The use of an artificial nucleotide for polymerase-based recognition of carcinogenic O\textsuperscript{6}-alkylguanine DNA adducts

Laura A. Wyss\textsuperscript{1,†}, Arman Nilforoushan\textsuperscript{1,†}, David M. Williams\textsuperscript{2}, Andreas Marx\textsuperscript{3} and Shana J. Sturla\textsuperscript{1,*}

\textsuperscript{1}Department of Health Sciences and Technology, ETH Zürich, 8092 Zürich, Switzerland, \textsuperscript{2}Center for Chemical Biology, Department of Chemistry, Krebs Institute, University of Sheffield, Sheffield S3 7HF, UK and \textsuperscript{3}Department of Chemistry, Konstanz Research School Chemical Biology, University of Konstanz, 78457 Konstanz, Germany

Received May 20, 2016; Revised June 20, 2016; Accepted June 21, 2016

ABSTRACT

Enzymatic approaches for locating alkylation adducts at single-base resolution in DNA could enable new technologies for understanding carcinogenesis and supporting personalized chemotherapy. Artificial nucleotides that specifically pair with alkylated bases offer a possible strategy for recognition and amplification of adducted DNA, and adduct-templated incorporation of an artificial nucleotide has been demonstrated for a model DNA adduct O\textsuperscript{6}-benzylguanine by a DNA polymerase. In this study, DNA adducts of biological relevance, O\textsuperscript{6}-methylguanine (O\textsuperscript{6}-MeG) and O\textsuperscript{6}-carboxymethylguanine (O\textsuperscript{6}-CMG), were characterized to be effective templates for the incorporation of benzimidazole-derived 2′-deoxynucleoside-5′-O-triphosphates (BenziTP and BIMTP) by an engineered KlenTaq DNA polymerase. The enzyme catalyzed specific incorporation of the artificial nucleotide Benzi opposite adducts, with up to 150-fold higher catalytic efficiency for O\textsuperscript{6}-MeG over guanine in the template. Furthermore, addition of artificial nucleotide Benzi was required for full-length DNA synthesis during bypass of O\textsuperscript{6}-CMG. The specific processing of BenziTP opposite biologically relevant O\textsuperscript{6}-alkylguanine adducts is characterized herein as a basis for potential future DNA adduct sequencing technologies.

INTRODUCTION

Human genetic material is under constant attack by harmful substances. For example, exposure to alkylating agents from the diet, tobacco smoke, environment, and chemotherapeutics, as well as endogenous sources can lead to DNA damage by chemical alkylation of nucleophilic sites on DNA bases giving rise to DNA adducts (1,2). Among the types of DNA adducts that may be formed, O\textsuperscript{6}-alkylguanine (O\textsuperscript{6}-alkylG) adducts are of important biological relevance because of their high propensity for inducing mutations (3,4), including G to A transitions prevalent in cancer, for example in codon 12 or 13 of the proto-oncogene K-ras (5), and the tumor suppressor gene p53 (6). Therefore, together with an understanding of the causal relationship between adduct formation and mutagenesis, strategies for locating alkylation adducts in DNA is an important basis to establish early biomarkers of carcinogenesis (7).

Among O\textsuperscript{6}-alkylG adducts, O\textsuperscript{6}-methylguanine (O\textsuperscript{6}-MeG) and O\textsuperscript{6}-carboxymethylguanine (O\textsuperscript{6}-CMG; Figure 1A) have been found to be present in human blood DNA (8) and tissue samples from meat-eaters and cancer patients (9,10). Putative sources of O\textsuperscript{6}-MeG include methylating agents like temozolomide (13).

© The Author(s) 2016. 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
sequence context, nor do they have the potential for amplification of the DNA adduct (16).

Discrimination between various DNA adducts and unmodified DNA with single-base resolution has been achieved by single molecule real time (SMRT) sequencing (17). This approach has excellent potential for future biological applications, but amplification or enrichment of the target sequence is still required prior to sequencing, along with further development of quantitation algorithms. Burrows and co-workers have reported two methods for sequencing oxidative lesions that are substrates for base excision repair (BER) (18,19). Following specific excision of the adducts with BER DNA glycosylases, they either marked the site with an amplifiable unnatural base pair or they identified the adduct location by introducing a deletion mutation (18,19). However, such specific enzymes do not exist for DNA alklylation adducts, thus an alternative strategy could be valuable.

Artificial nucleosides that specifically pair with DNA alklylation adducts together with DNA polymerases with the capacity to process altered base pairs offer a basis for alklylation adduct sensing at single base resolution. Polymerase-mediated incorporation of synthetic triphosphates opposite DNA damage has been reported for abasic sites, isoguanine, 8-oxoguanine (8-oxoG), cis-platinated guanine, and O6-alkylG adducts (20–25). For example, the synthetic nucleoside triphosphate dAdapTP discriminates 8-oxoG from G in single nucleotide incorporation studies by the A-family DNA polymerase Klenow (exo-) (22). Artificial bases like BIM and Benzi, and related analogues in oligonucleotides acted as O6-alkylG adduct-specific base pairing partners and resulted in the formation of more stable DNA duplexes when paired opposite O6-alkylG adducts vs unmodified G (26–28). Furthermore, these analogues were polymerase substrates for extending DNA primers terminated with some of the artificial bases paired opposite O6-alkylG adducts (29,30). Artificial nucleotides BIMTP and BenziTP acted as impedes substrates for human DNA polymerase η (hPol η) in replication of the major cisplatin DNA adduct (23). Recently, we communicated specific incorporation of BenziMP opposite O6-benzylguanine (O6-BnG) adducts versus nondamaged guanine templates by a mutant KlenTaq polymerase KTqM747K. Furthermore, BenziTP was required for full-length product formation in bypass of this bulky lesion and additionally was used for amplification of alkylated DNA in linear PCR by KTqM747K polymerase (25). This discovery was the first report of an artificial nucleotide being specifically incorporated opposite an O6-alkylG DNA adduct, but previous studies were carried out with the model adduct O6-BnG, which has not been observed in vivo (31).

