Published online 14 October 2008 *Nucleic Acids Research, 2008, Vol. 36, No. 20* ***6459–6471***
doi:10.1093/nar/gkn657

Transcription elongation past O^6-methylguanine by human RNA polymerase II and bacteriophage T7 RNA polymerase

Alexandra Dimitri, John A. Burns, Suse Broyde and David A. Scicchitano*

Department of Biology, New York University, 1009 Silver Center, 100 Washington Square East, New York, NY 10003, USA

Received August 2, 2008; Revised September 12, 2008; Accepted September 21, 2008

ABSTRACT

O^6-Methylguanine (O^6-meG) is a major mutagenic, carcinogenic and cytotoxic DNA adduct produced by various endogenous and exogenous methylating agents. We report the results of transcription past a site-specifically modified O^6-meG DNA template by bacteriophage T7 RNA polymerase and human RNA polymerase II. These data show that O^6-meG partially blocks T7 RNA polymerase and human RNA polymerase II elongation. In both cases, the sequences of the truncated transcripts indicate that both polymerases stop precisely at the damaged site without nucleotide incorporation opposite the lesion, while extensive misincorporation of uracil is observed in the full-length RNA. For both polymerases, computer models suggest that bypass occurs only when O^6-meG adopts an *anti* conformation around its glycosidic bond, with the methyl group in the *proximal* orientation; in contrast, blockage requires the methyl group to adopt a *distal* conformation. Furthermore, the selection of cytosine and uracil partners opposite O^6-meG is rationalized with modeled hydrogen-bonding patterns that agree with experimentally observed O^6-meG:C and O^6-meG:U pairing schemes. Thus, *in vitro*, O^6-meG contributes substantially to transcriptional mutagenesis. In addition, the partial blockage of RNA polymerase II suggests that transcription-coupled DNA repair could play an auxiliary role in the clearance of this lesion.

Methylating agents are among the array of chemicals that damage the genome. They react with nucleophilic sites in DNA, adding methyl groups to the purines, pyrimidines and the phosphodiester backbone, resulting in modifications that compromise the fidelity and efficiency of replication and transcription. These agents include exogenous compounds like *N*-methyl-*N*-nitrosourea, methyl methane sulfonate and *N*-methyl-*N*′-nitro-nitrosoguanidine (1), along with endogenous biochemical agents like *S*-adenosylmethionine (2) and metabolically nitrosated amines and amides such as betaine and choline (3). Nitrates and nitrites found in food and beverages can also form methylating agents in cells (4). Hence, aberrant DNA methylation is an unavoidable consequence of metabolism and exposure to the surrounding environment (5,6). Approximately 12 different methylated DNA lesions have been identified, with the relative amount of each depending largely on the chemical nature of the reactive alkylating species. The most frequent lesions occur at the N7- or O^6-positions of guanine and the N3 position of adenine (6,7).

O^6-Methylguanine (O^6-meG) (Figure 1) is a stable DNA lesion formed by the reaction of methylating agents, including the chemotherapeutic agents procarbazine and temozolomide, with guanine (8,9). O^6-MeG is highly mutagenic since it often instructs DNA polymerases to incorporate thymine instead of cytosine, resulting in GC to AT transitions (10–12). The highly mutagenic nature of O^6-meG offers an explanation as to why methylating agents that produce it are carcinogenic (13). But O^6-meG can also block replication (14), inducing chromosomal aberrations, sister chromatid exchange and intrachromosomal homologous recombination (15–17).

O^6-MeG is repaired in a non-enzymatic, stoichiometric fashion by the direct reversal protein O^6-methylguanine–DNA methyltransferase (MGMT) (18). MGMT transfers the methyl group from the lesion to a cysteine residue within itself, rendering it incapable of continued repair; hence, the number of O^6-meG lesions that can be cleared from a cell's genome is limited by the available MGMT molecules, unless more of the protein is made. In fact, variations in MGMT levels among cell types result in a differential ability of organs and tissues to repair O^6-meG via direct reversal, leading to the accumulation of O^6-meG after MGMT is depleted (19,20).

*To whom correspondence should be addressed. Tel: +1 212 998 8229; Fax: +1 212 995 4015; Email: das2@nyu.edu

Replication past O^6-meG by several DNA polymerases has been well studied (7,10,14,21–24); however, little is known about its effect on transcription elongation. In *Escherichia coli*, RNA polymerase bypasses O^6-meG with high efficiency, often misincorporating uracil opposite the lesion. The resulting alterations to the transcripts have been termed transcriptional mutagenesis (TM) (25). Here, we examined transcriptional processing of O^6-meG by bacteriophage T7 RNA polymerase (T7 RNAP) and human RNA polymerase II (hRNAPII). These two polymerases differ significantly in the architecture of their active sites and behave differently when facing DNA lesions (26–28). It is shown here for the first time that O^6-meG poses a minimal hurdle to transcription by T7 RNAP, but it does impede hRNAPII transcription to a significant extent. Consistent with the previous study with *E. coli* RNA polymerase (25), bypass of O^6-meG causes significant sequence alterations in the RNA transcripts, with predominant misincorporation of uracil by both polymerases. Computer modeling was utilized to provide a structural explanation of the experimental results. These models showed that the altered Watson–Crick hydrogen-bonding edge of O^6-meG (Figure 1) underlies the observed selectivity for uracil as well as cytosine incorporation during transcription past the lesion, with hydrogen-bonding possibilities between O^6-meG and uracil or cytosine that have been inferred in NMR solution studies (29–32) and observed in X-ray crystal structures of O^6-meG (22,33–36). In addition, the more crowded active site region of hRNAPII explained the greater blocking effect of the lesion in this polymerase.

MATERIALS AND METHODS

Materials

Unless stated otherwise, all reagents, enzymes and biologicals were purchased from Promega, Sigma-Aldrich, New England Biolabs, Ambion or BioRad. DNA oligomers were obtained from Sigma-Aldrich. Radioactive materials were purchased from Perkin Elmer. All DNA sequencing was done using the GeneAmp PCR System

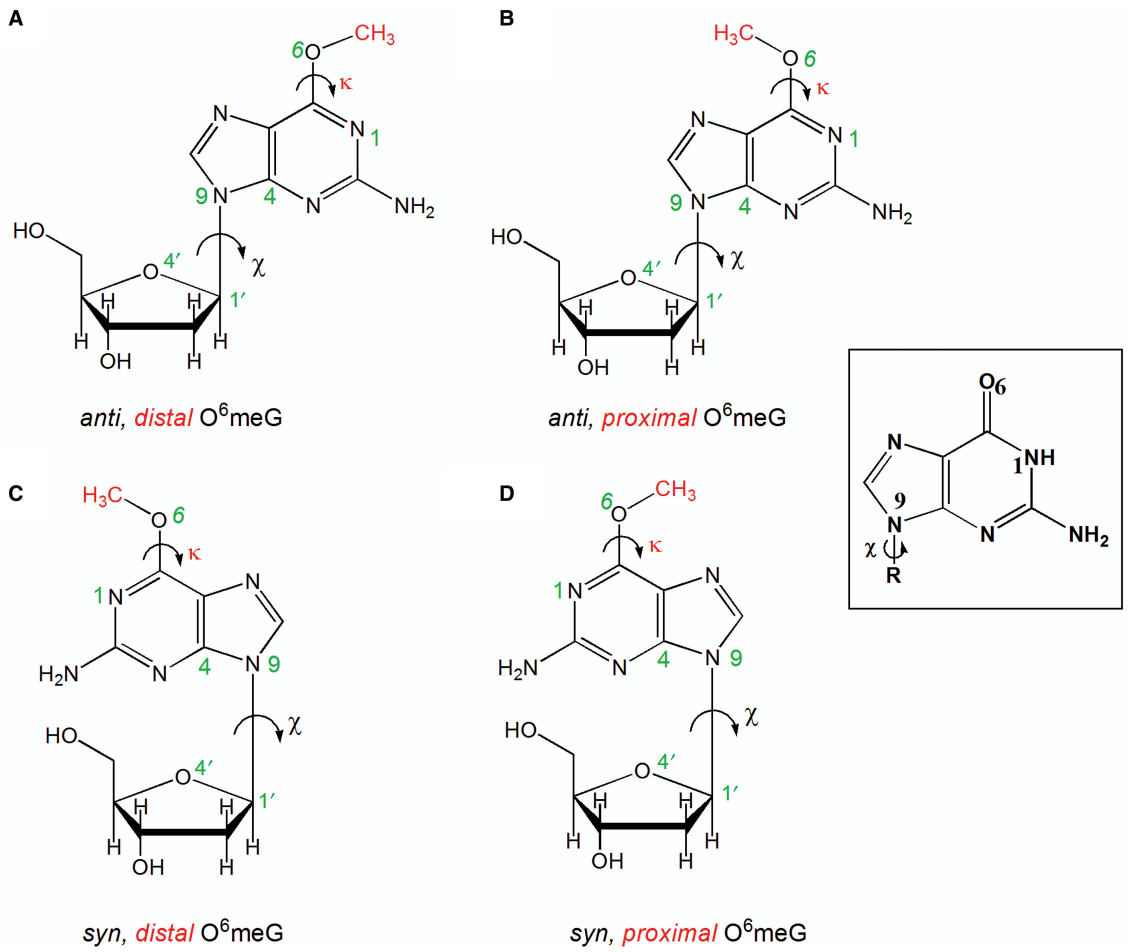

Figure 1. Structures of the four possible conformations of O^6-meG. The glycosidic torsion angle (χ) defined by the atoms $O4'$–$C1'$–N9–$C4$ is shown in black and the torsion angle (κ) defined by the atoms N1–C6–O^6–C is displayed in red. The angle χ is shown in both the *anti* (the six-membered ring points away from the sugar) or *syn* (the six-membered ring is over the sugar) conformations, while the torsion angle (κ) is in the *distal* (facing N1–C6) or the *proximal* (facing away from N1–C6) orientation. These torsional possibilities give rise to four conformational combinations for the lesion (**A**) *anti,distal* (**B**) *anti,proximal* (**C**) *syn,distal* and (**D**) *syn,proximal* O^6-meG. Guanine is shown as an inset to highlight the different hydrogen-bond donor/acceptor properties.

