA templates for transcription assays as described elsewhere (26–28). To construct the unmodified control vector, a 50-mer ODN with the sequence of 5′-CTAGCGGAT-GCATCGACTCAATTATAGCAC GCCATGGTCGACTCATCGCG-3′ was annealed with its complementary strand and ligated to an EcoRI-NheI restriction fragment from the pTGFP-T7-Hha10 plasmid (27). Using the similar method, we also constructed a competitor vector that contains three more nucleotides than the control plasmid near the relevant site of interest (Figure 1c). We next employed Nt.BstNBI to nick the unmodified control vector and generated a gapped plasmid by removing a 25-mer single-stranded ODN (Figure 1b), followed by filling the gap with a 13-mer lesion-free ODN (5′-AATTGAGTCGATG-3′) and a 12-mer lesion-containing ODN (5′-ATGGCGXGCTAT-3′, X = *N*3-EtdT, *O*^{2}-EtdT or *O*^{4}-EtdT) (20). We subsequently incubated the ligation products with ethidium bromide and purified the supercoiled lesion-bearing plasmids by agarose gel electrophoresis, as described previously (26,28). Purified lesion-bearing or lesion-free control plasmids were premixed with the competitor vector at specific molar ratios and used as transcription templates, where the molar ratios of dT, *N*3-, *O*^{2}- and *O*^{4}-EtdT-bearing plasmids over the competitor vector were 3:1, 12:1, 12:1 and 3:1, respectively.

In vitro transcription assay

Transcription assay using T7 RNA polymerase (T7 RNAP) or human RNA polymerase II (hRNAPII) was performed as described previously (27,29). Briefly, T7 RNAP-mediated reaction contained 50 ng of NotI-linearized DNA templates, 10 U of RNase inhibitor, 20 U of T7 RNAP, and 0.5 mM each of ATP, CTP, GTP and UTP in a 20-μl mixture and was incubated at 37°C for 1 h. The hRNAPII-mediated reaction contained 50 ng of NotI-linearized DNA templates, 10 U of RNase inhibitor, 8 U of HeLa nuclear extract and 0.4 mM each of the four ribonucleotides in a 25-μl mixture and incubated at 30°C for 1 h.

In vivo transcription assay

XPA-deficient (XP12RO) and XPA-complemented (GM15876A) cells in a 24-well plate at ∼70% confluence were transfected with 50 ng DNA templates and 450 ng carrier plasmid (self-ligated pGEM-T; Promega) using Lipofectamine 2000 (Invitrogen), following the manufacturer's protocol. For siRNA experiments, 293T cells in a 24-well plate at 40–60% confluence were transfected with ∼25 pmol siRNAs for each gene. After a 48-h incubation, 50 ng DNA templates, 450 ng carrier plasmid and another aliquot of siRNAs were co-transfected into the cells using Lipofectamine 2000 (Invitrogen). All the cells were harvested for RNA extraction 24 h after transfection with the DNA templates.

RNA extraction and reverse transcription-polymerase chain reaction

The RNA products were extracted using Total RNA Kit I (Omega), and were treated twice with the DNA-free kit (Ambion) to eliminate DNA contamination. cDNA synthesis was performed with M-MLV reverse transcriptase (Promega) and a mixture of oligo(dT)$_{16}$ and a gene-specific primer (5′-TCGGTGTTGCTGTGAT-3′). Reverse transcription-polymerase chain reaction (RT-PCR) amplification was then performed by using a pair of primers spanning the lesion site and Phusion high-fidelity DNA polymerase as described previously (27). For evaluating the efficiency of siRNA knockdown, real-time RT-PCR was performed with the iQ SYBR Green Supermix kit (Bio-Rad) and gene-specific primers for CSB, XPC or the control gene GAPDH as described elsewhere (27).

Polyacrylamide gel electrophoresis analysis

We performed NcoI- and SfaNI-mediated restriction digestion/postlabeling assay to resolve 13-mer ^{32}P-labeled nonmutagenic fragment d(CATGGCGAGCTAT) from the corresponding products carrying an A→T or A→C mutation, i.e. d(CATGGCGTGCTAT) and d(CATGGCGCGCTAT), respectively (Figures 1c, 2a and b). To this end, a portion of the above RT-PCR products was incubated in a 10-μl mixture containing 5 U NcoI, 1 U shrimp alkaline phosphatase, and 1× NEB buffer 3 at 37°C for 1 h and subsequently at 70°C for 20 min. The resulting dephosphorylated DNA was then incubated in a 15-μl solution containing 5 U T4 polynucleotide kinase, 1× NEB buffer 3 and ATP (50 pmol cold, premixed with 1.66 pmol [γ-^{32}P]ATP). The mixture was incubated at 37°C for 30 min and then at 70°C for 20 min, after which 2 U SfaNI was added and incubated at 37°C for 2 h. The resulting ^{32}P-labeled restriction fragments were resolved by using 30% native polyacrylamide gel (acrylamide:bis-acrylamide = 19:1) and quantified by phosphorimager analysis (27,28). Using a similar restriction digestion/postlabeling assay with MluCI and Cac8I, we were able to completely resolve