Herein, we addressed the scope of an alklylation adduct-artificial nucleotide replication system with regards to O6-alkylG adducts of biological relevance, namely O6-MeG and O6-CMG (Figure 1A). We investigated the bypass of these adducts by a mutant KlenTaq DNA polymerase using artificial nucleotides as substrates (Figure 1B). The DNA polymerase KTqM747K is a mutant of the N-terminally truncated A-family Taq polymerase. It is thermostable and can efficiently bypass various DNA lesions (25,32,33). We found that the O6-alkylG adducts template the specific incorporation of the artificial nucleotide BenziMP when DNA synthesis was carried out by KTqM747K. Furthermore, full-length products were formed following effective incorporation and extension of Benzi nucleotide. A 2′,3′-dideoxy Benzi-nucleotide was newly synthesized and allowed verification that the artificial nucleotide does not impede KTqM747K polymerase in processive DNA replication past natural templates. Furthermore, it enabled marking of the adduct site with the artificial nucleotide. Finally, we found that with BenziTP biologically relevant O6-alkylG adducts could be recognized in mixtures of damaged and non-damaged DNA. The findings demonstrated herein represent a chemical basis for enzymatic O6-alkylG adduct detection technologies at single-base resolution that are required for establishing biomarkers of cancer risk or chemotherapeutic drug efficacy.

**MATERIALS AND METHODS**

**Chemical reagents and materials**

Reagents were purchased from Sigma-Aldrich and used without further purification. Nucleoside analogues BIM and Benzi (30,34), and triphosphates BIMTP and BenziTP were synthesized as described previously (25). O6-MeG and unmodified 5′-O-dimethoxytrityl phosphoramidites were purchased from Link Technologies Ltd. The O6-CMG phosphoramidite was prepared as reported (35). Unlabeled dNTPs were obtained from Invitrogen and [γ-32P]ATP was purchased from PerkinElmer Life Sciences. KTqM747K mutant DNA polymerase was kindly provided by myPOLS Biotec GmbH. Pyrophosphatase (Inorganic, Escherichia coli) was purchased from New England Biolabs. Silica gel 60 F254 plates with aluminum backing were used for thin layer chromatography. Flash column chromatography was performed on a Biotage system with pre-packed Flash+ KP-SiO2 cartridges. 1H, 13C and 31P NMR spectra were recorded on a Bruker Biospin 400 MHz NMR instrument, and chemical shifts are reported in parts per million (ppm, δ) relative to the chemical shift of the respective NMR solvent. High resolution mass spectra were recorded on
Thermo Scientific exactive mass spectrometer with electro-spray ionization.

**Oligonucleotides**

Oligonucleotides were either purchased from Eurofins, Microsynth or Eurogentec. DNA sequences are listed in Supplementary Table S1. Modified oligonucleotides containing $O^6$-CMG were synthesized as described elsewhere (36). Oligonucleotides were purified by reverse phase HPLC on a Phenomenex Luna C-18 column (5 μm, 4.6 x 250 mm). The $O^6$-CMG DNA 28mer was purified with a mobile phase gradient of 10.5–14.5% acetonitrile in 50 mM TEAA over 50 min and eluted at 31 minutes (Supplementary Figures S1 and S2). 48mer $O^6$-CMG DNA was prepared and purified as reported elsewhere (35). Corresponding oligonucleotide fractions were collected and combined, dried on a centrifugal vacuum concentrator and stored at −20°C until further use. The ssDNA concentration was determined by UV spectroscopy at 260 nm on a NanoDrop 1000 spectrophotometer. Theoretical molar extinction coefficients of the DNA sequences were determined using Integrated DNA technologies online at [http://eu.idtdna.com/analyzer/Applications/OligoAnalyzer/](http://eu.idtdna.com/analyzer/Applications/OligoAnalyzer/).

**Primer extension assays**

Radioactive labeling of primer strands at their 5’ end was carried out using T4 polynucleotide kinase (Promega) and [γ-32P] ATP following manufacturer protocol. Primer and templates were annealed by incubating at 95°C for 5 min and slow cooling over 12 h. Final concentrations were 1 μM primer and 1.5 μM template. Standard primer extension reactions (10 μl) contained 1× KTq reaction buffer, 5 mM enzyme, 15 nM DNA (15 nM primer and 22.5 nM template), and 10 μM dNTPs. In full-length DNA synthesis experiments, reactions contained all four natural dNTPs (10 μM total) with or without BenziTP (10 μM). Primer/template, nucleotides and DNA polymerase were incubated at 55°C for 10 min. 1× KTq reaction buffer contained 50 mM Tris–HCl (pH 9.2), 16 mM (NH₄)₂SO₄, 2.5 mM MgCl₂, and 0.1% Tween 20. Reactions were quenched by adding 20 μl PAGE gel loading buffer (80% formamide, 20 mM EDTA, 0.05% bromophenol blue, 0.05% xylene cyanole FF) and the product mixtures were analyzed by 7M urea denaturing gels and subjected to autoradiography (Bio-Rad). Quantification was carried out with Bio-Rad Quantity One software.

**Steady-state kinetic analysis**

Steady-state kinetics parameters for single nucleotide incorporation by DNA polymerase KTqM747K were determined under single completed hit conditions (37,38). For various dNTP concentrations, the quantity of $n + 1$ product formed by performing the reaction at 55°C was measured. Reaction mixtures included 5 nM enzyme, 100 nM primer, 150 nM template, and 1× KTq reaction buffer. Reactions were initiated by adding pre-warmed enzyme and DNA mix to pre-warmed dNTPs. Reactions were quenched by adding PAGE loading buffer. Products were separated on a 15% polyacrylamide/7M urea denaturing gel, visualized by autoradiography, and quantified with Quantity One Software (Bio-Rad). To obtain kinetic parameters $v_{max}$, $K_M$ and $k_{cat}$, the intensities of $n + 1$ bands (quantified on the Quantity One Software, Bio-Rad) were fit to a Michaelis–Menten rectangular hyperbola using SigmaPlot12 Software (Systat Software). Reactions were performed in triplicate and for $K_M$ values, means (± standard deviations) are reported.