9700 (BioRad) and 3100 Capillary Sequencer (Applied Biosystems).

Transcription template construction

DNA templates suitable for transcription by T7 RNAP and hRNAPII were prepared as described in (26,27,37) (Figure 2). In brief, the restriction endonuclease BbsI was used to linearize the plasmid pCI-neo-G-less-T7, which contained the two promoters needed for both RNAPs (Supplementary Figure S1). A double-strand tail was assembled by annealing three synthetic single-strand DNA oligomers—an 11-mer and a 96-mer, to a complementary 90-mer containing a 5′-biotin tag. The 11-mer was either unmodified or modified with a site-specific O^6-meG lesion, depending on whether a control or adduct-modified template was being assembled, respectively. DNA oligomers are shown in Table 1. T4 DNA ligase was used to ligate the tail to the digested vector for 16 h at 16°C. The product was precipitated using streptavidin-linked beads, gel purified and tested for nicks as described elsewhere (26,27,37).

Transcription reactions with bacteriophage T7 RNA polymerase

Transcription using T7 RNAP was performed as recommended by the manufacturer Promega and as described elsewhere (26,27). Briefly, the transcription reactions contained 20 units of T7 RNAP (a unit is defined as the amount of enzyme required to catalyze the incorporation of 5 nmol of CTP into acid-insoluble product in 60 min at 37°C in a total volume of 100 µl), 25 ng (0.033 pmols) of DNA template, 2.5 mM each of ATP, GTP and UTP, 200 µM [α-^{32}P]CTP (50 Ci/mmol), 100 mM DTT and 10 units RNasin, 40 mM Tris–HCl (pH 7.9 at 25°C), 10 mM NaCl, 6 mM $MgCl_2$ and 2 mM spermidine in a final volume of 20 µl. The molar concentrations provided are the stock solution concentrations. The reaction was incubated at 37°C for 1 h, the unincorporated nucleotides were removed, and an equal volume of loading dye (80% formamide, 5 mM Tris–boric acid, 1 mM EDTA, 0.1% bromophenol blue and 0.1% xylene cyanol) was added to the product, which was then resolved using denaturing PAGE.

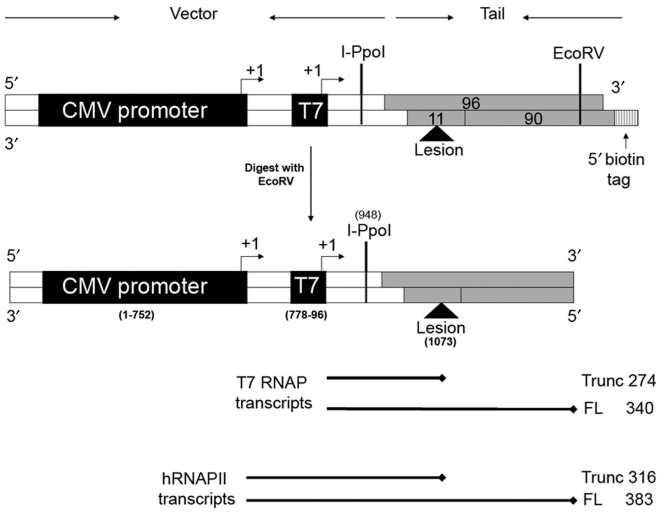

Figure 2. Site specifically modified DNA templates prepared for transcriptional assays. The vector displayed in *white* is comprised of the cloned T7 promoter and the CMV promoter/enhancer (*black boxes*). The tail region is formed by annealing a 96-mer, a 90-mer and an 11-mer (*grey boxes*). The position of the O^6-meG lesion is shown with a triangle. The size of the template is 1140 bp after digestion with EcoRV, the last step during the purification of the template which removes the biotin tag from the tail end. The arrowheads indicate the +1 transcription start site. The expected sizes of the RNA products obtained during transcription by T7 RNAP and hRNAPII are shown below the template.

To test the efficiency of lesion bypass, a transcription time course was performed. The reaction was initiated as described earlier, and a radioactively labeled 108-base DNA oligomer was added to the master mix to act as a standard during PAGE. Aliquots were removed from the reaction mix at time points 0, 1, 10, 20, 30, 40, 50 and 60 min. Each aliquot was quenched by adding loading dye and heating to 65°C for 15 min to ensure polymerase inactivation. ImageQuant software (Molecular Dynamics) was used to visualize and quantify the radiolabeled transcripts, and for each time point, the RNA band intensity was compared to the DNA standard intensity.

Transcription reactions with human RNA polymerase II

Multiple rounds of transcription reactions using hRNAPII were performed as recommended by the manufacturer Promega and as described in (26,27,37,38). HeLa nuclear extract served as the source of the hRNAPII machinery, and the presence of the CMV enhancer/promoter on the DNA templates, which is specific for promoting transcription by RNAPII, ensured that only hRNAPII complexes formed that could carry out RNA elongation past the lesion (Promega and references herein). Briefly, transcription was carried out using 8 units of HeLa nuclear extract (a unit is defined as the amount

Table 1. Sequence of DNA oligomers and primers used

Oligomer name	DNA Sequence
96-mer[a]	5′-TTGCGGTAGCGATGGGTGCTGTACTCAGGTGTGG AATCAACCCACAGCTGACAGGGCAGGTCTTGGCCAG TTGGGATATCCAAAACATCTTGTTGA-3′
90-mer[a]	5′-TTTTTTTTTTCAACAAGATGTTTTGGATATCCCAA CTGGCCAAGACCTGCCCTGTCAGCTGTGGGTTGATTCC ACACCTGAGTACAGCAC-3′
11-mer[a]	5′-CCATC**G̲**CTACC-3′
49L[b]	5′-GCTTCCTCATCCTTCTCCATCTCTCTGTGCAGCTG TGGGTTGATTCACG
24F[c]	5′-AACGCAGTCAGTGCTTCTGACACA-3′
25R[c]	5′-GAGAGATGGAGAAGGATGAGGAAGC -3′
22R trunc	5′-TACAGTGCTCGCATACGCTCAA-3′
24 chimera	5′-UUUUGAGCGTATGCGAGCACTGTAidt-3′

[a]The 96-mer, 90-mer and the 11-mer were annealed to make the tail of the DNA templates. The bold underscored G in the 11-mer represents either the undamaged guanine or the damaged O^6-meG.
[b]49L is the primer used in the reverse transcription of the full-length RNA obtained from transcription of T7 RNAP past the DNA templates.
[c]24F and 25R are the forward and reverse PCR primers.
22R: The reverse transcription primer used in the ligation mediated reverse transcription. It is complementary to the chimera, which was ligated to the RNA prior to reverse transcription.

of extract required for the incorporation of 50 fmol of nucleotides into a 363-nt runoff transcript 'generated from the CMV immediate early promoter fragment' per hour at 30 °C), 100 ng (0.15 pmols) DNA template, 10 units of RNase inhibitor, 3 mM $MgCl_2$, 10 mM each of ATP, GTP and UTP, 0.4 μM CTP, 1 μl [α-^{32}P]CTP (25 Ci/mmol), 20 mM HEPES (pH 7.9 at 25°C), 100 mM KCl, 0.2 mM EDTA, 0.5 mM DTT, 20% glycerol in a final volume of 25 μl (Molar concentrations provided are the stock solution concentrations). The reaction was incubated at 30°C for 60 min, and then quenched using a stop solution [0.3 M Tris–HCl (pH 7.4 at 25°C), 0.3 M sodium acetate, 0.5% SDS, 2 mM EDTA and 3 μg/ml tRNA] provided by the manufacturer (Promega). The resulting RNA was purified by phenol:chloroform extraction and resolved using denaturing 7% PAGE.

For a single round of transcription, GTP was omitted from the NTP mix, and the polymerase complex was allowed to transcribe up to the point at which the first GTP was required. After 10 min, heparin was added to bind all free hRNAPII. Finally, the addition of 100 μM GTP allowed transcription to continue, and after an additional 40 min, the reaction was quenched. The resulting RNA was purified by phenol:chloroform extraction and resolved using denaturing 7% PAGE.

To test the efficiency of lesion bypass, a transcription time course was performed. In order to prevent repair of O^6-meG, HeLa nuclear extracts were pre-incubated with 0.5 μM O^6-benzylguanine (BZG), a potent inhibitor of the MGMT that repairs O^6-meG (39). A loading standard 175 bases in length was added to the reaction mix, and the transcription reaction was initiated by the addition of the DNA template as described earlier. Aliquots were removed from the reaction mix, and quenched in stop solution at time points 0, 1, 5, 10, 20, 30, 40 and 60 min. The resulting RNA was purified and resolved on 7% PAGE. ImageQuant software (Molecular Dynamics) was used to visualize and quantify the radiolabeled transcripts. At each time point, RNA band intensity was used to determine the percentage of each RNA species present.

RNA sequencing

To sequence full-length transcripts, transcription reactions were performed without radioactive CTP but with all four non-radioactive nucleotides, each at a concentration of 500 μM for T7 RNAP and 400 μM for hRNAPII. Transcription reactions were followed by digestion with RNase-free DNase at 37°C for 30 min to destroy any residual DNA template. The resulting RNA was purified and sequenced using the reverse transcriptase polymerase chain reaction (RT–PCR). In brief, for the reverse transcriptase reaction, 20 pmol of primers were annealed to the RNA template, ImProm II (Promega) was used as a source of reverse transcriptase, and the reaction was performed following the manufacturer's protocol. Residual RNA was digested with RNase, and the RNase was subsequently heat inactivated. Excess primers and small RNA fragments were eliminated by using Microcon Centrifugal Filter Columns. Finally, PCR was performed with the forward and reverse primers 24F and 25R (Table 1).