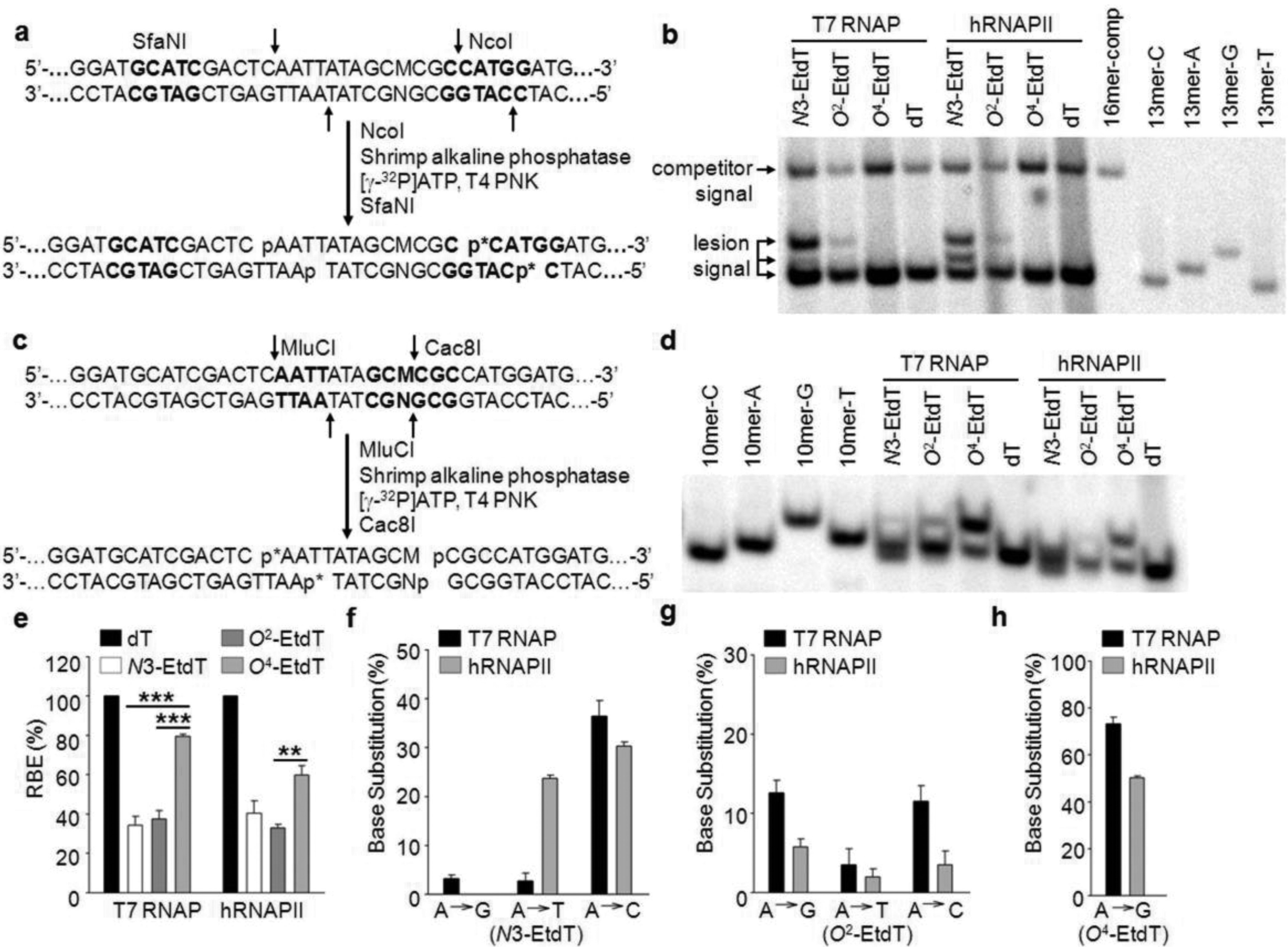

Figure 2. The effects of *N*3-EtdT, O^2-EtdT and O^4-EtdT on DNA transcription *in vitro*. (**a**) The effects of EtdT lesions on DNA transcription *in vitro*. (**a**) Sample processing for NcoI- and SfaNI-mediated restriction digestion/postlabeling assay (p* indicates ^{32}P-labeled phosphate group). (**b**) Representative gel images showing the NcoI- and SfaNI-treated restriction fragments of interest. '16mer-Comp' represents the standard ODN d(CATGGCGATATGCTAT), which corresponds to the restriction fragment arising from the competitor vector; '13mer-C', '13mer-A', '13mer-G' and '13mer-T' represent the standard ODN d(CATGGCGNGCTAT), where 'N' is C, A, G, T, respectively. (**c**) Sample processing for MluCI- and Cac8I-mediated restriction digestion/postlabeling assay. (**d**) Representative gel images showing the MluCI- and Cac8I-treated restriction fragments of interest. '10mer-C', '10mer-A', '10mer-G' and '10mer-T' represent the standard ODN d(AATTATAGCM), where 'M' is C, A, G, T, respectively. (**e**) The RBE values of *N*3-, O^2- and O^4-EtdT in *in vitro* transcription systems using T7 RNAP and HeLa nuclear extract (hRNAPII). (**f**–**h**) Mutagenic properties of *N*3- (f), O^2- (g) and O^4-EtdT (h) in *in vitro* transcription systems. The data represent the mean and standard error of results from three independent experiments. '**', $P < 0.01$; '***', $P < 0.001$. The *P*-values were calculated by using unpaired two-tailed Student's *t*-test.