**Linear amplification of $O^6$-alkylG DNA**

Reactions were incubated in a Biometra T3000 thermocycler, where 30 cycles of denaturation, annealing, and elongation were performed under the following conditions: 30 s at 95°C, 30 s at 42°C, and 30 s at 55°C. Reactions contained 0.5 ng of corresponding $O^6$-alkylG template DNA (28 nt), 300 nM primer (19 nt), 25 nM KTqM747K DNA polymerase, 1× KTq reaction buffer, all four dNTPs (10 μM), and 10 μM or no BenziTP. Product mixtures were separated on a 20% polyacrylamide/7M urea denaturing gel, stained with SybrGold nucleic acid gel stain (Invitrogen) and visualized on a Bio-Rad molecular imager Gel Doc XR+ Imaging System. Product bands were quantified using the software Image Lab 3.0 (Bio-Rad) and normalized to the known quantity of the 28 nt marker loaded on the same gel, indicating a yield of 14.6 ng ampiclon (theoretical yield on the basis of 30 cycles is 15 ng).

**Molecular modeling studies**

Structures were computed with the Molecular Operating Environment software suite (Chemical Computing Group). Crystal structures of a KTq mutant polymerase with incoming ddCTP opposite template G (PDB code: 3PY8) (33) and structures of Bst DNA polymerase with incoming dCTP opposite template G (PDB code: 1LV5) or incoming ddTTP opposite template $O^6$-MeG (PDB code: 2HHW) were used (39). For modeling studies with the KTq mutant, crystal structure (PDB code: 3PY8) was modified by attaching an $O^6$-CMG group to the templating G and replacing incoming ddCTP by BenziTP or BIMTP (in favored syn conformation). In studies involving Bst DNA polymerase with template G incoming dCTP (PDB code: 1LV5) was replaced by syn BenziTP and for template $O^6$-MeG, ddTTP was altered to syn BenziTP (PDB code: 2HHW). For energy minimizations the potential energy of the protein was fixed and followed by applying the Amber99 force field. Visualization was performed in the PyMol software (Schrodinger).

**Application of ddBenziTP to mark $O^6$-CMG adducts in DNA**

The same protocol was followed as described for primer extension reactions. ddBenziTP was added at increasing concentrations (0, 100, 500, 1000 and 2000 μM) to all four natural dNTPs (10 μM) and BenziTP (10 μM). Reaction mixtures (10 μl) contained 1x KTq reaction buffer, 5 nM KTqM747K DNA polymerase, 15 nM DNA (15 nM radiolabeled primer annealed to 22.5 nM template), and 1/60 units pyrophosphatase, and were allowed to react at 55°C for 10 min. Reactions were quenched by adding 20 μl PAGE gel loading buffer, and analyzed by separating on
Sensing $O^6$-alkyG-containing DNA mixed with unmodified DNA

Mixtures of G and $O^6$-alkyG DNA were prepared with a constant amount of template (22.5 nM template annealed to 15 nM radiolabeled primer) at varying ratios of G:$O^6$-alkyG DNA: 1:0, 1000:1, 100:1, 10:1, 5:1, 3:1, 2:1, 1:1. Primer extension reaction mixtures (10 μl) contained 1 × KTq reaction buffer, 5 nM enzyme, 15 nM DNA, and 10 μM BenziTP DNA mixtures were prewarmed and BenziTP/Polymerase mixture was added and allowed to react at 55°C for 10 min. Products were separated on a 15% polyacrylamide/7M urea denaturing gel and visualized by autoradiography. Intensity of n + 1 bands was quantified on the Quantity One Software (Bio-Rad). The experiment was performed in triplicate and mean values ± standard deviations are reported.

RESULTS AND DISCUSSION

Translesion DNA synthesis past $O^6$-alkyG adducts by KTqM747K

To investigate the capacity of the mutant KlenTaq polymerase KTqM747K to replicate DNA containing $O^6$-MeG or $O^6$-CMG, primer extension studies were performed. Thus, a 5'-end radiolabeled 23 nucleotide (nt) primer and a 28 nt template containing either G, $O^6$-MeG, or $O^6$-CMG (referring to X positioned at nucleotide 24, Figure 2A) were allowed to react with KTqM747K polymerase and dNTPs, (Figure 2). Extension products were analyzed by gel shift assay on denaturing polyacrylamide gels and visualized by autoradiography. For results presented in Figure 2, the level of nucleotide incorporation is indicated as percent primer extension and was calculated as a ratio of the amount of n + 1 extension product formed to initial amount of primer.

The capacity of KTqM747K to replicate DNA containing $O^6$-alkylG adducts in the presence of four natural dNTPs depend on adduct structure. Replication in the presence of $O^6$-MeG gave rise to full-length product (26%, Figure 2B, X = $O^6$-MeG, lane 4), whereas in the presence of $O^6$-CMG significantly less full-length product was observed (7%, Figure 2B, X = $O^6$-CMG, lane 4). In both cases, misincorporation of dTMP was favored over correct dCMP incorporation (Figure 2B). For $O^6$-MeG, dTTP was incorporated to a large extent (92%) and dCMP was also incorporated (41%). For $O^6$-CMG templates, natural nucleotides were incorporated less than they were for $O^6$-MeG templates. Thus, there was 32% dTMP misincorporation and 17% correct dCMP incorporation. In a previous study concerning $O^6$-BnG (25), the polymerase was stalled. Considering the proficiency of $O^6$-MeG bypass, stalling in the other two cases may be attributed to the larger sizes of the adducts. For $O^6$-BnG, dTTP misincorporation was favored over dCTP incorporation (32 vs 18%) (25). Misincorporation of dTTP has been observed in bypass of $O^6$-MeG by bacterial Escherichia coli KF DNA polymerase (40), B. stearothermophilus Bst polymerase, Vent (exo-) (41), viral

T4 (42), T7 (exo-) DNA polymerase and HIV reverse transcriptase (43), or eukaryotic Drosophila melanogaster Pol α (44). Also, human Pol β and translesion Pol λ had a similar preference for dTTP when replicating over $O^6$-MeG (4,45).