To sequence the truncated transcripts, 1 μM of a chimeric RNA–DNA oligomer, 24 chimera (Table 1), was ligated to the 3′-ends of the transcripts using T4 RNA ligase and DMSO (40). The chimeric RNA–DNA oligomer contains four ribonucleotides at its 5′-end to facilitate ligation to the transcripts using RNA ligase; the other bases of the oligomer are deoxyribonucleotides that make the chimeric oligomer resistant to degradation. The reverse transcriptase primer 22RTrunc was annealed to the tail, and RT–PCR was subsequently performed. This approach selected all RNA, both the full-length and the truncated transcripts to which the chimeric oligomer was ligated.

For both full-length and truncated RNA, the RT–PCR products were cloned and sequenced. The RT–PCR products were purified using a 1.2% non-denaturing agarose matrix and subsequently inserted into the pDrive vector (Promega) specifically designed for cloning PCR products. The cloned PCR products were used to transform Qiagen EZ competent cells grown on agar plates containing kanamycin and IPTG. Only white colonies, indicating the insertion of the PCR product in the vector, were selected and grown overnight in Luria broth supplemented with ampicillin. Plasmids were extracted from the cells using the QIAprep Spin Miniprep Kit (Qiagen) and were then sequenced. For each polymerase, 100 clones were sequenced for full-length products and 50 clones were sequenced for the truncated products (26,27).

Molecular modeling of O^6-meG in T7 RNAP

To model the O^6-meG adduct in the active site of T7 RNAP, four available X-ray crystal structures of the protein were used. The crystal structures represent four phases of one cycle of nucleotide addition during the elongation phase of transcription: the pre-insertion substrate complex, the insertion substrate complex, the pre-translocation product complex and the post-translocation complex with PDB codes 1S0V (41), 1S76 (42), 1S77 (42) and 1MSW (43), respectively. These structures are shown in Supplementary Figure S2. In each phase, the protein was crystallized in a ternary complex with a DNA duplex of approximately 30 bases, a 9 base DNA:RNA hybrid, and an incoming ribonucleotide, except for the post-translocation phase. Sequences of the template/non-template DNA and the RNA transcript were re-modeled to match the sequences used in the study reported here (Figure 3).

The structure of the damaged O^6-meG base was obtained from a crystal structure of the lesion paired with a cytosine [PDB code 1D24 (33)]. The unmodified guanine at position $i+1$ or $i+2$ (only in the post-translocation complex) of the template strand was replaced by this O^6-meG structure (Figure 3). The subsequent search for optimal structures employed the following procedure. The glycosidic torsion angle of the lesion was retained in the original *anti* conformation present in the crystal (χ = 252°) (Figure 1A and B). The base of the crystal NTP was replaced by cytosine or uracil while retaining the original *anti* conformation of the glycosidic torsion angle (χ = 196°). The torsion angle (κ) (Figure 1) governing

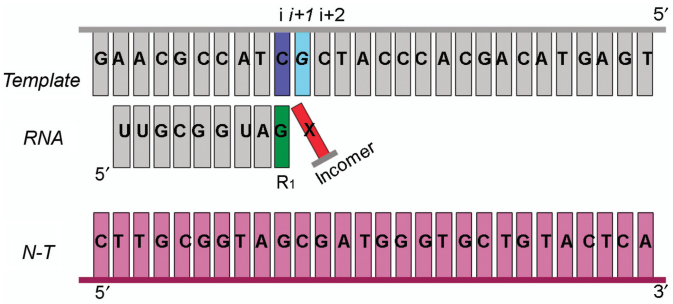

Figure 3. Sequences employed in the experimental and modeling studies in T7 RNAP and yRNAPII complexes. The insertion position $i + 1$ occupied by the template nucleotide to be transcribed, G or O^6-meG is shown in *cyan*. The pre-insertion position $i + 2$ occupied by the next to be transcribed nucleotide is in grey. The ribonucleotide previously incorporated at the 3′-end of the RNA is designated R_1 and shown in *green*. The rNTP to be inserted opposite the lesion is called Incomer and colored *red* and the non-template DNA strand is colored in *purple*.

the methyl group conformation of the adduct was then rotated utilizing the INSIGHTII (Accelrys) utility bump-check in a search for structures with minimal collisions. This torsion angle has been observed in crystal structures to adopt two conformational ranges termed *distal* and *proximal* with (κ) values in the range of 0–76° and 159–208°, respectively (22,33–36,44). However, to accommodate the incoming pyrimidines, only the *proximal* conformation for (κ), ranging from 182° to 191°, provided structures without clashes. The structure with κ = 183° was used for further study, as this value fell in the middle of the range of *proximal* κ-values (159–208°) previously observed in crystal structures of O^6-meG (22,34–36,44). Subsequently, the slight, close contacts remaining (<2 Å) were relieved by rotation of the side chain of Arg-632, namely around the C_α–C_β, C_β–C_γ and C_γ–C_δ bonds. Finally, additional adjustments were made to the torsion angle (χ) of the NTP (<15°) in order to locate structures with optimal hydrogen bonds between O^6-meG and the incoming NTP.

In the case of an incoming ATP, the original *anti* conformation of O^6-meG exhibited severe collisions with the ribonucleotide. To relieve the clashes observed in the model and explain the insertion of adenine in the growing RNA as identified by the biochemical results, the torsion angle (χ) of O^6-meG was continuously rotated to find the range of minimal collision, which was in the *syn* domain. As with the incoming pyrimidines, slight close contacts (<2 Å) were relieved by rotation of the side chain of the same Arg-632, namely around the C_β–C_γ, C_γ–C_δ and C_δ–N_ε bonds. When a collision-free range was obtained, the torsion angle (χ) of the lesion was further rotated slightly (<3°) to locate a conformation with optimal hydrogen-bonding possibilities between the incoming ATP and O^6-meG. No NTP was present opposite the base at $i+2$ in the post-translocation complex. The hydrogen-bonding criteria in the insertion complex were in the following observed ranges (22,45–50): (1) donor to acceptor atom distance of 2.5 Å < X–A< 3.9 Å; (2) hydrogen to acceptor atom distance of 1.5 Å < H–A < 2.9 Å; and (3) angle θ defined by the donor–hydrogen–acceptor atoms (X–H–A) >90°. Full details for the torsion angles and the resulting hydrogen bonds in all the modeled structures are provided in Tables S1 and S2, respectively.

Molecular modeling of O^6-meG in yeast RNA polymerase II

Crystal structures for the hRNAPII are not available at this writing. However, crystal structures for yeast *S. cerevisiae* RNAPII (yRNAPII) shares 53% sequence identity with the human polymerase, and the conserved residues are distributed over the entire structure (51). In addition, cryo-electron microscopy with image reconstruction and variance analysis of the hRNAPII together with homology modeling of the human enzyme based on the yRNAPII show that the human and the yeast enzymes' cores that contain the active sites are structurally indistinguishable (52). This makes the yeast polymerase a good model for the human enzyme. The X-ray crystal structure of *S. cerevisiae* RNAPII, PDB code 2NVZ (53), was in a ternary complex form with a 29-base DNA template strand, a 10-base RNA transcript and a 14-base non-template DNA strand complementary to the template strand downstream of the RNA. This structure was used as the starting model for yRNAPII containing the O^6-meG lesion in its active site. The modeling procedures employed for yRNAPII were very similar to the procedures described earlier for T7 RNAP. First, the undamaged templating base present in the crystal was replaced by the lesion while retaining its original glycosidic torsion angle (χ = 187°), placing the damaged guanine in an *anti* conformation. Then, the torsion angle κ was continuously rotated to find a collision-free range for the methyl group (κ = 200–250°). The original crystal contained a uracil as an incoming ribonucleotide. Therefore, to explain the incorporation of uracil opposite the lesion as shown in the experimental results, the possibility of hydrogen bonding between O^6-meG and uracil was evaluated. Hydrogen-bonding requirements followed the criteria specified earlier for the T7 RNAP models. Next, the incoming ribonucleotide was remodeled to a cytosine, to interpret its observed incorporation in the RNA; again the glycosidic torsion angle was retained as in the crystal (χ = 220°), and finally, hydrogen-bonding possibilities with O^6-meG were investigated. All modeling was carried out with the Biopolymer module of INSIGHTII (Accelrys), and all images were generated using the PYMOL program.

RESULTS

DNA templates suitable for transcription were synthesized and confirmed to be free of nicks

DNA templates were constructed with a site-specific O^6-meG on the transcribed strand. The lesion was positioned downstream from the CMV enhancer/promoter and the T7 RNAP promoter to allow for comparisons between the two RNAPs transcribing past O^6-meG (Figure 2). Unmodified templates acted as control DNA. To test for the presence of nicks in the ligated products, the constructed templates were digested with I-PpoI, an endonuclease that cuts with high specificity at a 15-bp recognition sequence. The fragments were radiolabeled

at the 5′-ends and analyzed on a denaturing PAGE. When ligation is complete, the resulting DNA fragments should be 187, 191 and 953 bases in length. In contrast, incomplete ligation that results in the presence of nicks in the template would result in fragments 73, 77, 110 or 118 bases in length. The absence of the shorter bands indicates that ligation was indeed efficient, and the templates for transcription were free of nicks (Supplementary Figure S3).

O^6-MeG minimally blocks T7 RNA polymerase elongation

Transcription of the unmodified DNA with T7 RNAP produced only full-length RNA 340 bases in length. In contrast, transcription of the modified DNA template primarily produced full-length transcripts 340 bases in length, with some truncated RNA of approximately 274 bases (Figure 4A). The percentage of full-length RNA averaged from nine separate reactions was $95 \pm 1\%$. The one percent value represents the standard deviation from the mean. The shorter transcripts coincided with the position of O^6-meG on the template strand, indicating that this lesion impedes elongating T7 RNAP. The shorter RNA molecules observed between 274 and 340 bases were also obtained with the unmodified templates; hence, inherent pause sites on the DNA likely cause their formation, not the damaged base.