the 10-mer ^{32}P-labeled fragment d(AATTATAGCG), which corresponds to the product arising from A→G mutation opposite the lesion site, from the nonmutagenic fragment d(AATTATAGCA) and the fragments with an A→T or A→C mutation (Figure 2c and d). The relative bypass efficiency (RBE) was calculated using the following formula, RBE = (lesion signal/competitor signal)/(unmodified control signal/competitor signal) (27,30).

Liquid chromatography-tandem mass spectrometry analysis

The RT-PCR products were treated with 20 U SfaNI and 20 U shrimp alkaline phosphatase in 120 μl NEB buffer 3 at 37°C for 2 h, and then at 70°C for 20 min. To the mixture, 50 U NcoI was added, and the mixture was incubated at 37°C for 2 h. The resulting solution was extracted with phenol/chloroform/isoamyl alcohol (25:24:1, v/v) and the aqueous portion was dried with Speed-vac, desalted with high performance liquid chromatography and dissolved in water. The resultant ODN mixture was subjected to liquid chromatography-tandem mass spectrometry (LC-MS/MS) analysis following previously described conditions (27,28). Briefly, a 0.50 × 150 mm Zorbax SB-C18 column (Agilent Technologies) was used. The flow rate was 8.0 μl/min, and a 5-min linear gradient of 5–20% methanol followed by a 25 min of 20–60% methanol in 400 mM 1,1,1,3,3,3-hexafluoro-2-propanol buffer (pH was adjusted to 7.0 by the addition of triethylamine) was employed for the separation. The LTQ linear ion trap mass spectrometer (Thermo Fisher Scientific) was set up for monitoring the fragmentation of the $[M-3H]^{3-}$ ions of the complementary 13-mer ODNs, i.e.

d(AATTATAGCMCGC), where 'M' designates A, T, C or G.

RESULTS

Using our recently developed competitive transcription and adduct bypass assay (27), we investigated how the regioisomeric *N*3-, O^2- and O^4-EtdT lesions perturb DNA transcription *in vitro* and in human cells. To this end, we employed a well-established protocol (26,28) to incorporate *N*3-, O^2- and O^4-EtdT at a defined site in the transcribed strand of a nonreplicative double-stranded plasmid, placing the lesions downstream from the cytomegalovirus (CMV) and T7 promoters as described in 'Materials and Methods' (Figure 1b and c). The lesion-containing or unmodified control plasmid was premixed with a competitor construct at given molar ratios and used as DNA templates for the *in vitro* or *in vivo* transcription assays. We subsequently amplified the run-off transcripts of interest with RT-PCR, digested the resulting RT-PCR products with suitable restriction enzymes and subjected the restriction digestion mixture to polyacrylamide gel electrophoresis (PAGE) and LC-MS/MS analyses (Figure 1c).

We performed the *in vitro* transcription assays using human RNA polymerase II (hRNAPII) in HeLa cell nuclear extract or purified T7 RNA polymerase (T7 RNAP). The latter is a single-subunit enzyme that is structurally homologous to eukaryotic mitochondrial RNA polymerases (31). The RT-PCR products were digested with NcoI and SfaNI, and the digestion products were subjected to PAGE analysis. In this context, the ^{32}P-labeled wild-type fragment d(p*CATGGCGTGCTAT) could be resolved from the corresponding products harboring an A→T or A→C mutation opposite the lesion site, i.e. d(p*CATGGCGAGCTAT) and d(p*CATGGCGGGCTAT), but could not be resolved from the corresponding product carrying an A→G mutation, i.e. d(p*CATGGCGCGCTAT) (Figure 2a and b). However, by digesting the RT-PCR products with MluCI and Cac8I, we were able to resolve the DNA fragment emanating from A→G mutation opposite the lesion site, i.e. d(p*AATTATAGCG), from the wild-type fragment and the products with an A→T or A→C mutation (Figure 2c and d).

The quantification data from PAGE analysis showed that O^4-EtdT did not considerably inhibit T7 RNAP-mediated transcription elongation, whereas *N*3-EtdT and O^2-EtdT impeded substantially DNA transcription mediated by T7 RNAP, with RBE values being ∼34 and ∼38%, respectively (Figure 2e). In addition, we found that the RBE value for O^4-EtdT was ∼60% during hRNAPII-mediated transcription reaction, whereas *N*3-EtdT and O^2-EtdT conferred strong inhibitory effects on the *in vitro* transcription mediated by hRNAPII (Figure 2e).