Specific incorporation of an artificial nucleotide opposite $O^6$-alkyG DNA adducts

Having established how KTqM747K polymerase bypasses $O^6$-alkyG adducts with natural nucleotides as substrates, the incorporation of two benzimidazole-derived base-modified nucleotide analogues (BIMTP and BenziTP) was investigated in the same manner as described above, adding BIMTP or BenziTP in primer extension reactions. Both artificial nucleotides were effective substrates for KTqM747K polymerase (Figure 2B). However, BIMTP yielded little extension product in general: G (20%), $O^6$-MeG (11%), and $O^6$-CMG (14%; Figure 2B). On the other hand, BenziTP was a good substrate and there were high incorporation percentages for both adducts (Figure 2B, X = $O^6$-MeG: 95%; X = $O^6$-CMG: 87%) but, importantly, not with G (Figure 2B, X = G: 12%). This observation matched the preference for alkylated templates previously reported for $O^6$-
BnG (25). Since natural nucleotides also were incorporated readily under the standard reaction conditions (55°C, 10 min) when O6-MeG was in the template, conditions were adjusted to improve selectivity for BenziTP over natural dNTPs in the case of O6-MeG (SI, Figure S3). By carrying out the reaction at 72°C for 2 min, incorporation of natural nucleotides was dramatically reduced (25% dTTP, 6% dCTP), whereas BenziMP incorporation resulted in 80% extension product and only 4% incorporation opposite G. In summary, BenziMP was specifically incorporated opposite O6-alkyl-G and DNA adducts over G.

To investigate the influence of the bases flanking the DNA adduct in this process, we studied two alternative sequences besides –CXT- with O6-CMG at position X (SI, Figures S4 and S5). In one sequence, we changed the 5’ pyrimidine to a purine (–GXT–) and in the second, the 3’ T was replaced by a C (–GXC–). In both cases, BenziTP was favored as a substrate over natural dNTPs when O6-CMG was in the template. Furthermore, full-length products were formed to a significantly higher extent in the presence of all four natural dNTPs plus BenziTP than in reactions without BenziTP (SI, Figures S4 and S5; <10% full-length product with 4 dNTPs only vs up to 50% full-length product in presence of 4 dNTPs plus BenziTP).

**Artificial nucleotide is required for efficient full-length DNA synthesis of O6-CMG templates**

Knowing that BenziMP is efficiently incorporated opposite O6-alkylG adducts, we characterized how the artificial nucleotide impacted full-length DNA synthesis in lesion bypass. Thus, we examined whether a 19 nt primer could be elongated if the template contained O6-alkylG five bases downstream from the primer terminus (at nt 24) in the presence of all four natural dNTPs and BenziTP (Figure 3). The O6-MeG adduct was readily bypassed by KTqM747K polymerase and full-length products (over 90%) were formed with natural dNTPs only or with supplemented BenziTP (Figure 3B, X = O6-MeG). However, when O6-CMG was in the template, replication was stalled (Figure 3B, X = O6-CMG, lane 4; Bands at 23 nt and at adduct site X, 24 nt). The addition of BenziTP significantly promoted the formation of full-length products in this case. With both alkylated templates, a prominent band at 29 nt was visible, likely due to template-independent incorporation of an additional nucleotide (46). BenziTP was required for efficient bypass of O6-CMG by overcoming the stalling of the polymerase at the adduct site to result in full-length products.

**Steady-state kinetic analysis of translesion DNA synthesis**

In order to quantitatively compare efficiencies of nucleotide incorporation in replication of DNA containing O6-MeG or O6-CMG, steady-state kinetic parameters for KTqM747K polymerase catalysis were determined (Table 1). In this experiment, single nucleotide incorporation (n+1 product formation) was followed over time, and kinetic parameters $K_M$ and $k_{cat}$ were derived (37). In general, the presence of the O6-alkylG DNA adducts did not greatly influence the catalytic turnover $k_{cat}$, but decreased the binding affinity (increasing $K_M$) compared to replication of unmodified DNA with natural dNTPs. The catalytic efficiencies $k_{cat}/K_M$ for synthesis past O6-MeG (0.038 μM⁻¹min⁻¹), O6-CMG (0.007 μM⁻¹min⁻¹), and O6-BnG (0.021 μM⁻¹min⁻¹) (25) by KTqM747K were similar to values measured previously for incorporation of dTMP opposite O6-MeG by the A-family Bst DNA polymerase (0.075 μM⁻¹min⁻¹) (39). Compared to misincorporation of dTMP by the archaeal Sulfolobus solfataricus DNA polymerase Dpo4 (dTTP opposite O6-MeG: 0.0044 μM⁻¹min⁻¹) (47), the KTqM747K polymerase tested herein was 9-fold more efficient in replicating O6-MeG adducts. The human translesion DNA polymerase η on the other hand is more efficient in translesion synthesis past O6-alkylG adducts than KTqM747K; 20-fold more for O6-MeG (48), and 9-fold more in bypass of O6-CMG (35).

The $K_M$ value for processing of BenziTP was the same order of magnitude for unmodified as well as adducted templates (Table 1). However, catalytic turnover $k_{cat}$ was significantly higher for replication over O6-alkylG adducts versus G (17-fold for O6-CMG; 55-fold for O6-MeG). The highest catalytic efficiency $k_{cat}/K_M$ for BenziMP incorporation was observed when O6-alkylG was in the template, and was 13-fold higher than for O6-CMG or 6-fold than for O6-BnG ($k_{cat}/K_M$ 0.120 μM⁻¹min⁻¹) (25). Finally, comparing these values with catalytic efficiencies for incorporation of BenziMP opposite natural templates (25), incorporation opposite A was almost 2-fold more efficient than with templating O6-MeG. However, processing of BenziTP opposite A was 30-fold less efficient than incorporation of dTMP (25). When comparing incorporation of BenziMP vs dTMP, the difference in selectivity was 19-fold with the O6-MeG template, 8-fold for O6-CMG, and 6-fold for O6-BnG (25). We previously demonstrated that BenziMP was also incorporated by hPol η during replication of unmodified or platinated G templates (23). However, incorporation of BenziMP by KTqM747K opposite O6-MeG was 29-fold more efficient than its incorporation by hPol η opposite the major platinum intrastrand cross-link product cis-Pr-1,2-d(GpG) ($k_{cat}/K_M$ 0.025 μM⁻¹min⁻¹) (23). In summary, kinetic data confirmed the specific incorporation of BenziMP opposite O6-alkylG adducts compared to guanine with the catalytic efficiency of incorporation opposite G being reduced 150-fold compared to O6-MeG, 24-fold for O6-BnG (25), and 12-fold for O6-CMG. Steady-state kinetic data re-
revealed the fastest incorporation efficiencies for Benzi were opposite $O^6$-MeG, and the highest selectivity for BenziTP over natural dNTPs was observed for $O^6$-MeG templates. Findings from primer extension and steady-state kinetic experiments indicate that KTem747K polymerase can readily bypass $O^6$-MeG, with natural nucleotides being well incorporated. Since KTem747K was significantly stalled by $O^6$-CMG and adding BenziTP promoted bypass of this adduct, we focused further attention on $O^6$-CMG.

