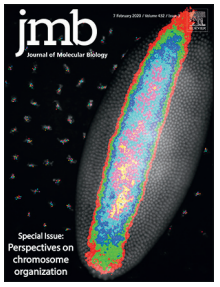
Communication



Enzymatic Hydroxylation and Excision of Extended 5-Methylcytosine Analogues

Miglė Tomkuvienė [1], Diana Ikasalaitė [1], Anton Slyvka [2], Audronė Rukšėnaitė [1], Mirunalini Ravichandran [3], Tomasz P. Jurkowski [4], Matthias Bochtler [2,5] and Saulius Klimašauskas [1]*

1 - ***Institute of Biotechnology,*** *Life Sciences Center, Vilnius University, Vilnius LT-10257, Lithuania*
2 - ***International Institute of Molecular and Cell Biology,*** *02-109 Warsaw, Poland*
3 - ***School of Medicine,*** *University of California, San Francisco, CA 94143, USA*
4 - ***School of Biosciences,*** *Cardiff University, CF10 3AX Cardiff, Wales, UK*
5 - ***Polish Academy of Sciences,*** *Institute of Biochemistry and Biophysics, 02-106 Warsaw, Poland*

Correspondence to Saulius Klimašauskas: *saulius.klimasauskas@bti.vu.lt (S. Klimašauskas)*
https://doi.org/10.1016/j.jmb.2020.10.011
Edited by Sepideh Khorasanizadeh

Abstract

Methylation of cytosine to 5-methylcytosine (mC) is a prevalent reversible epigenetic mark in vertebrates established by DNA methyltransferases (MTases); the methylation mark can be actively erased *via* a multistep demethylation mechanism involving oxidation by Ten-eleven translocation (TET) enzyme family dioxygenases, excision of the latter oxidation products by thymine DNA (TDG) or Nei-like 1 (NEIL1) glycosylases followed by base excision repair to restore the unmodified state. Here we probed the activity of the mouse TET1 (mTET1) and *Naegleria gruberi* TET (nTET) oxygenases with DNA substrates containing extended derivatives of the 5-methylcytosine carrying linear carbon chains and adjacent unsaturated C—C bonds. We found that the nTET and mTET1 enzymes were active on modified mC residues in single-stranded and double-stranded DNA *in vitro*, while the extent of the reactions diminished with the size of the extended group. Iterative rounds of nTET hydroxylations of ssDNA proceeded with high stereo specificity and included not only the natural alpha position but also the adjoining carbon atom in the extended side chain. The regioselectivity of hydroxylation was broken when the reactive carbon was adjoined with an sp^1 or sp^2 system. We also found that NEIL1 but not TDG was active with bulky TET-oxidation products. These findings provide important insights into the mechanism of these biologically important enzymatic reactions.
© 2020 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

0022-2836/© 2020 The Author(s). Published by Elsevier Ltd.This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Introduction

Methylation of cytosine at the 5th position is a well-characterized natural DNA modification. The resulting 5-methylcytosine (mC)‡ is found in DNA from all kingdoms of organisms, ranging from viruses to humans. mC can be regarded as an additional layer of information in a genome with its own writers, readers and erasers.[1] In bacteria, it is mostly employed by restriction-modification systems, while some evidence shows also participation in regulation of gene expression.[2] In Eukarya, from fungi to plants and animals, it is a widely accepted epigenetic mark enabling precise regulation of genomic expression, which is crucially important for differentiation, development and health of an organism.[3]