PAGE analysis also allowed us to determine the effects of *N*3-, O^2- and O^4-EtdT on the fidelity of transcription mediated by T7 RNAP or hRNAPII *in vitro*. The quantification data revealed that O^4-EtdT induced predominantly one type of mutant transcript that contains a guanine misincorporation opposite the lesion (A→G) during transcription by T7 RNAP and hRNAPII, with base substitution frequencies being ∼73 and ∼50%, respectively (Figure 2f). Unlike O^4-EtdT, *N*3-EtdT and O^2-EtdT displayed promiscuous miscoding properties during T7 RNAP-mediated transcription: A→G, A→T and A→C mutations were observed for *N*3-EtdT at frequencies of ∼3, ∼3 and ∼36%, respectively, and for O^2-EtdT at frequencies of ∼13, ∼4 and ∼12%, respectively (Figure 2f–h). During hRNAPII-mediated transcription, *N*3-EtdT induced A→T and A→C mutations at frequencies of ∼24 and ∼30%, respectively, whereas O^2-EtdT induced A→G, A→T and A→C mutations at frequencies of ∼6, ∼2 and ∼4%, respectively (Figure 2f and g). We also confirmed the identities of the above mutant products by LC-MS/MS analysis (Supplementary Figures S1 and S2).

We next asked how *N*3-, O^2- and O^4-EtdT compromise DNA transcription in human cells. For this purpose, we premixed either lesion-bearing or unmodified control plasmids with the competitor vector and co-transfected them into XPA-deficient (XP12RO) and XPA-complemented (GM15876A) cells. XPA is a core component of nucleotide excision repair (NER) that is one of the most versatile DNA damage removal pathways (32). After a 24-h incubation, RNA products were extracted from the cells and amplified with RT-PCR, followed by PAGE and LC-MS/MS analyses of the restriction digestion mixture of RT-PCR products as described above.

Our results showed that *N*3-EtdT, O^2-EtdT and O^4-EtdT considerably inhibited DNA transcription in GM15876A cells, with the RBE values being ∼37, ∼42 and ∼43%, respectively (Figure 3a and Supplementary Figure S3a). Moreover, we found that the RBE values for *N*3-EtdT, O^2-EtdT and O^4-EtdT were significantly lower in XPA-deficient (XP12RO) cells (∼28, ∼26 and 32%, respectively, Figure 3a and Supplementary Figure S3a). We also found that transcriptional bypass of *N*3-EtdT could induce A→T and A→C mutations at significantly higher frequencies (∼34 and ∼18%, respectively) in XPA-deficient cells than in XPA-complemented cells (∼20 and ∼11%, respectively, Figure 3b and Supplementary Figure S3a and b). On the other hand, transcriptional bypass of O^4-EtdT induced A→G mutation at a frequency of ∼34%, and O^2-EtdT induced three types of mutant transcripts (A→G, A→T and A→C) at frequencies of ∼2–5% in GM15876A cells; however, depletion of XPA did not considerably alter the mutagenic properties of O^2-EtdT or O^4-EtdT in human cells (Figure 3c and d and Supplementary Figure S3a and b).

We further examined the potential roles of XPC and CSB, which are key players in global-genome NER (GG-NER) and transcription-coupled NER (TC-NER), respectively (32), in transcriptional alternations induced by *N*3-EtdT, O^2-EtdT and O^4-EtdT in human cells. Our results showed that, compared to nontargeting control siRNA treatment, siRNA-mediated downregulation of CSB led to a significant decrease in transcriptional bypass efficiency of *N*3-EtdT, along with a marked elevation in transcriptional mutations induced by this lesion in 293T cells (Figure 3e and f and Supplementary Figures S4 and S5a). We also found that siRNA knockdown of CSB caused a slight, yet statistically significant increase in A→G mutation induced by O^4-EtdT; however, depletion of CSB did not considerably change the transcriptional bypass efficiency of O^4-EtdT in 293T cells