**Molecular modeling of $O^6$-CMG: Benzi base pair**

Molecular modeling was performed to visualize a possible structural basis for the specific incorporation of Benzi opposite $O^6$-CMG. Thus, molecular mechanics simulations of base pairing interactions between BenziTP and G or $O^6$-CMG in the active site of a KTemTaq mutant polymerase (I614K, M747K; PDB code: 3PY8) (33) were performed. Original crystal structures containing an incoming ddCTP opposite templating G were modified by replacing ddCTP with BenziTP and adding a carboxymethyl group to the templating G. The modified nucleotides were placed in anti conformation, since in syn conformation a steric clash was evident between the substituent at position 2 on the nucleobase and the oxygen of the sugar moiety (23). Following energy minimization (Amber 99 force field, Figure 4 and SI, Figure S6) possible differences in base pair geometries were considered. For incoming BenziTP in the KTem mutant active site BenziTP was computed to be flipped out when paired with G (Figure 4A). Whereas Benzi:$O^6$-CMG was predicted to adopt a planar Watson-Crick-like geometry and potentially form two hydrogen bonds (Figure 4B): one between the –NH donor on Benzi and the N2 of $O^6$-CMG (2.4 Å); the other was predicted between the carbonyl group of Benzi and the –NH2 donor on $O^6$-CMG (2.1 Å). Similar findings were observed with hPol £ in replication of the major cisplatin adduct where two hydrogen bonds were possibly formed with Benzi when it was incorporated opposite the first base of a platinated GG site (23).

Additionally, we modelled BenziTP in the active site of another thermostable A-family polymerase, Bst DNA Pol, for which a crystal structure is available with an incoming ddTTP opposite $O^6$-MeG (PDB code: 2HHW) (39). Furthermore, Bst DNA Pol has the closest amino acid sequence homology to KTemTaq polymerase with 51% identity matches (blastp against PDB protein database, http://www.ncbi.nlm.nih.gov/, 15/01/16) They share high sequence similarities in three conserved motifs among A-family DNA polymerases (Supplementary Figure S6A). The structures were analyzed in the same manner described above, and the data indicate that BenziTP is anticipated to be extruded approaching a pairing relationship with a template G and adopt a co-planar, two hydrogen bond, configuration when paired with $O^6$-MeG (Supplementary Figure S7A). The modeling results suggest the importance of hydrogen bonding interactions within the polymerase active site in explaining the experimental data.

**Linear amplification of $O^6$-alkylG DNA**

With the knowledge that BenziTP specifically promotes full-length synthesis past $O^6$-CMG adducts, we further investigated whether adducted DNA could be linearly amplified using BenziTP, which is specifically incorporated opposite $O^6$-alkylG as a marker for the adduct (Supplementary Figure S7A). Thus, to a low amount of $O^6$-CMG DNA (28 nt, $X = O^6$-CMG at position 24 nt, 0.5 ng DNA) was added four dNTPs plus BenziTP and extension of a 19 nt primer to a full-length 28 nt product was monitored after 30 amplification cycles (95°C, 30 s; 42°C, 30 s; 55°C, 30 s) by

| dNTP       | $K_M$ [μM] | $k_{cat}$ [min$^{-1}$] | $k_{cat}/K_M$ [μM$^{-1}$min$^{-1}$] |
|------------|------------|------------------------|-------------------------------------|
| $X = G$    |            |                        |                                     |
| dCTP       | 0.07 ± 0.02| 14                     | 190                                 |
| BenziTP    | 48 ± 8     | 0.24                   | 0.005                               |
| $X = O^6$-CMG |
| dTTP       | 519 ± 54   | 3.6                    | 0.007                               |
| BenziTP    | 70 ± 10    | 4.06                   | 0.058                               |
| $X = O^6$-MeG |
| dTTP       | 410 ± 66   | 15.6                   | 0.038                               |
| BenziTP    | 18 ± 4     | 13.2                   | 0.730                               |

Table 1. Steady-state kinetic parameters for nucleotide incorporation by KTemM747K DNA polymerase.
With the newly synthesized ddBenziTP at hand, we used it to test whether Benzi is exclusively incorporated opposite O6-alkylG or if Benzi is also a substrate in replication of natural templates. Thus, templates containing G or O6-CMG (28 nt. X at 24 nt) were annealed to a 19 nt primer and allowed to react with KTqM747K polymerase, in the presence of the four natural dNTPs (10 μM) and BenziTP (10 μM) plus increasing concentrations of ddBenziTP (Figure 5). In reactions containing template G, only desired full-length products were formed in the presence of four natural dNTPs, BenziTP, and ddBenziTP.