Experiments were carried out to examine T7 RNAP lesion bypass efficiency over time. Aliquots were removed from the reaction mix at different times as described in the Materials and Methods section. For each time point, the RNA band intensity was compared to a DNA standard of 108 bases in length. The experiment was performed in triplicate, and in each case, the undamaged templates produced increasing amounts of full-length RNA over time (Figure 4B, lanes 1 through 8), while the O^6-meG-damaged templates produced mainly full-length transcripts (~95%), with a small amount of truncated RNA (~5%) (Figure 4B, lanes 9 through 16). The results suggest that rapid RNA synthesis is mildly obstructed by O^6-meG.

O^6-MeG partially blocks human RNA polymerase II transcription

HeLa nuclear extract was used as the source of hRNAPII machinery. The presence of the CMV enhancer/promoter on the DNA templates supports hRNAPII transcription initiation alone, as described in the manufacturer's

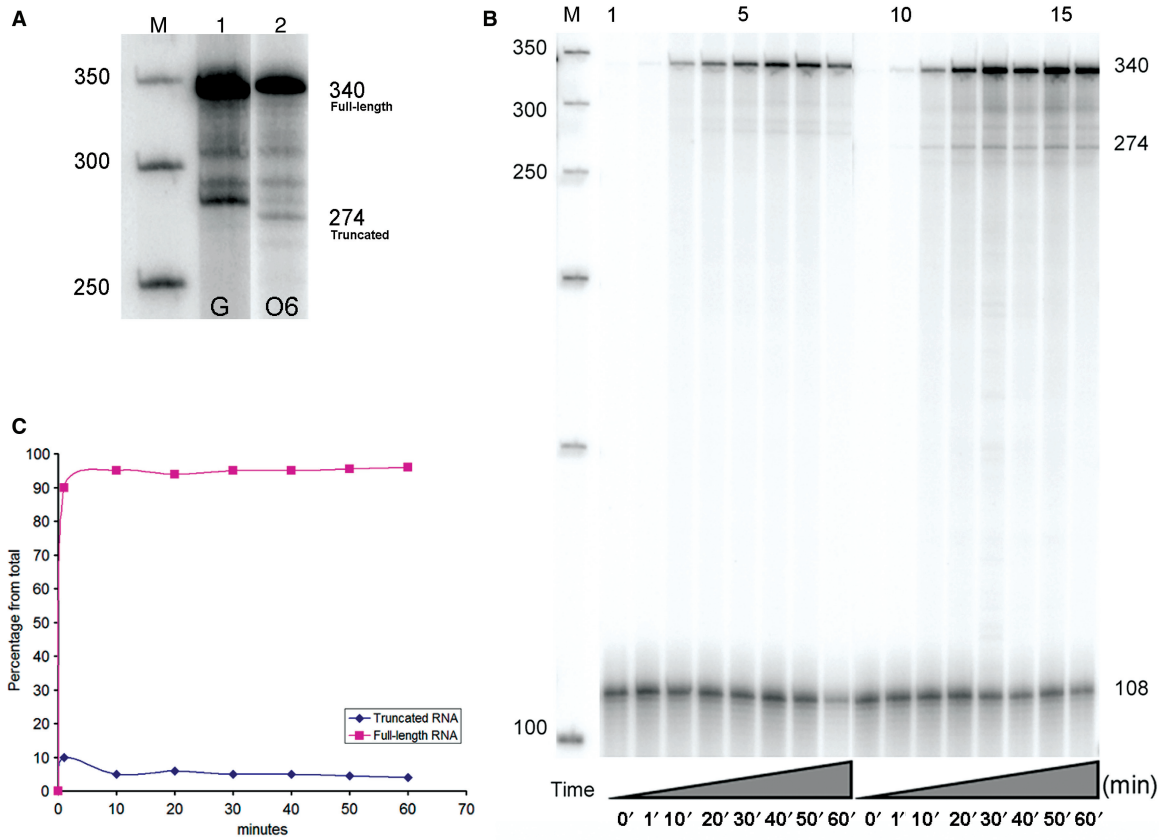

Figure 4. Effect of O^6-meG modified-DNA on transcription by T7 RNAP. (**A**) T7 RNAP transcription of the unmodified DNA resulted in full-length RNA (FL) approximately 340 bases in length (lane 1) while the modified template resulted in a mixture of two RNA species approximately 274 and 340 bases in length (lane 2). A 50 bp DNA step ladder is labeled M. (**B**) Time course of transcription with T7 RNAP past the undamaged DNA (lanes 1 through 8) and the O^6-meG template (lanes 9 through 16). Aliquots were removed every 10 min to measure the extent of lesion blockage or bypass over time. Fragments were compared to a DNA standard of 108 bases in length used to minimize loading errors. The increased intensity across lanes indicates an increased amount of RNA over time. Transcription past the damaged templates shows that the lesion poses a minor obstruction to the rapid RNA synthesis. (**C**) The graph illustrates the percent of full-length RNA. It indicates that the amount of full-length RNA remains the same over time (95%) and that the lesion causes the production of ~5% of truncated RNA at any time point.

protocol (www.promega.com/tbs/tb123/tb123.html). Both single and multiple rounds of transcription were carried out on undamaged and damaged templates. For single rounds of transcription, a G-less cassette was utilized to synchronize the reaction so that a template could be transcribed only once by a single RNA polymerase (54). For multiple rounds of transcription, a template could be transcribed multiple times by several polymerase complexes. After 60 min, both single and multiple rounds of transcription using the unmodified control resulted in the formation of full-length transcripts 383 bases in length, with no truncated transcripts occurring in the vicinity of the O^6-meG (Figure 5A, lane 1). However, transcription using the damaged template produced truncated and full-length RNA of 316 and 383 bases in length, respectively. Single round of transcription produced $64 \pm 6\%$ of full-length RNA, while multiple rounds of transcription produced $74 \pm 2\%$ of full-length RNA. The data obtained represent the average of 10 separate transcription experiments, with standard deviations from the mean given.

As with T7RNAP, time course experiments were carried out to examine hRNAPII lesion bypass efficiency over time. We carried out these transcription reactions in the presence of 0.5 μM BZG to ensure that repair of O^6-meG does not occur. The time course shows that undamaged templates only produce full-length RNA over time (Figure 5B, lanes 1 through 8), while templates containing O^6-meG produced a constant ratio of full-length (~60%) and truncated RNA (~40%) at each time point after the first 10 min (Figure 5B, lanes 10 through 17). These results suggest that RNA synthesis is partially obstructed by O^6-meG and that repair of the damaged templates is unlikely to occur during the course of the performed transcription reaction as performed here.

Note that the addition of BZG to the HeLa nuclear extracts resulted in a decrease in the amount of full-length RNA, from 74% to ~58% at 60 min, indicating that the extracts may have some MGMT activity which is thwarted in the presence of BZG, eliminating repair of the lesion. Note that nucleotide excision repair (NER) does not recognize O^6-meG in human cell extracts (55), indicating that this pathway is not responsible for the repair of the lesion. Furthermore, mismatch repair does not remove O^6-meG but replaces the mismatched thymine on the opposite strand (56); hence, this pathway is also not responsible for the clearance of this lesion in the extracts.

T7 RNAP and hRNAPII generate extensive base misincorporation opposite O^6-meG in full-length RNA

RT–PCR was used to sequence the full-length and truncated RNA following transcription of DNA templates that contain O^6-meG. The resulting PCR products consisted of a heterogeneous population of DNA, necessitating subcloning to analyze individual RT-PCR products. For full-length transcripts resulting from T7 RNAP transcription past the O^6-meG templates, the polymerase

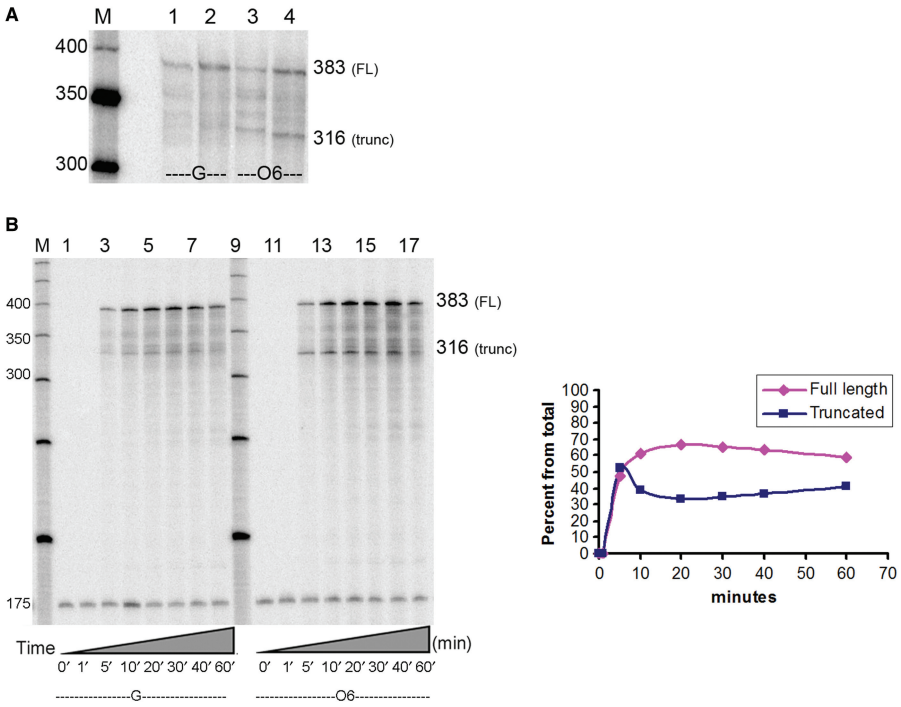

Figure 5. Effect of O^6-meG modified-DNA on transcription by hRNAPII. (**A**) Single and multiple rounds of transcription using HeLa nuclear extracts. Undamaged templates generated full-length (FL) transcripts of approximately 383 bases in length (lanes 1 and 2) while O^6-meG-damaged templates produced two RNA species of 383 and 316 bases in length (lanes 3 and 4). A 50 bp DNA step ladder is shown (M). (**B**) Time course of transcription with hRNAPII past the undamaged DNA (lanes 1 through 8) and the O^6-meG template (lanes 10 through 17). Aliquots were removed at the time points indicated on the figure to measure the extent of lesion blockage or bypass over time. Fragments were compared to a DNA standard of 175 bases in length used to minimize loading errors. Transcription past the damaged templates shows that the lesion partially obstructs RNA synthesis (only producing ~60% of full-length RNA at time points 20 min through 60 min). The graph illustrates the % of full-length and truncated RNA over time (*x*-axis). It indicates that the amount of full-length and truncated RNA remain almost constant after the first ten minutes of the transcription reaction.

primarily incorporated both pyrimidines equally (47% each) and adenine to a very small extent (6%) (Table 2). Undamaged templates produced only full-length transcripts having the correct base sequence (data not shown). In the case of full-length RNA obtained for hRNAPII bypass of O^6-meG, cytosine and uracil were exclusively incorporated opposite the lesion with a 3:1 ratio. Misincorporation of adenine was not observed in full-length RNA generated by hRNAPII (Table 2).