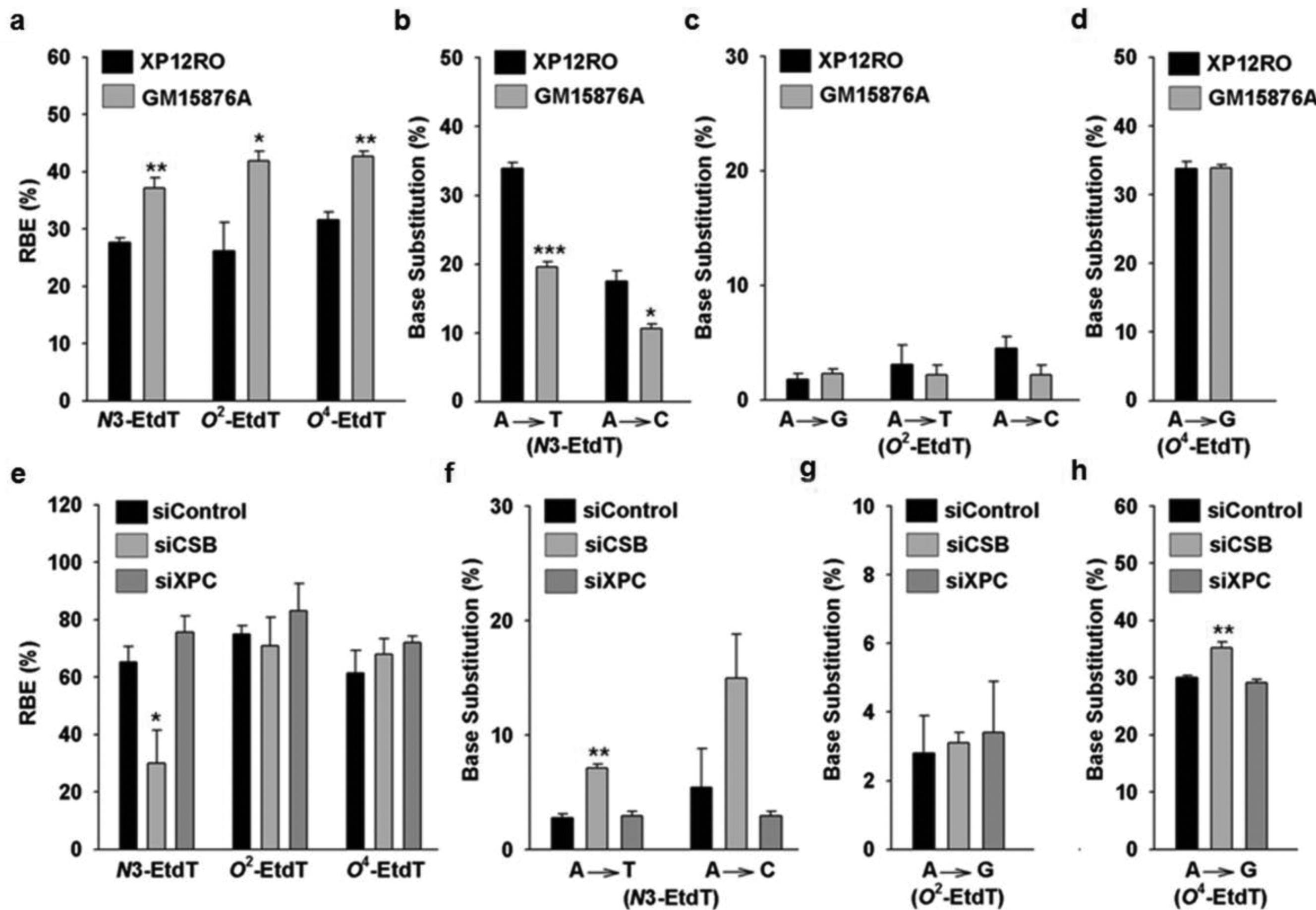

Figure 3. The effects of $N3$-EtdT, O^2-EtdT and O^4-EtdT on DNA transcription in human cells. (**a**) The RBE values of $N3$-, O^2- and O^4-EtdT in XPA-deficient (XP12RO) and XPA-complemented (GM15876A) cells. (**b**–**d**) Mutagenic properties of $N3$- (b), O^2- (c) and O^4-EtdT (d) in XP12RO and GM15876A cells. (**e**) The RBE values of $N3$-, O^2- and O^4-EtdT in 293T cells treated with *CSB* or *XPC* siRNAs. (**f**–**h**) Mutagenic properties of $N3$- (f), O^2- (g) and O^4-EtdT (h) in 293T cells treated with *CSB* or *XPC* siRNAs. The data represent the mean and standard error of results from three independent experiments. '*', $P < 0.05$; '**', $P < 0.01$; '***', $P < 0.001$. The P-values were calculated by using unpaired two-tailed Student's t-test.

(Figure 3e and h and Supplementary Figure S5a and b). In addition, depletion of CSB exerted no significant effect on the transcriptional bypass efficiency or mutation frequency of O^2-EtdT in 293T cells (Figure 3e and g and Supplementary Figure S5a and b). On the other hand, siRNA knockdown of XPC did not significantly change the effects of all three EtdT lesions on the efficiency and fidelity of DNA transcription in 293T cells (Figure 3e–h and Supplementary Figures S4 and S5a and b).

DISCUSSION

Alkylative damage to DNA is generally unavoidable because of the abundant presence of alkylating agents in the environment and within cells (1). Thymine is known to be alkylated at the $N3$, O^2 and O^4 positions (8,9,11–13,24). In this context, the regioisomeric $N3$-EtdT, O^2-EtdT and O^4-EtdT lesions have been detected at significantly higher levels in smokers than in nonsmokers, suggesting their potential roles as useful biomarkers for exposure to ethylating agents and possibly for cancer risk assessment (9,14). Although little is known about the repair of $N3$-EtdT, O^2-EtdT and O^4-EtdT are known to be poorly repaired and thus accumulate as highly persistent DNA lesions in mammalian tissues (12,19,33).