Verification of Benzi incorporation opposite O6-CMG adducts

In order to verify that Benzi is only incorporated opposite the O6-alkylG adduct and not opposite canonical bases we performed a primer extension assay with a terminaling Benzi nucleotide. This experiment required a 2',3'-dideoxynucleoside analogue of BenziTP, which was not previously reported. Thus, we prepared the base-modified 2',3'-dideoxynucleoside-5'-O-triphosphate ddBenziTP (Scheme 1). The Benzi nucleoside 1 (34) was mesylated at the 3'-position affording the respective 5'-O-dimethoxytrityl-3'-O-mesyI Benzi nucleoside 2 (49). Protected nucleoside 2 was allowed to react with sodium iodide, and the 5'-dimethoxytrityl group was removed with acid treatment, yielding the 3'-iodo derivative 3 (50). Elimination of the iodide from 3'-iodo Benzi nucleoside 3 by palladium-catalyzed hydrogenation yielded 2',3'-dideoxy Benzi nucleoside 4 (50). Reaction with (4-chlorobutyl)(methyl)phosphoramic dichloride gave rise to benzotriazole intermediate 5 (51,52), which following reaction with benzyl alcohol resulted in the phosphoramidate 6. Next, monophosphate 6 was activated by catalytic hydrogenolysis with palladium on activated carbon. Filtration of the catalyst and subsequent addition of pyrophosphate resulted in formation of the desired triphosphate 7 (ddBenziTP) (25).
Recognition of DNA adducts in DNA mixtures with unmodified templates

Knowing that Benzi is incorporated selectively opposite O6-alkylG adducts, we examined whether the adducts could also be sensed in a mixture with non-damaged DNA. Thus, we performed a primer extension experiment with various dilutions of O6-alkyl- and G-containing DNA (at a constant template concentration). Reactions were carried out with BenziTP and KTqM747K, and the formation of n + 1 extension products was monitored (Figure 6). Reactions contained a 23 nt primer annealed to corresponding 28 nt templates with G or O6-alkylG adducts at nucleotide 24. Concentration-dependent formation of extension products was observed that correlated with increasing fraction of O6-alkylG present in a mixture with G templates (Figure 6, Supplementary Figure S8). For a 1:1 mixture of O6-CMG:G DNA, 35% product was observed. In the absence of the O6-CMG adduct, no extension product was formed (3%; Figure 6, dashed line). The lowest visually observable product (7%) was in the case of a 1:10 ratio of O6-CMG:G. When the experiment was performed with O6-MeG, 51% extension product was formed from a 1:1 mixture of O6-MeG:G DNA, and the lowest detectable dilution was also 1:10 O6-MeG:G (9%; Supplementary Figure S8). Critical limitations for addressing DNA adducts are that they exist at much lower levels and in the presence of a far larger excess unmodified DNA, as well as the lack of suitability of phosphorylating as a basis of an optimized detection strategy. Nonetheless, the capacity for artificial nucleotide incorporation in a mixture was achieved, and furthermore, the amount of O6-alkylG template present in the reactions, which corresponds to 0.17 ng O6-alkylG DNA mixed with 1.7 ng unmodified DNA, was 3-fold lower than was detected by linear amplification (0.5 ng O6-alkylG DNA; SI, Figure S7).

CONCLUSIONS

We investigated artificial nucleotides as substrates for the replication of DNA containing carcinogenic alkylation adducts O6-MeG and O6-CMG. We demonstrated that KTqM747K DNA polymerase specifically incorporated artificial nucleotide BenziMP opposite O6-MeG and O6-CMG independent of its sequence context, and was competent in further extension. The specific incorporation of BenziMP opposite O6-alkylG adducts versus G was 150-fold higher for O6-MeG, and 12-fold for O6-CMG. O6-MeG was readily bypassed by KTqM747K polymerase and full-length products were formed in the presence of natural dNTPs, whereas O6-CMG stalled the polymerase and BenziTP was required for efficient full-length DNA synthesis. A structural basis for O6-alkylG adduct-specific incorporation of Benzi may be due to favorable hydrogen bonding interactions with O6-alkylG while Benzi is extruded from the duplex when it is opposite G.

An additional advance described in this study was the preparation of the 2′,3′-dideoxynucleoside BenziTP, which was used to mark the adduct site in O6-CMG-containing DNA and confirm that Benzi is not incorporated opposite natural templates in full-length synthesis. This demonstration is the first example of an artificial dideoxynucleotide used in a sequencing experiment involving marking a DNA adduct site. The combination of artificial dideoxynucleotide and KTqM747K polymerase allowed us to demonstrate a basis whereby O6-alkylG adducts in mixtures with non-damaged G containing DNA may be sensed. While significant further adaptations for real biological applications, including enrichment strategies, sensitive analytical read-out methods, and further refinement of polymerase characteristics are needed, these findings are a chemical basis that suggest novel approaches for single-base resolution determination of mutagenic DNA adduct occurrence.

Figure 6. Recognition of DNA adduct in a mixture of O6-alkylG- and G-containing oligonucleotides. Top: Primer extension reactions for O6-CMG:G oligonucleotide mixtures (total concentration 22.5 μM template). Bottom: Percentages of n + 1 bands accounting for incorporation of BenziMP (10 μM) incubated with corresponding O6-CMG:G DNA mixtures (error bars represent standard error from triplicate measurements). Dashed horizontal line refers to background level from reactions that contained unmodified oligonucleotides only. Reactions were carried out at 55°C for 10 min.
SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENT

We are grateful to Michael Raez (ETH Zurich, Switzerland) for preparing oligonucleotides containing 6-O'-CMG.

FUNDING

European Research Council [260341]; Swiss National Science Foundation [156280]; ETH research commission [ETH-43 14-1]. Funding for open access charge: ETH Zurich

Conflict of interest statement. None declared.