For the truncated transcripts, T7 RNAP often stalled exactly at the site of the lesion and did not permit nucleotide incorporation opposite it. In some cases, T7 RNAP was able to misincorporate either one adenine or a cytosine followed by two non-template instructed adenines before terminating transcription (Table 3). In contrast, hRNAPII typically stopped at the site of O^6-meG, but without nucleotide incorporation opposite the lesion. Additionally, hRNAPII misincorporated cytosine opposite the base immediately preceding O^6-meG, which was cytosine, and then stopped (Table 3).

Modeling suggests hydrogen-bonding possibilities that explain observed nucleotide selectivity

Bacteriophage T7 RNAP. Molecular modeling was employed to provide a structural rationale for the experimentally observed findings: full-length RNA exhibited significant base misincorporation events opposite the lesion with base selectivity C = U >> A, and shorter transcripts were truncated one residue before the lesion. Two glycosidic bond conformations (χ) for the O^6-meG nucleoside were investigated, namely *anti* and *syn* (Figure 1). Our models showed that when the O^6-meG nucleoside was in the *anti* conformation, it allowed the insertion of pyrimidine NTPs with no steric collisions, provided that the methyl group adopted the *proximal* conformation relative to the N1 atom (Figure 1B). This conformation avoided collision with incoming pyrimidines and allowed base pairing to occur. For CTP, two hydrogen-bonding interactions were identified, one between N4 of cytosine and O^6 of O^6-meG and another between O^2 of cytosine and N^2 of O^6-meG. In case of an N3-protonated cytosine, an additional hydrogen bond could occur with the N1 of O^6-meG. A protonated cytosine would allow a Watson–Crick-like pairing between the lesion and the incoming NTP similar to what was identified crystallographically (29,33,36) (Figure 6B). In the case of UTP, two hydrogen bonds were found possible, one between N3 of uridine and N1 of O^6-meG, and the second between O^2 of uridine and N^2 of O^6-meG. This O^6-meG:U pair is in a 'Watson–Crick isosteric' conformation (36), also termed 'pseudoWatson–Crick' (22), and was observed crystallographically (33,34,36) (Figure 6C).

In order to accommodate an incoming adenine, the O^6-meG nucleoside had to adopt the *syn* conformation around its glycosidic bond (χ) to avoid overlap with the six-membered ring of the adenine. In addition, the methyl group had to be oriented in the *distal* conformation (Figure 1C). In this conformation, the methyl could be housed, without collision, in the polymerase active site pocket on the major groove side of the damaged guanine, allowing the incorporation of adenine opposite it. Interestingly, this *syn* conformation was stabilized by one hydrogen bond between the O^6 of O^6-meG and a protein side chain of Gln-649 (Figure 7 and Supplementary Table S2B). The *syn* (χ) *distal* (κ) conformation of the lesion allowed for one hydrogen-bonding interaction between the N^6 of the incoming ATP and the O^6 of the O^6-meG. For the case of an N1-protonated adenine, as observed in crystal structures (57,58), an additional hydrogen bond could form between the N1 hydrogen of the ATP with the N7 of the lesion (Figure 6D).

These hydrogen-bonding possibilities offer an explanation for the incorporation of cytosine and the misincorporation of uracil and adenine in the RNA opposite O^6-meG during transcription elongation by T7 RNAP: equal insertion of cytosine and uracil is consistent with their equivalent hydrogen-bonding possibilities with O^6-meG, while the much smaller selection for adenine is explained by the requirement for the abnormal and less-prevalent *syn* conformation of the templating lesion in order to allow hydrogen bonding with the incoming ATP (Figures 6 and 8). We have given detailed results for the substrate insertion complex, but the models were essentially identical for the pre-translocation product complex and quite similar in the post-translocation and the substrate pre-insertion complexes. Detailed results of the torsion angles and the hydrogen-bonding partners are given in Tables S1 and S2.

Yeast RNAPII (yRNAPII). As detailed in the Materials and Methods section, current evidence indicates that the active site regions of yeast and human RNA polymerase II are in fact nearly identical (51,52). This makes the yeast polymerase a suitable model for the human enzyme. Hence, molecular modeling was employed to explain the

Table 2. Bases incorporated opposite O^6-meG in the full-length RNA during transcription with T7 RNAP and hRNAPII

% rNTP	C	U	A
RNAP			
T7 RNAP	47	47	6
hRNAPII	75	25	–

Table 3. Bases incorporated opposite O^6-meG in the truncated RNA during transcription with T7 RNAP and hRNAPII[a]

	Sequence	RNA length (bases)	Percentage
Expected RNA sequence	...UCUUGCGGUAG<u>C</u>GAUGG		
Sequence obtained with T7 RNAP	...UCUUGCGGUAG	272	25
	...UCUUGCGGUAG<u>**A**</u>	273	25
	...UCUUGCGGUAG<u>**CAA**</u>	275	50
Sequence obtained with hRNAPII	...UCUUGCGGUAG	272	80
	...UCUUGCGGUA**C**		20

[a]Bold, underscored bases represent misincorporated bases after the lesion.

observed incorporation of C and to lesser extent U opposite the lesion in the full-length RNA using yRNAPII. No adenine incorporation was observed. In the case of the truncated transcripts, nucleotide incorporation did not occur opposite the O^6-meG lesion. To explain the observed bypass events, the models had the glycosidic bond (χ) for the O^6-meG nucleoside in the *anti* conformation and there was a range in κ around the O^6–C6 bond which was collision free (κ = 200–250°), similar to the T7 RNAP bypass models. In this domain, the lesion was in an *anti, proximal* orientation (Figure 1B). We term this orientation the bypass conformation because it is compatible with nucleotide incorporation into the RNA. The *anti, proximal* orientation of the O^6-meG allowed for the formation of two hydrogen bonds with each of the incoming pyrimidines. For CTP, the hydrogen bonds involved the N^4 of cytosine with O^6 of O^6-meG, and the O^2 of cytosine with N^2 of O^6-meG. Again, in the case of an N3-protonated cytosine, an additional hydrogen bond could form with the N1 of O^6-meG. For UTP, the two hydrogen bonds involved the N3 of uridine with N1 of O^6-meG and the O^2 of uridine with N2 of O^6-meG. These

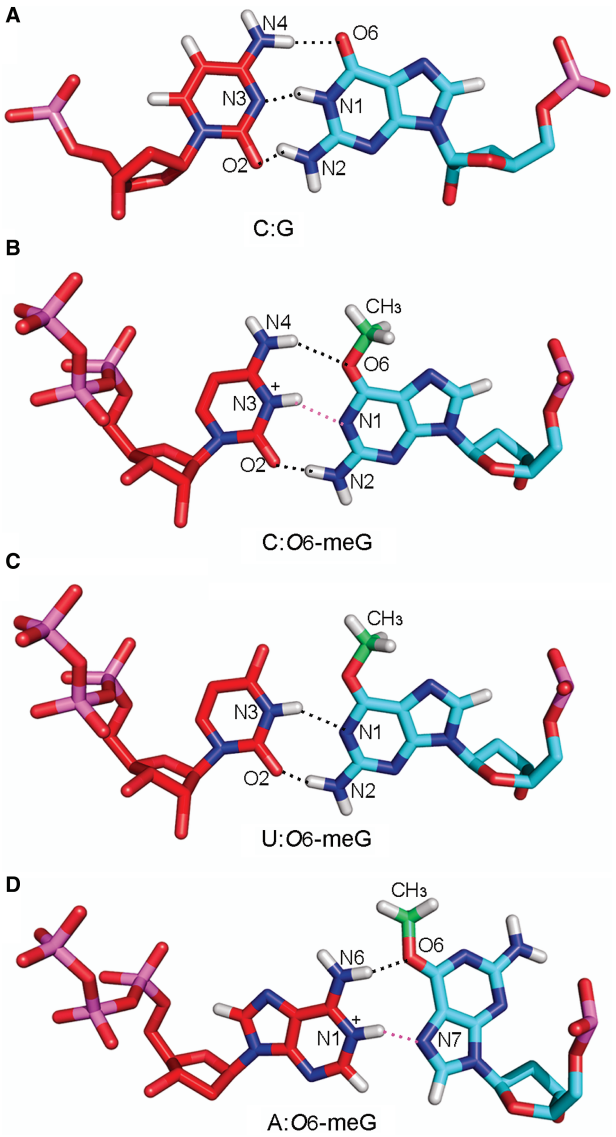

Figure 6. Hydrogen-bonding interactions between (**A**) a G:C pair, O^6-meG and (**B**) cytosine, (**C**) uridine and (**D**) adenine, respectively (see text). (A) G:C pair; G(O^6):C(N^4), G(N1):C(N3), and G(N^2):C(O^2), (B) *anti,proximal* O^6-meG:C pair; O^6-meG(O^6):C(N^4) and O^6-meG(N^2): C(O^2), (C) *anti,proximal* O^6-meG:U pair; O^6-meG(N1):U(N3) and O^6-meG(N^2):U(O^2) and (D) *syn, distal* O^6-meG:A pair; O^6-meG(O^6): A(N^6). Hydrogen bonds involving potential protonation sites, O^6-meG(N1):C($N3^+$) and O^6-meG(N7):A($N1^+$) are shown in pink. The methyl group carbon is colored in green. Color code is the same as in Figure 3. Details of the potential hydrogen-bonding partners are described in Supplementary Table S2.