The methyl group is introduced in a post-replicative manner by enzymes DNA cytosine-5 methyltransferases.[4] In mammals, mC is found mostly in a CG dinucleotide context and is primarily associated with transcriptional silencing.[5] Being a reversible epigenetic mark, mC is subject to active erasure by the ten-eleven translocation (TET) dioxygenases. These enzymes catalyze successive rounds of oxidation of the 5-methyl group to yield 5-hydroxymethylcytosine (hmC), 5-formylcytosine (fC) and 5-carboxylcytosine (caC), in an α-ketoglutarate (α-KG) and Fe(II)-dependent manner.[6–9] Mammals have three TET paralogues. TET3 has the earliest role, it contributes to zygote demethylation.[10] TET1 and TET2 are important for demethylation in embryonic stem cells.[11] In the adult, TET2 plays a role for the hematopoietic system,[12] and TET1 is primarily responsible for the high hmC levels in the brain.[13] Representatives of this enzyme family are found in metazoa and at least some protists.[14] A smaller TET paralogue from *Naegleria gruberi*, nTET, has proven a robust model system for structural and mechanistic studies of the TET enzymes and found applications for analysis of epigenetic DNA modifications.[15,16]

Besides being intermediates in replication-dependent[17] and independent DNA demethylation, hmC/fC/caC could also serve as epigenetic marks in their own right[18–20]; they are chemically stable under physiological conditions, thus their active removal requires dedicated enzymatic action. At this stage, thymine DNA glycosylase (TDG) plays a key role by excising fC and caC and initiating base excision repair (BER).[21–23] NEIL1 glycosylase participates in the demethylation pathway by facilitating TDG substrate turnover,[24] can excise caC directly,[25] and can partially rescue the loss of TDG in TET-oxidation dependant reactivation of epigenetically silenced genes.[26] Besides the demethylation pathway, these glycosylases are also responsible for the excision of a wide spectrum of other modified bases, which may occur due to chemical damage such as deamination, oxidation or alkylation of DNA bases.[27–31]

In spite of their significance, mechanistic studies of the TET enzymes have remained limited. Based on evidence from other homologues in their class and available crystal structures of nTET and human TET2,[32,33] the TET-mediated C—H hydroxylation has been proposed to proceed through a radical pathway, whereby abstraction of a H atom by reactive Fe(IV)-oxo intermediate is followed by binding of a generated OH radical. mTET2 has recently been interrogated using larger chemical groups replacing the methyl group of mC,[34,35] which hinted at the importance of conjugative stabilization of the intermediate radical potentially defining the transition-state barrier for the proposed C—H activation.

Here we sought to query the catalytic plasticity of nTET and mTET1 oxygenases in the context of varied stereoelectronic environment of the transition state. For this we prepared a series of modified DNA substrates in which the 5-methyl group of the target mC residue is appended with a linear side chain that brings distinct electronic environments (sp^1, sp^2 or sp^3 carbon) and distinct numbers of C—H bonds available for hydroxylation (Figure 1(a)). We found that the activity of the TET enzymes is exerted not only at the alpha methylene group but also at other positions in the extended side chain. We also observed base excision activity of the bulky TET-oxidized derivatives by NEIL1 glycosylase, but not by TDG.

Results

The mCG dinucleotides are the prevalent targets of the TET proteins in mammalian DNA.[32] A series of modified oligonucleotide substrates containing extended moieties at the 5-position in a unique target mC residue were prepared using the chemoenzymatic mTAG approach.[36] The approach makes use of a DNA methyltransferase and synthetic cofactor AdoMet analogues carrying sulfonium-bound corresponding transferable groups.[37] In particular, we used an engineered version of the M. HhaI cytosine-5 MTase,[38] which modifies the inner cytosine (underlined) in the G<u>C</u>GC sequence, for preparing a set of oligonucleotide substrates in which one strand contained a modified residue