Accurate transmission of genetic information is essential for normal physiological processes of a living organism (15,16,34). Emerging evidence indicates that, aside from errors in DNA replication, transient transcriptional mutagenesis may contribute to genomic instability and stable phenotypic changes, which may ultimately result in carcinogenesis and the development of other diseases (15,16,34–39). It has been suggested that transient errors in the mRNA encoding a transcription factor involved in a bistable switch can promote heritable change in cellular phenotype (36,37). In addition, a recent study showed that mutagenic transcriptional bypass of a site-specifically inserted 8-oxoguanine lesion in the transcribed strand of the *Ras* oncogene could lead to a constitutively active mutant Ras protein and activation of downstream oncogenic signaling (35). Whereas the effects of $N3$-EtdT, O^2-EtdT and O^4-EtdT on DNA replication have been well studied (20–24), it is desirable to determine how the three regioisomeric EtdT lesions compromise the efficiency and fidelity of DNA transcription.

In the present study, we showed that $N3$-EtdT and O^2-EtdT, but not O^4-EtdT, strongly block DNA transcription by single-subunit T7 RNA polymerase or multisubunit human RNA polymerase II *in vitro*. In addition, we found that transcriptional bypass of O^4-EtdT is highly mutagenic and exclusively induces the misincorporation of guanine opposite the lesion (i.e. A→G); $N3$-EtdT and O^2-EtdT are also miscoding lesions, but display promiscuous mutagenic properties during transcription. These findings are consistent with results from previous replication studies of the corresponding alkylated DNA adducts (20–24), suggesting the similar mechanisms underlying base misincorporation opposite these lesions during DNA replication and transcription. The distinct miscoding properties of three regioisomeric EtdT lesions may be attributed to their unique chemical properties. In this respect, it has been shown that alkylation of thymine at the O^4 position promotes its favorable pairing with guanine (40). The addition of an alkyl group to the $N3$ position, however, can block the Watson–Crick hydrogen bonding face of thymine and thus abolish its base-pairing capability; alkylation of thymine at the O^2 position confers the inability of the nucleobase to pair preferentially with any of the canonical nucleobases (20,41). Thus, relative to O^4-EtdT, nucleotide incorporation opposite $N3$-EtdT and O^2-EtdT is much less selective.

We further demonstrated that all three regioisomeric EtdT lesions exhibit strong inhibitory and mutagenic effects on DNA transcription in certain human cells. Moreover, we found that XPA and CSB, but not XPC, are involved in the repair of $N3$-EtdT when located on the template DNA strand. Thus, these results suggest that the nonbulky $N3$-EtdT lesion is an efficient substrate for TC-NER, which is a subpathway of NER that preferentially repairs DNA lesions in the transcribed strand of active genes (32). Along this line, TC-NER has also been shown to efficiently repair several other nonbulky lesions including abasic sites in the transcribed strand in cells (27,29,34,42). On the other hand, although we found that depletion of XPA reduces the transcriptional bypass efficiencies of O^2-EtdT and O^4-EtdT, such depletion does not appreciably affect the mutagenic properties of the two lesions in human cells. In addition, siRNA-mediated knockdown of CSB has no or little impact on the transcriptional alternations induced by O^2-EtdT and O^4-EtdT in human cells. Thus, it is very likely that TC-NER has negligible roles in the removal of O^2-EtdT and O^4-EtdT from the template strand of actively transcribed genes in human cells.

In conclusion, we have examined, for the first time, how the regioisomeric $N3$-EtdT, O^2-EtdT and O^4-EtdT lesions perturb the flow of genetic information during transcription mediated by single-subunit T7 RNA polymerase or multisubunit human RNA polymerase II *in vitro* and in human cells. We have also investigated the potential roles of NER proteins, including XPA, CSB and XPC, in the repair of these EtdT lesions in human cells. Our results demonstrated that TC-NER has an important role in the removal of $N3$-EtdT, but not O^2-EtdT or O^4-EtdT, in the DNA template strand in human cells. Together, our findings provide novel and important insights into the biological consequences of these ethylated DNA lesions in living cells.

SUPPLEMENTARY DATA

Supplementary data are available at NAR Online.

ACKNOWLEDGMENT

We thank Professor Timothy R. O'Connor for providing the initial pTGFP-Hha10 vector, and Professor Karlene A Cimprich for kindly providing XPA-deficient (XP12RO) and XPA-complemented (GM15876A) cell lines.