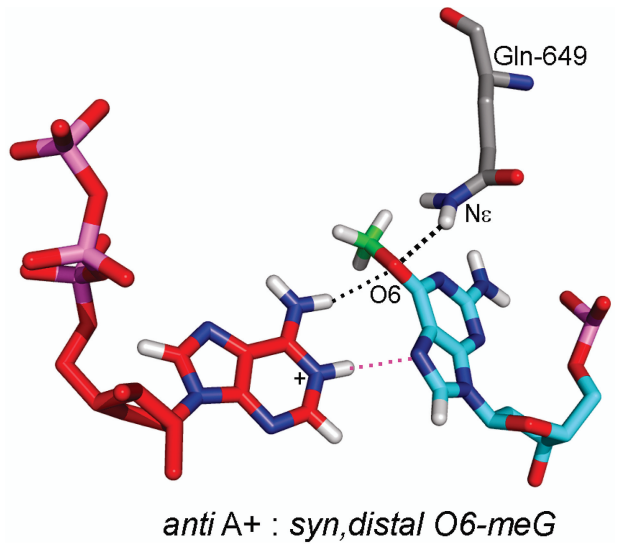

Figure 7. H-bond interaction between O^6-meG and the T7 RNAP residue, Gln-649. The O^6-meG in the *syn* conformation is stabilized by one hydrogen bond between the O^6 of O^6-meG and the N_ε of Gln-649. The hydrogen bond is shown as a dotted *black* line. Hydrogens of the backbone were removed for clarity. Color code is the same as in Figures 3 and 6.

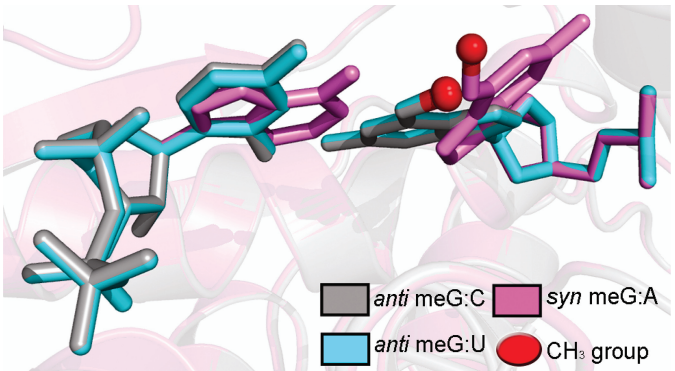

Figure 8. View of the O^6-meG lesion pairing with cytosine (gray), uracil (cyan) and adenine (violet) inside the active site of T7 RNAP. The backbone of the protein is rendered in transparent cartoons, superimposed on the three pairs and colored using the same code as the DNA:RNA pair shown in the figure. The methyl group of the lesion is shown as a *red* sphere. For lesion pairing with cytosine or uracil, the O^6-meG:C and O^6-meG:U pairs superimpose completely and mimic a normal Watson–Crick base pair. With the O^6-meG:A pair, the O^6-meG must adopt the less-prevalent *syn* conformation for collision-free placement of the incoming rATP.

hydrogen-bonding schemes, which are detailed in Figure 6 and Supplementary Table S3, are similar to what was obtained for T7 RNAP, and explain the incorporation of uracil and cytosine in the growing RNA.

When the methyl group was outside the allowed κ range, it collided with either the incoming ribonucleotide or neighboring protein residues, namely Thr-827, Ala-828, Val-829 of the bridge helix. These collisions could not be alleviated by minimal rotation of the side chains of the helix, because slight rotations of the bridge helix residues cause severe distortions to the yRNAPII structure (53). This conformational range ($\kappa \neq 200$–250°), termed the blocking structure, provides a model to explain the halting of the polymerase opposite O^6-meG.

The modeling studies showed that ATP could not be accommodated with the O^6-meG in the *anti* glycosidic bond conformation due to collisions between the lesion and the large purine. While, it was possible to obtain a collision-free structure containing ATP, this structure was otherwise unfavorable. The lesion could only be placed in a narrow region of the *syn* domain ($\chi = 238$–243°) with κ in the *distal* conformation ($\kappa = 172$–238°) (Figure 1C). This structure had two hydrogen bonds between the Hoogsteen edge of the O^6-meG and the A [N^6(A):O^6(meG) and N1$^+$(A):N7(meG)], but it contained an unfavorable repulsive interaction between the O^6 and an oxygen of Ala-828 of subunit 1 of yRNAPII, with close distance of ~2.6 Å (Supplementary Figure S4). This repulsive interaction together with the small region of conformational space in χ that is permissible sterically would make this structure unlikely to occur. Thus, the observed failure to incorporate adenine is reasonable on the basis of these models.

DISCUSSION

Gene expression is a tightly regulated process that requires RNA polymerase to read a transcription unit when directed. To achieve this, an unperturbed, processive elongation complex centered around an RNA polymerase is essential. However, transcription is often obstructed by pausing sequences, protein roadblocks (59) or DNA damage (60). More specifically, unrepaired damage present on the transcribed strand of a gene can either block an elongating RNA polymerase or lead to bypass with nucleotide insertion opposite or near the site of the damage (60). While arrest of transcription is believed to trigger TCR, lesion bypass can induce TM (25).

Here we report that, in the current sequence context, O^6-meG partially blocks transcription elongation of both T7 RNAP and hRNAPII, albeit to different extents (Figures 4 and 5); the lesion poses a more significant block to hRNAPII. Although, there have been contradictory reports on the preferential repair of O^6-meG from the transcribed strand of a gene (61,62), our results indicate that partial polymerase blockage by O^6-meG raises the possibility that this lesion might invoke TCR (63). Hence, the blocked polymerase could provide an additional means for cells to identify O^6-meG, not recognized by the MGMT protein. This possibility needs to be investigated further.

In addition, we show the contribution of O^6-meG to TM *in vitro*. T7 RNAP bypasses the lesion efficiently but with low fidelity, resulting in a high frequency of nucleotide misinsertion opposite the lesion with selectivity C = U = 47% >> A = 6%. Similarly, hRNAPII bypasses O^6-meG with misincorporation of uracil in 25% of the full-length RNA. During gene expression and mRNA translation, each protein-encoding mRNA is translated about 40 times on average; therefore, an error in the primary code of the RNA, such as produced by O^6-meG, can lead to a burst effect of flawed proteins (64). In this connection, it is interesting that the incorporation of uracil was evidenced during a transcriptional assay past O^6-meG in human cells (Burns,J.A. Dreij,K. and Scicchitano,D.A. unpublished data). Hence, our work adds O^6-meG to the pool of smaller DNA lesions that partially block transcription and may induce TM; others include the oxidative lesions 8-oxoguanine (65,66), thymine glycols (67) and cisplatin 1,2-intrastrand d(GpG) crosslinks (68).

Sequences of the truncated RNAs showed that when T7 RNAP and hRNAPII stopped at the lesion, nucleotide incorporation opposite O^6-meG was not permitted (Table 3). Modeling indicated that in both polymerases transcription blockage would ensue if the lesion adopts an *anti, distal* orientation, shown in Figure 1A. In the *distal* orientation, the methyl group is in close proximity to the incoming ribonucleotide, producing steric collisions that thwart incorporation and elongation. Moreover, the greater blockage observed with hRNAPII could be attributed to the occurrence of a *distal* conformation of the methyl group which highly crowds the incoming NTP in the more confined hRNAPII. Interestingly, two additional truncated RNAs were obtained during transcription by T7 RNAP with either an additional adenine or a cytosine followed by two non-template instructed adenines (Table 3). These two truncated RNA species were also observed in other studies where T7 RNAP stalled at damaged guanines namely 1,N^2-ethenogunaine (1,N^2-εG) (26) and 5-guanidino-4-nitroimidazole (NI) (27). This behavior may therefore be inherent to the stalled T7 RNAP and independent of the lesion. It has been suggested that stalled RNA polymerases tend to add one or two bases at the 3′-end of transcripts (69) further supporting our observation.

The bypass events observed with O^6-meG can be explained on the basis of the altered hydrogen-bonding possibilities of O^6-meG compared to guanine (Figure 1). Specifically, the hydrogen-bond acceptor atom N1 of guanine becomes a hydrogen-bond donor in O^6-meG. With the lesion in an *anti-proximal* conformation our models showed that, in both polymerases, there were no collisions with the incoming UTP or CTP. We readily modeled a pseudo Watson–Crick base pair with uracil (22) and a Watson–Crick-like pair with cytosine (36). Each pair allowed two hydrogen bonds between the lesion and the incoming nucleotide (Figure 6B and C). A protonated cytosine allowed three hydrogen bonds to form with an alignment that is very near Watson–Crick. These pairing schemes were feasible in both T7 RNAP

and yRNAPII, and a number of crystal structures (22,33–36) and NMR solution studies (29–32) have revealed the base-pairing possibilities for O^6-meG present in our models. While a wobble interaction with cytosine [N^4(C):N1(O^6-meG) and N3(C):N^2(O^6-meG)] was also seen in crystal structures (70) and inferred by NMR solution studies (32), this scheme requires rearrangement in the DNA backbone which could be energetically costly. The preference for cytosine over uracil incorporation opposite O^6-meG in yRNAPII probably stems from the more confining active site region of yRNAPII, which likely selects more stringently for a hydrogen-bonding scheme closest to Watson–Crick. Our models show that adenine could only be accommodated in T7 RNAP when O^6-meG was in the abnormal *syn* orientation around its glycosidic bond; explaining the low adenine misincorporation (Figure 7). No favorable models were possible for the yRNAPII, explaining why adenine incorporation was not seen in the full-length transcripts.