REFERENCES

1. Gates,K.S. (2009) An overview of chemical processes that damage cellular DNA: spontaneous hydrolysis, alkylation, and reactions with radicals. Chem. Res. Toxicol., 22, 1747–1760.
2. 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.
3. Pauly,G.T. and Moschel,R.C. (2001) Mutagenesis by O(6)-methyl-, O(6)-ethyl-, and O(6)-benzyguanine and O(4)-methylhyamine in human cells: effects of O(6)-alkylguanine-DNA alkyltransferase and mismatch repair. Chem. Res. Toxicol., 14, 894–900.
4. Pence,M.G., Choi,J.Y., Egli,M. and Guengerich,F.P. (2010) Structural basis for proficient incorporation of dITP opposite O6-methylguanine by human DNA polymerase iota. J. Biol. Chem., 285, 40666–40672.
5. Cancer Genome Atlas, N. (2012) Comprehensive molecular characterization of human colon and rectal cancer. Nature, 487, 330–337.
6. Pfeifer,G.P., Denissenko,M.F., Olivier,M., Tretjakova,N., Hecht,S.S. and Hainaut,P. (2002) Tobacco smoke carcinogens, DNA damage and p53 mutations in smoking-associated cancers. Oncogene, 21, 7435–7451.
7. Hecht,S.S. (2003) Tobacco carcinogens, their biomarkers and tobacco-induced cancer. Nat. Rev. Cancer, 3, 733–744.
8. EUROGAST. (1994) O6-methylguanine in blood leucocyte DNA: an association with the geographic prevalence of gastric cancer and with low levels of serum pepsinogen A, a marker of severe chronic atrophic gastritis. The EUROGAST Study Group. Carcinogenesis, 15, 1815–1820.
9. Hall,C.N., Badawi,A.F., O’Connor,P.J. and Saffhill,R. (1991) The detection of alkylation damage in the DNA of human gastrointestinal tissues. Br. J. Cancer, 64, 59–63.
10. Lewin,M.H., Bailey,N., Bandeletova,T., Bowman,R., Cross,A.J., Pollock,J., Shaker,D.E. and Bingham,S.A. (2006) Red meat enhances the colonic formation of the DNA adduct O6-carboxymethylguanine: implications for colorectal cancer risk. Cancer Res., 66, 1859–1865.
11. Hecht,S.S. (1999) DNA adduct formation from tobacco-specific N-nitrosamines. Mutat. Res., 424, 127–142.
12. 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.
13. Middleton,M.R., Lee,S.M., Arance,A., Wood,M., Thaxter,N. and Margison,G.P. (2000) O6-methylguanine formation, repair protein depletion and clinical outcome with a 4 hr schedule of temozolomide in the treatment of advanced melanoma: results of a phase II study. Int. J. Cancer, 88, 469–473.
14. Curud,B.C., Zeng,Z., Singh,R. and Shuker,D.E. (2004) Detection of O6-carboxymethyl-2'-deoxyguanosine in DNA following reaction of nitric oxide with glycine and in human blood DNA using a quantitative immunoslot blot assay. Chem. Res. Toxicol., 17, 294–300.
15. Bouvard,V., Loomis,D., Guyton,K.Z., Grosse,Y., Ghissassi,F.E., Benbrahim-Tallaa,L., Guha,N., Mattock,H., Straif,K. and International Agency for Research on Cancer Monograph Working Group (2015) Carcinogenicity of consumption of red and processed meat. Lancet Oncol., 16, 1599–1600.
16. Himmelstein,M.W., Boogaard,P.J., Cadet,J., Farmer,P.B., Kim,J.H., Martin,E.A., Persaud,R. and Shuker,D.E. (2009) Creating context for the use of DNA adduct data in cancer risk assessment: II. Overview of methods of identification and quantitation of DNA damage. Crit. Rev. Toxicol., 39, 679–694.
17. Clark,T.A., Spittle,K.E., Turner,S.W. and Korlach,J. (2011) Direct detection and sequencing of damaged DNA bases. Genome Integr., 2, 10.
18. Riedl,J., Ding,Y., Fleming,A.M. and Burrows,C.J. (2015) Identification of DNA lesions using a third base pair for amplification and nanopore sequencing. Nat. Commun., 6, 8807.
19. Riedl,J., Fleming,A.M. and Burrows,C.J. (2016) Sequencing of DNA Lesions Facilitated by Site-Specific Excision via Base Excision Repair DNA Glycosylases Yielding Ligatable Gaps. J. Am. Chem. Soc., 138, 491–494.
20. Switzer,C.Y., Moroney,S.E. and Benner,S.A. (1993) Enzymatic recognition of the base pair between isocytidine and isoguanosine. Biochemistry, 32, 10489–10496.
21. Matray,T.J. and Kooi,E.T. (1999) A specific partner for abasic damage in DNA. Nature, 399, 704–708.
22. Taniguchi,Y., Kikukawa,Y. and Sasaki,S. (2015) Discrimination between 8-oxo-2’-deoxyguanosine and 2’-deoxyguanosine in DNA by the single nucleotide primer extension reaction with Adap triphosphate. Angew. Chem. Int. Ed. Engl., 54, 5147–5151.
23. Nilforoushan,A., Furrer,A., Wyss,L.A., van Loon,B. and Sturla,S.J. (2015) Nucleotides with altered hydrogen bonding capacities impede human DNA polymerase eta by reducing synthesis in the presence of the major cisplatin DNA adduct. J. Am. Chem. Soc., 137, 4728–4734.
24. Stornetta,A., Angelov,T., Guengerich,F.P. and Sturla,S.J. (2013) Incorporation of nucleoside probes opposite O6-methylguanine by Sulfolobus solfataricus DNA polymerase Dpo4: importance of hydrogen bonding. Chembiochem, 14, 1634–1639.
25. Wyss,L.A., Nilforoushan,A., Eichenseher,F., Suter,U., Blatter,N., Marx,A. and Sturla,S.J. (2015) Specific incorporation of an artificial nucleotide opposite a mutagenic DNA adduct by a DNA polymerase. J. Am. Chem. Soc., 137, 30–33.
26. Gong,J. and Sturla,S.J. (2007) A synthetic nucleoside probe that discerns a DNA adduct from unmodified DNA. J. Am. Chem. Soc., 129, 4882–4883.
27. Trantakis,I.A. and Sturla,S.J. (2014) Gold nanoprobe for detecting DNA adducts. Chem. Commun., 50, 15517–15520.
28. Trantakis,I.A., Nilforoushan,A., Dahlmann,H.A., Stauble,C.K. and Sturla,S.J. (2016) In-gene quantification of O6-methylguanine with elongated nucleoside analogues on gold nanoprobes. J. Am. Chem. Soc., 138, 2229–2232.
29. Gahan,H.L., Boby,M.L. and Sturla,S.J. (2014) O6-alkylguanine postlesion DNA synthesis is correct with the right complement of hydrogen bonding. ACS Chem. Biol., 9, 2807–2814.
30. Gahan,H.L., Schweizer,W.B. and Sturla,S.J. (2013) Tolerance of base pair size and shape in postlesion DNA synthesis. J. Am. Chem. Soc., 135, 6384–6387.
31. Peterson,L.A. (1997) N-Nitrosobenzylmethylamine is activated to a DNA benzylating agent in rats. Chem. Res. Toxicol., 10, 19–26.
32. Gloeckner,C., Sauter,K.B. and Marx,A. (2007) Evolving a thermostable DNA polymerase that amplifies from highly damaged templates. Angew. Chem. Int. Ed. Engl., 46, 3115–3117.
33. Obeid,S., Schmar,A., Gloeckner,C., Blatter,N., Weite,W., Diederichs,K. and Marx,A. (2011) Learning from directed evolution: Thermus aquaticus DNA polymerase mutants with translesion synthesis activity. ChemBiochem., 12, 1574–1580.
34. Gahan,H.L. and Sturla,S.J. (2013) Hydrogen bonding or stacking interactions in differentiating duplex stability in oligonucleotides containing synthetic nucleoside probes for alkylated DNA. Chem. Eur. J., 19, 11062–11067.
35. Raez,M.H., Dexter,H., Millington,C.L., van Loon,B., Williams,D.M. and Sturla,S.J. (2016) Translesion DNA synthesis past O6-carboxymethylguanine DNA by human Y- and B-Family polymerases promote error-free and error-prone bypass. Chem. Res. Toxicol., 29, in press.
36. Millington, C.L., Watson, A.J., Marriott, A.S., Margison, G.P., Povey, A.C. and Williams, D.M. (2012) Convenient and efficient syntheses of oligodeoxyribonucleotides containing O(6)-(carboxymethyl)guanine and O(6)-(4-oxo-4-(3-pyridyl)butyl)guanine. *Nucleos. Nucleot. Nucl.*, 31, 328–338.