FUNDING

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

REFERENCES

1. 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.
2. Beranek,D.T. (1990) Distribution of methyl and ethyl adducts following alkylation with monofunctional alkylating agents. *Mutat. Res.*, **231**, 11–30.
3. Rydberg,B. and Lindahl,T. (1982) Nonenzymatic methylation of DNA by the intracellular methyl group donor *S*-adenosyl-L-methionine is a potentially mutagenic reaction. *EMBO J.*, **1**, 211–216.
4. Sedgwick,B. (1997) Nitrosated peptides and polyamines as endogenous mutagens in O^6-alkylguanine-DNA alkyltransferase deficient cells. *Carcinogenesis*, **18**, 1561–1567.
5. Hecht,S.S. (1999) DNA adduct formation from tobacco-specific N-nitrosamines. *Mutat. Res.*, **424**, 127–142.
6. Thompson,C.L., McCoy,Z., Lambert,J.M., Andries,M.J. and Lucier,G.W. (1989) Relationships among benzo[*a*]pyrene metabolism, benzo[*a*]pyrene-diol-epoxide:DNA adduct formation, and sister chromatid exchanges in human lymphocytes from smokers and nonsmokers. *Cancer Res.*, **49**, 6503–6511.
7. Chen,L., Wang,M., Villalta,P.W. and Hecht,S.S. (2007) Liquid chromatography-electrospray ionization tandem mass spectrometry analysis of 7-ethylguanine in human liver DNA. *Chem. Res. Toxicol.*, **20**, 1498–1502.
8. Anna,L., Kovacs,K., Gyorffy,E., Schoket,B. and Nair,J. (2011) Smoking-related O^4-ethylthymidine formation in human lung tissue and comparisons with bulky DNA adducts. *Mutagenesis*, **26**, 523–527.
9. Chen,H.J., Wang,Y.C. and Lin,W.P. (2012) Analysis of ethylated thymidine adducts in human leukocyte DNA by stable isotope dilution nanoflow liquid chromatography-nanospray ionization tandem mass spectrometry. *Anal Chem.*, **84**, 2521–2527.
10. Chao,M.R., Wang,C.J., Chang,L.W. and Hu,C.W. (2006) Quantitative determination of urinary *N*7-ethylguanine in smokers and non-smokers using an isotope dilution liquid chromatography/tandem mass spectrometry with on-line analyte enrichment. *Carcinogenesis*, **27**, 146–151.
11. Swenberg,J.A., Dyroff,M.C., Bedell,M.A., Popp,J.A., Huh,N., Kirstein,U. and Rajewsky,M.F. (1984) O^4-ethyldeoxythymidine, but not O^6-ethyldeoxyguanosine, accumulates in hepatocyte DNA of rats exposed continuously to diethylnitrosamine. *Proc. Natl Acad. Sci. U.S.A.*, **81**, 1692–1695.
12. Den Engelse,L., De Graaf,A., De Brij,R.J. and Menkveld,G.J. (1987) O^2- and O^4-ethylthymine and the ethylphosphotriester dTp(Et)dT are highly persistent DNA modifications in slowly dividing tissues of the ethylnitrosourea-treated rat. *Carcinogenesis*, **8**, 751–757.
13. Godschalk,R., Nair,J., van Schooten,F.J., Risch,A., Drings,P., Kayser,K., Dienemann,H. and Bartsch,H. (2002) Comparison of multiple DNA adduct types in tumor adjacent human lung tissue: effect of cigarette smoking. *Carcinogenesis*, **23**, 2081–2086.
14. Chen,H.J. and Lee,C.R. (2014) Detection and simultaneous quantification of three smoking-related ethylthymidine adducts in