CONCLUSION

The present work extends our understanding of lesion processing by RNA polymerases and the possible attendant sequelae of either TM or TCR. Our work and that of others have provided studies of bulky lesions derived from polycyclic aromatic carcinogens [reviewed in (60,71)] with diverse shapes and stereochemical properties, as well as the rigid 1,N^2-εG (26), the flexible NI (27) and here the smaller O^6-meG. From these investigations, we deduce that the unique and specific hydrogen-bonding capabilities of the lesion with different incoming NTPs play an important role in TM. For the specific case of O^6-meG, the altered hydrogen-bond donor/acceptor properties (Figure 1) allow for enhanced selectivity for uracil compared to guanine in both T7 RNAP and hRNAPII. In addition, lesion flexibility has been pinpointed as an important structural factor based on our studies of the conformationally heterogenous NI lesion (27). Our results show that for O^6-meG, the possible variation in the torsion angle κ, which governs the rotation about the O^6–C6 bond, and hence the placement of the bulky methyl group in the polymerase active site is key. Both *distal* and *proximal* conformations have been explicitly observed by NMR (29–32) and X-ray crystallography (33,35,36). The conformation of the methyl group is the determinant for whether bypass or blockage occurs. Polymerase blockage and the subsequent induction of TCR is connected to steric hindrance by the lesion in the polymerase active site. Blockage structures stem from a *distal* orientation of the methyl group which places this sterically challenging group in collision with the templating base. Our results are consistent with earlier modeling studies suggesting that multiple conformations of O^6-meG account for the observed blockage and pyrimidine insertion during *in vitro* replication (72). Thus, the subtle structural distinctions between guanine and its O^6-methylated derivative produce significant consequences, not only as a mutational damage during replication, but through its impact on transcription. Further research is needed to discern the impact of O^6-meG on the production of faulty proteins in cells via TM and the potential impact of TCR on the lesion's clearance.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Professor Nicholas Geacintov for his critical reading of the manuscript and helpful discussions. We also thank the members of the Scicchitano Lab for reading and evaluation of the manuscript.

FUNDING

National Institutes of Health (ES010581 to D.A.S., 2R01 CA75449 and 5R01 CA28038 for partial support of computational infrastructure and systems management) Funding for open access charge: ES010581 to D.A.S.

Conflict of interest statement. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Cancer Institute or the National Institutes of Health.

REFERENCES

1. Beranek,D.T. (1990) Distribution of methyl and ethyl adducts following alkylation with monofunctional alkylating agents. *Mutat. Res.*, **231**, 11–30.
2. 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.
3. Sedgwick,B. (1997) Nitrosated peptides and polyamines as endogenous mutagens in O6-alkylguanine-DNA alkyltransferase deficient cells. *Carcinogenesis*, **18**, 1561–1567.
4. Liu,J.G. and Li,M.H. (1989) Roussin red methyl ester, a tumor promoter isolated from pickled vegetables. *Carcinogenesis*, **10**, 617–620.
5. De Bont,R. and van Larebeke,N. (2004) Endogenous DNA damage in humans: a review of quantitative data. *Mutagenesis*, **19**, 169–185.
6. Wyatt,M.D. and Pittman,D.L. (2006) Methylating agents and DNA repair responses: methylated bases and sources of strand breaks. *Chem. Res. Toxicol.*, **19**, 1580–1594.
7. Loechler,E.L. (1994) A violation of the Swain-Scott principle, and not SN1 versus SN2 reaction mechanisms, explains why carcinogenic alkylating agents can form different proportions of adducts at oxygen versus nitrogen in DNA. *Chem. Res. Toxicol.*, **7**, 277–280.
8. Souliotis,V.L., Kaila,S., Boussiotis,V.A., Pangalis,G.A. and Kyrtopoulos,S.A. (1990) Accumulation of O6-methylguanine in human blood leukocyte DNA during exposure to procarbazine and its relationships with dose and repair. *Cancer Res.*, **50**, 2759–2764.
9. Taverna,P., Catapano,C.V., Citti,L., Bonfanti,M. and D'Incalci,M. (1992) Influence of O6-methylguanine on DNA damage and cytotoxicity of temozolomide in L1210 mouse leukemia sensitive and resistant to chloroethylnitrosoureas. *Anticancer Drugs*, **3**, 401–405.
10. Toorchen,D. and Topal,M.D. (1983) Mechanisms of chemical mutagenesis and carcinogenesis: effects on DNA replication of methylation at the O6-guanine position of dGTP. *Carcinogenesis*, **4**, 1591–1597.
11. Loechler,E.L., Green,C.L. and Essigmann,J.M. (1984) In vivo mutagenesis by O6-methylguanine built into a unique site in a viral genome. *Proc. Natl Acad. Sci. USA*, **81**, 6271–6275.