37. Boosalis, M.S., Petruska, J. and Goodman, M.F. (1987) DNA polymerase insertion fidelity. Gel assay for site-specific kinetics. *J. Biol. Chem.*, 262, 14689–14696.

38. Creighton, S., Bloom, L.B. and Goodman, M.F. (1995) Gel fidelity assay measuring nucleotide misinsertion, exonucleolytic proofreading, and lesion bypass efficiencies. *Methods Enzymol.*, 262, 232–256.

39. Warren, J.J., Forsberg, L.J. and Beese, L.S. (2006) The structural basis for the mutagenicity of O(6)-methylguanine lesions. *Proc. Natl. Acad. Sci. U.S.A.*, 103, 19701–19706.

40. Tan, H.B., Swann, P.F. and Chance, E.M. (1994) Kinetic analysis of the coding properties of O6-methylguanine in DNA: the crucial role of the conformation of the phosphodiester bond. *Biochemistry*, 33, 5335–5346.

41. Lim, S., Song, L., Guengerich, F.P. and Choi, J.Y. (2012) Effects of N(2)-alkylguanine, O(6)-alkylguanine, and abasic lesions on DNA binding and bypass synthesis by the euryarchaeal B-family DNA polymerase Vent (exo(-)). *Chem. Res. Toxicol.*, 25, 1699–1707.

42. Chavarria, D., Ramos-Serrano, A., Hirao, I. and Berdis, A.J. (2011) Exploring the roles of nucleobase desolvation and shape complementarity during the misreplication of O(6)-methylguanine. *J. Mol. Biol.*, 412, 325–339.

43. 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.

44. Singer, B., Chavez, F., Goodman, M.F., Essigmann, J.M. and Dosanjh, M.K. (1989) Effect of 3’ flanking neighbors on kinetics of pairing of dCTP or dTTP opposite O6-methylguanine in a defined primed oligonucleotide when Escherichia coli DNA polymerase I is used. *Proc. Natl. Acad. Sci. U.S.A.*, 86, 8271–8274.

45. Singh, I., Su, L. and Snow, E.T. (1996) Replication across O6-methylguanine by human DNA polymerase beta in vitro. Insights into the futile cytotoxic repair and mutagenesis of O6-methylguanine. *J. Biol. Chem.*, 271, 28391–28398.

46. Obeid, S., Blatter, N., Kranaster, R., Schnur, A., Diederichs, K., Wele, W. and Marx, A. (2010) Replication through an abasic DNA lesion: structural basis for adenine selectivity. *EMBO J.*, 29, 1738–1747.

47. 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.

48. 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.

49. Srivastava, N.C., Shaky, N., Mak, M., Agrawal, B., Tyrrell, D.L. and Kumar, R. (2010) Antiviral activity of various 1-(2'-deoxy-beta-D-lyxofuranosyl), 1-(2'-fluoro-beta-D-xylofuranosyl), 1-(3'-fluoro-beta-D-arabinofuranosyl), and 2'-fluoro-2',3'-dideoxy-2',3'-dideoxyribose pyrimidine nucleoside analogues against duck hepatitis B virus (DHBV) and human hepatitis B virus (HBV) replication. *J. Med. Chem.*, 53, 7156–7166.

50. Chu, C.K., Schinazi, R.F., Ahn, M.K., Ulass, G.V. and Gu, Z.P. (1989) Structure-activity relationships of pyrimidine nucleosides as antiviral agents for human immunodeficiency virus type 1 in peripheral blood mononuclear cells. *J. Med. Chem.*, 32, 612–617.

51. Garrido-Hernandez, H., Moon, K.D., Geahlen, R.L. and Borch, R.F. (2006) Design and synthesis of phosphotyrosine peptidomimetic prodrugs. *J. Med. Chem.*, 49, 3368–3376.

52. Tobias, S.C. and Borch, R.F. (2001) Synthesis and biological studies of novel nucleoside phosphoramidate prodrugs. *J. Med. Chem.*, 44, 4475–4480.