human salivary DNA by liquid chromatography tandem mass spectrometry. *Toxicol. Lett.*, **224**, 101–107.
15. Bregeon,D. and Doetsch,P.W. (2011) Transcriptional mutagenesis: causes and involvement in tumour development. *Nat. Rev. Cancer*, **11**, 218–227.
16. Morreall,J.F., Petrova,L. and Doetsch,P.W. (2013) Transcriptional mutagenesis and its potential roles in the etiology of cancer and bacterial antibiotic resistance. *J. Cell Physiol.*, **228**, 2257–2261.
17. Sale,J.E., Lehmann,A.R. and Woodgate,R. (2012) Y-family DNA polymerases and their role in tolerance of cellular DNA damage. *Nat. Rev. Mol. Cell Biol.*, **13**, 141–152.
18. Lange,S.S., Takata,K. and Wood,R.D. (2011) DNA polymerases and cancer. *Nat. Rev. Cancer*, **11**, 96–110.
19. Scherer,E., Timmer,A.P. and Emmelot,P. (1980) Formation by diethylnitrosamine and persistence of O^4-ethylthymidine in rat liver DNA in vivo. *Cancer Lett.*, **10**, 1–6.
20. Andersen,N., Wang,P. and Wang,Y. (2013) Replication across regioisomeric ethylated thymidine lesions by purified DNA polymerases. *Chem. Res. Toxicol.*, **26**, 1730–1738.
21. Zhai,Q., Wang,P. and Wang,Y. (2014) Cytotoxic and mutagenic properties of regioisomeric O^2-, *N*3- and O^4-ethylthymidines in bacterial cells. *Carcinogenesis*, **35**, 2002–2006.
22. Klein,J.C., Bleeker,M.J., Lutgerink,J.T., van Dijk,W.J., Brugghe,H.F., van den Elst,H., van der Marel,G.A., van Boom,J.H., Westra,J.G., Berns,A.J. *et al.* (1990) Use of shuttle vectors to study the molecular processing of defined carcinogen-induced DNA damage: mutagenicity of single O^4-ethylthymine adducts in HeLa cells. *Nucleic Acids Res.*, **18**, 4131–4137.
23. Zhai,Q., Wang,P., Cai,Q. and Wang,Y. (2014) Syntheses and characterizations of the in vivo replicative bypass and mutagenic properties of the minor-groove O^2-alkylthymidine lesions. *Nucleic Acids Res.* **42**, 10529–10537.
24. Grevatt,P.C., Donahue,J.M. and Bhanot,O.S. (1991) The role of *N*3-ethyldeoxythymidine in mutagenesis and cytotoxicity by ethylating agents. *J. Biol. Chem.*, **266**, 1269–1275.
25. Bomgarden,R.D., Lupardus,P.J., Soni,D.V., Yee,M.C., Ford,J.M. and Cimprich,K.A. (2006) Opposing effects of the UV lesion repair protein XPA and UV bypass polymerase eta on ATR checkpoint signaling. *EMBO J.*, **25**, 2605–2614.
26. Baker,D.J., Wuenschell,G., Xia,L., Termini,J., Bates,S.E., Riggs,A.D. and O'Connor,T.R. (2007) Nucleotide excision repair eliminates unique DNA-protein cross-links from mammalian cells. *J. Biol. Chem.*, **282**, 22592–22604.
27. You,C., Dai,X., Yuan,B., 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.
28. Yuan,B., You,C., Andersen,N., Jiang,Y., Moriya,M., O'Connor,T.R. and Wang,Y. (2011) The roles of DNA polymerases kappa and iota in the error-free bypass of N^2-carboxyalkyl-2′-deoxyguanosine lesions in mammalian cells. *J. Biol. Chem.*, **286**, 17503–17511.
29. You,C., Dai,X., Yuan,B. and Wang,Y. (2012) Effects of 6-thioguanine and S^6-methylthioguanine on transcription in vitro and in human cells. *J. Biol. Chem.*, **287**, 40915–40923.
30. Delaney,J.C. and Essigmann,J.M. (2006) Assays for determining lesion bypass efficiency and mutagenicity of site-specific DNA lesions in vivo. *Methods Enzymol.*, **408**, 1–15.
31. Masters,B.S., Stohl,L.L. and Clayton,D.A. (1987) Yeast mitochondrial RNA polymerase is homologous to those encoded by bacteriophages T3 and T7. *Cell*, **51**, 89–99.
32. 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.
33. Thomale,J., Huh,N.H., Nehls,P., Eberle,G. and Rajewsky,M.F. (1990) Repair of O^6-ethylguanine in DNA protects rat 208F cells from tumorigenic conversion by N-ethyl-N-nitrosourea. *Proc. Natl Acad. Sci. U.S.A.*, **87**, 9883–9887.
34. Gaillard,H., Herrera-Moyano,E. and Aguilera,A. (2013) Transcription-associated genome instability. *Chem. Rev.*, **113**, 8638–8661.
35. Saxowsky,T.T., Meadows,K.L., Klungland,A. and Doetsch,P.W. (2008) 8-Oxoguanine-mediated transcriptional mutagenesis causes Ras activation in mammalian cells. *Proc. Natl Acad. Sci. U.S.A.*, **105**, 18877–18882.
36. Gordon,A.J., Halliday,J.A., Blankschien,M.D., Burns,P.A., Yatagai,F. and Herman,C. (2009) Transcriptional infidelity promotes heritable phenotypic change in a bistable gene network. *PLoS Biol.*, **7**, e44.
37. Gordon,A.J., Satory,D., Halliday,J.A. and Herman,C. (2013) Heritable change caused by transient transcription errors. *PLoS Genet.*, **9**, e1003595.
38. Strathern,J.N., Jin,D.J., Court,D.L. and Kashlev,M. (2012) Isolation and characterization of transcription fidelity mutants. *Biochim. Biophys. Acta*, **1819**, 694–699.
39. Bregeon,D., Peignon,P.A. and Sarasin,A. (2009) Transcriptional mutagenesis induced by 8-oxoguanine in mammalian cells. *PLoS Genet.*, **5**, e1000577.
40. Brennan,R.G., Pyzalska,D., Blonski,W.J., Hruska,F.E. and Sundaralingam,M. (1986) Crystal structure of the promutagen O^4-methylthymidine: importance of the anti conformation of the O^4 methoxy group and possible mispairing of O^4-methylthymidine with guanine. *Biochemistry*, **25**, 1181–1185.
41. Huff,A.C. and Topal,M.D. (1987) DNA damage at thymine *N*-3 abolishes base-pairing capacity during DNA synthesis. *J. Biol. Chem.*, **262**, 12843–12850.
42. 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.