12. Singer,B. and Dosanjh,M.K. (1990) Site-directed mutagenesis for quantitation of base-base interactions at defined sites. *Mutat. Res.*, **233**, 45–51.
13. Marchesi,F., Turriziani,M., Tortorelli,G., Avvisati,G., Torino,F. and De Vecchis,L. (2007) Triazene compounds: mechanism of action and related DNA repair systems. *Pharmacol. Res.*, **56**, 275–287.
14. Voigt,J.M. and Topal,M.D. (1995) O6-methylguanine-induced replication blocks. *Carcinogenesis*, **16**, 1775–1782.
15. Galloway,S.M., Greenwood,S.K., Hill,R.B., Bradt,C.I. and Bean,C.L. (1995) A role for mismatch repair in production of chromosome aberrations by methylating agents in human cells. *Mutat. Res.*, **346**, 231–245.
16. Kaina,B., Fritz,G. and Coquerelle,T. (1993) Contribution of O6-alkylguanine and N-alkylpurines to the formation of sister chromatid exchanges, chromosomal aberrations, and gene mutations: new insights gained from studies of genetically engineered mammalian cell lines. *Environ. Mol. Mutagen.*, **22**, 283–292.
17. Zhang,H., Tsujimura,T., Bhattacharyya,N.P., Maher,V.M. and McCormick,J.J. (1996) O6-methylguanine induces intrachromosomal homologous recombination in human cells. *Carcinogenesis*, **17**, 2229–2235.
18. Pegg,A.E., Dolan,M.E., Scicchitano,D. and Morimoto,K. (1985) Studies of the repair of O6-alkylguanine and O4-alkylthymine in DNA by alkyltransferases from mammalian cells and bacteria. *Environ. Health Perspect.*, **62**, 109–114.
19. Pegg,A.E., Fang,Q. and Loktionova,N.A. (2007) Human variants of O(6)-alkylguanine-DNA alkyltransferase. *DNA Repair (Amst).*, **6**, 1071–1078.
20. Chen,J.M., Zhang,Y.P., Wang,C., Sun,Y., Fujimoto,J. and Ikenaga,M. (1992) O6-methylguanine-DNA methyltransferase activity in human tumors. *Carcinogenesis*, **13**, 1503–1507.
21. Choi,J.Y., Chowdhury,G., Zang,H., Angel,K.C., Vu,C.C., Peterson,L.A. and Guengerich,F.P. (2006) Translesion synthesis across O6-alkylguanine DNA adducts by recombinant human DNA polymerases. *J. Biol. Chem.*, **281**, 38244–38256.
22. Eoff,R.L., Irimia,A., Egli,M. and Guengerich,F.P. (2007) Sulfolobus solfataricus DNA polymerase Dpo4 is partially inhibited by "wobble" pairing between O6-methylguanine and cytosine, but accurate bypass is preferred. *J. Biol. Chem.*, **282**, 1456–1467.
23. Woodside,A.M. and Guengerich,F.P. (2002) Misincorporation and stalling at O(6)-methylguanine and O(6)-benzylguanine: evidence for inactive polymerase complexes. *Biochemistry*, **41**, 1039–1050.
24. Woodside,A.M. and Guengerich,F.P. (2002) Effect of the O6 substituent on misincorporation kinetics catalyzed by DNA polymerases at O(6)-methylguanine and O(6)-benzylguanine. *Biochemistry*, **41**, 1027–1038.
25. Viswanathan,A., Liu,J. and Doetsch,P.W. (1999) E. coli RNA polymerase bypass of DNA base damage. Mutagenesis at the level of transcription. *Ann. NY Acad. Sci.*, **870**, 386–388.
26. Dimitri,A., Goodenough,A.K., Guengerich,F.P., Broyde,S. and Scicchitano,D.A. (2008) Transcription processing at 1,N2-ethenoguanine by human RNA polymerase II and bacteriophage T7 RNA polymerase. *J. Mol. Biol.*, **375**, 353–366.
27. Dimitri,A., Jia,L., Shafirovich,V., Geacintov,N.E., Broyde,S. and Scicchitano,D.A. (2008) Transcription of DNA containing the 5-guanidino-4-nitroimidazole lesion by human RNA polymerase II and bacteriophage T7 RNA polymerase. *DNA Repair (Amst)*, **7**, 1276–1288.
28. Perlow,R.A., Kolbanovskii,A., Hingerty,B.E., Geacintov,N.E., Broyde,S. and Scicchitano,D.A. (2002) DNA adducts from a tumorigenic metabolite of benzo[a]pyrene block human RNA polymerase II elongation in a sequence- and stereochemistry-dependent manner. *J. Mol. Biol.*, **321**, 29–47.
29. Williams,L.D. and Shaw,B.R. (1987) Protonated base pairs explain the ambiguous pairing properties of O6-methylguanine. *Proc. Natl Acad. Sci. USA*, **84**, 1779–1783.
30. Patel,D.J., Shapiro,L., Kozlowski,S.A., Gaffney,B.L. and Jones,R.A. (1986) Structural studies of the O6meG.C interaction in the d(C-G-C-G-A-A-T-T-C-O6meG-C-G) duplex. *Biochemistry*, **25**, 1027–1036.
31. Patel,D.J., Shapiro,L., Kozlowski,S.A., Gaffney,B.L. and Jones,R.A. (1986) Structural studies of the O6meG.T interaction in the d(C-G-T-G-A-A-T-T-C-O6meG-C-G) duplex. *Biochemistry*, **25**, 1036–1042.
32. Gaffney,B.L., Goswani,B. and Jones,R.A. (1993) Nitrogen-15-labeled oligodeoxynucleotides. Use of 15H NMR to probe H-bonding in an O6MeG.C base pair. *J. Am. Chem. Soc.*, **115**, 12607–12608.
33. Ginell,S.L., Kuzmich,S., Jones,R.A. and Berman,H.M. (1990) Crystal and molecular structure of a DNA duplex containing the carcinogenic lesion O6-methylguanine. *Biochemistry*, **29**, 10461–10465.
34. Ginell,S.L., Vojtechovsky,J., Gaffney,B., Jones,R. and Berman,H.M. (1994) Crystal structure of a mispaired dodecamer, d(CGAGAATTC(O6Me)GCG)2, containing a carcinogenic O6-methylguanine. *Biochemistry*, **33**, 3487–3493.
35. Vojtechovsky,J., Eaton,M.D., Gaffney,B., Jones,R. and Berman,H.M. (1995) Structure of a new crystal form of a DNA dodecamer containing T.(O6Me)G base pairs. *Biochemistry*, **34**, 16632–16640.
36. Warren,J.J., Forsberg,L.J. and Beese,L.S. (2006) The structural basis for the mutagenicity of O(6)-methyl-guanine lesions. *Proc. Natl Acad. Sci. USA*, **103**, 19701–19706.
37. Perlow,R.A., Schinecker,T.M., Kim,S.J., Geacintov,N.E. and Scicchitano,D.A. (2003) Construction and purification of site-specifically modified DNA templates for transcription assays. *Nucleic Acids Res.*, **31**, e40.
38. Schinecker,T.M., Perlow,R.A., Broyde,S., Geacintov,N.E. and Scicchitano,D.A. (2003) Human RNA polymerase II is partially blocked by DNA adducts derived from tumorigenic benzo[c]phenanthrene diol epoxides: relating biological consequences to conformational preferences. *Nucleic Acids Res.*, **31**, 6004–6015.
39. Dolan,M.E., Moschel,R.C. and Pegg,A.E. (1990) Depletion of mammalian O6-alkylguanine-DNA alkyltransferase activity by O6-benzylguanine provides a means to evaluate the role of this protein in protection against carcinogenic and therapeutic alkylating agents. *Proc. Natl Acad. Sci. USA*, **87**, 5368–5372.
40. Elbashir,S.M., Lendeckel,W. and Tuschl,T. (2001) RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes Dev.*, **15**, 188–200.
41. Temiakov,D., Patlan,V., Anikin,M., McAllister,W.T., Yokoyama,S. and Vassylyev,D.G. (2004) Structural basis for substrate selection by t7 RNA polymerase. *Cell*, **116**, 381–391.
42. Yin,Y.W. and Steitz,T.A. (2004) The structural mechanism of translocation and helicase activity in T7 RNA polymerase. *Cell*, **116**, 393–404.
43. Yin,Y.W. and Steitz,T.A. (2002) Structural basis for the transition from initiation to elongation transcription in T7 RNA polymerase. *Science*, **298**, 1387–1395.
44. Yamagata,Y., Kohda,K. and Tomita,K. (1988) Structural studies of O6-methyldeoxyguanosine and related compounds: a promutagenic DNA lesion by methylating carcinogens. *Nucleic Acids Res.*, **16**, 9307–9321.
45. Tabernero,L., Bella,J. and Aleman,C. (1996) Hydrogen bond geometry in DNA-minor groove binding drug complexes. *Nucleic Acids Res.*, **24**, 3458–3466.
46. Panigrahi,S.K. and Desiraju,G.R. (2007) Strong and weak hydrogen bonds in the protein-ligand interface. *Proteins*, **67**, 128–141.
47. Torshin,I.Y., Weber,I.T. and Harrison,R.W. (2002) Geometric criteria of hydrogen bonds in proteins and identification of "bifurcated" hydrogen bonds. *Protein Eng.*, **15**, 359–363.
48. McDonald,I.K. and Thornton,J.M. (1994) Satisfying hydrogen bonding potential in proteins. *J. Mol. Biol.*, **238**, 777–793.
49. Fabiola,F., Bertram,R., Korostelev,A. and Chapman,M.S. (2002) An improved hydrogen bond potential: impact on medium resolution protein structures. *Protein Sci.*, **11**, 1415–1423.
50. Spratt,T.E. and Levy,D.E. (1997) Structure of the hydrogen bonding complex of O6-methylguanine with cytosine and thymine during DNA replication. *Nucleic Acids Res.*, **25**, 3354–3361.
51. Cramer,P., Bushnell,D.A. and Kornberg,R.D. (2001) Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. *Science*, **292**, 1863–1876.
52. Kostek,S.A., Grob,P., De Carlo,S., Lipscomb,J.S., Garczarek,F. and Nogales,E. (2006) Molecular architecture and conformational flexibility of human RNA polymerase II. *Structure*, **14**, 1691–1700.

53. Wang,D., Bushnell,D.A., Westover,K.D., Kaplan,C.D. and Kornberg,R.D. (2006) Structural basis of transcription: role of the trigger loop in substrate specificity and catalysis. *Cell*, **127**, 941–954.
54. Viswanathan,A. and Doetsch,P.W. (1998) Effects of nonbulky DNA base damages on Escherichia coli RNA polymerase-mediated elongation and promoter clearance. *J. Biol. Chem.*, **273**, 21276–21281.
55. Kusumoto,R., Masutani,C., Sugasawa,K., Iwai,S., Araki,M., Uchida,A., Mizukoshi,T. and Hanaoka,F. (2001) Diversity of the damage recognition step in the global genomic nucleotide excision repair in vitro. *Mutat. Res.*, **485**, 219–227.
56. Karran,P. and Bignami,M. (1994) DNA damage tolerance, mismatch repair and genome instability. *Bioessays*, **16**, 833–839.
57. Brown,T., Leonard,G.A., Booth,E.D. and Chambers,J. (1989) Crystal structure and stability of a DNA duplex containing A(anti).G(syn) base-pairs. *J. Mol. Biol.*, **207**, 455–457.
58. Pan,B., Mitra,S.N. and Sundaralingam,M. (1999) Crystal structure of an RNA 16-mer duplex R(GCAGAGUUAAAUCUGC)2 with nonadjacent G(syn).A+(anti) mispairs. *Biochemistry*, **38**, 2826–2831.
59. Pavco,P.A. and Steege,D.A. (1990) Elongation by Escherichia coli RNA polymerase is blocked in vitro by a site-specific DNA binding protein. *J. Biol. Chem.*, **265**, 9960–9969.
60. Scicchitano,D.A., Olesnicky,E.C. and Dimitri,A. (2004) Transcription and DNA adducts: what happens when the message gets cut off? *DNA Repair (Amst)*, **3**, 1537–1548.
61. Souliotis,V.L., Sfikakis,P.P., Anderson,L.M. and Kyrtopoulos,S.A. (2004) Intra- and intercellular variations in the repair efficiency of O6-methylguanine, and their contribution to kinetic complexity. *Mutat. Res.*, **568**, 155–170.
62. Pegg,A.E., Scicchitano,D. and Dolan,M.E. (1984) Comparison of the rates of repair of O6-alkylguanines in DNA by rat liver and bacterial O6-alkylguanine-DNA alkyltransferase. *Cancer Res.*, **44**, 3806–3811.
63. Brueckner,F., Hennecke,U., Carell,T. and Cramer,P. (2007) CPD damage recognition by transcribing RNA polymerase II. *Science*, **315**, 859–862.
64. Libby,R.T. and Gallant,J.A. (1991) The role of RNA polymerase in transcriptional fidelity. *Mol. Microbiol.*, **5**, 999–1004.
65. Tornaletti,S., Maeda,L.S., Kolodner,R.D. and Hanawalt,P.C. (2004) Effect of 8-oxoguanine on transcription elongation by T7 RNA polymerase and mammalian RNA polymerase II. *DNA Repair (Amst)*, **3**, 483–494.
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.
67. 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.
68. Cullinane,C., Mazur,S.J., Essigmann,J.M., Phillips,D.R. and Bohr,V.A. (1999) Inhibition of RNA polymerase II transcription in human cell extracts by cisplatin DNA damage. *Biochemistry*, **38**, 6204–6212.
69. Milligan,J.F., Groebe,D.R., Witherell,G.W. and Uhlenbeck,O.C. (1987) Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. *Nucleic Acids Res.*, **15**, 8783–8798.
70. Sriram,M., van der Marel,G.A., Roelen,H.L., van Boom,J.H. and Wang,A.H. (1992) Structural consequences of a carcinogenic alkylation lesion on DNA: effect of O6-ethylguanine on the molecular structure of the d(CGC[e6G]AATTCGCG)-netropsin complex. *Biochemistry*, **31**, 11823–11834.
71. Scicchitano,D.A. (2005) Transcription past DNA adducts derived from polycyclic aromatic hydrocarbons. *Mutat. Res.*, **577**, 146–154.
72. Dosanjh,M.K., Loechler,E.L. and Singer,B. (1993) Evidence from in vitro replication that O6-methylguanine can adopt multiple conformations. *Proc. Natl Acad. Sci. USA*, **90**, 3983–3987.