‡ Abbreviations used: mC, 5-methylcytosine; MTase, methyltransferase; TET, Ten-eleven translocation (dioxygenase); TDG, thymine DNA glycosylase; NEIL1, Nei-like 1 (glycosylase); mTET1, mouse TET 1; nTET, *Naegleria gruberi* TET; hmC, 5-hydroxymethylcytosine; fC, 5-formylcytosine; caC, 5-carboxylcytosine; α-KG, α-ketoglutarate; BER, base excision repair; xC, extended 5-methylcytosine used in this study; eC, 5-ethylcytosine; peC, 5-(prop-2-enyl)cytosine; byC, 5-(but-2-ynyl)cytosine; hyC, 5-(6-azidohex-2-ynyl)cytosine; ss, single-stranded (DNA substrate); ds, double-stranded (DNA substrate); HPLC-ESI/MS, high performance liquid chromatography electrospray ionization mass spectrometry; EIC, Extracted-ion chromatogram; eM. HhaI, engineered version of the DNA methyltransferase HhaI; MRM, Multiple reaction monitoring; TDGcd, thymine DNA glycosylase catalytic domain; TET1 KRAK, mouse TET1 catalytic domain; mTAG, methyltransferase-directed Transfer of Activated Groups.

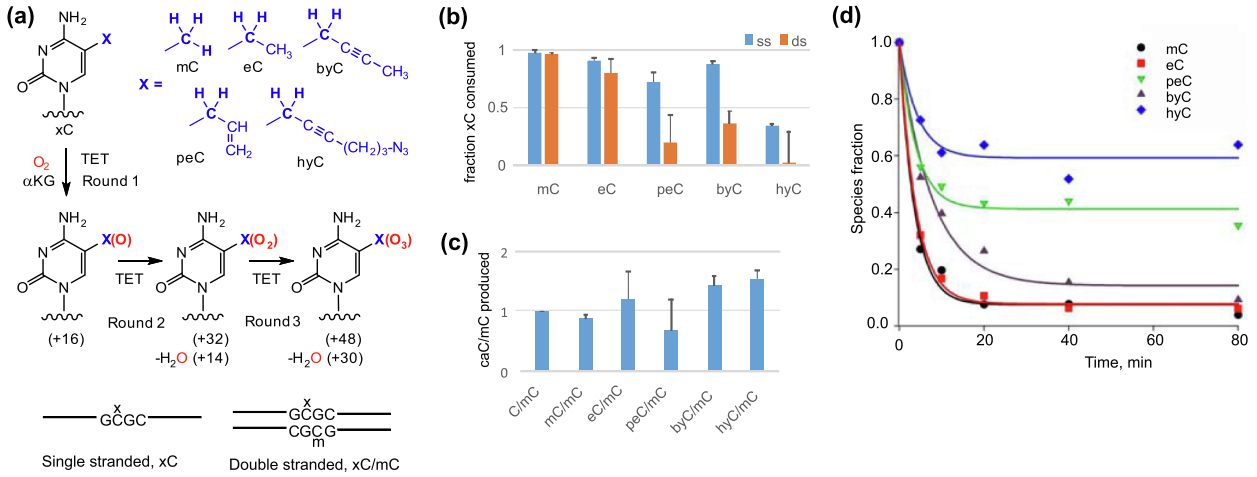


Figure 1. Oxidation of extended 5-methylcytosines (xC) by TET oxygenases. (a) Schematic representation of the oxidation reactions and oligonucleotide substrates described in this study. mC, 5-methylcytosine; eC, 5-ethylcytosine; peC, 5-(prop-2-enyl)cytosine; byC, 5-(but-2-ynyl)cytosine; hyC, 5-(6-azidohex-2-ynyl)cytosine. (b) Activity of nTET on xC in single-stranded (ss) and double-stranded (ds) substrates. (c) Oxidation of mC to caC by nTET in ds substrates carrying extended methylcytosines, represented as amounts of caC produced per mC, and normalized relative to hemimethylated duplex C/mC. (d) Kinetic traces of xC consumption by nTET activity in ssDNA substrates. DNA substrates were incubated with nTET and corresponding 2′-deoxyribonucleosides quantified by HPLC-MS.

(xC) while the complementary strand had synthetically incorporated mC (see Figure S1).

DNA strand selection by the TET enzymes

nTET and mTET1 have been shown to be active on both single-stranded and double-stranded DNA.[15] To assess the activity of nTET and mTET1 towards modified methylcytosines in double-stranded and single-stranded DNA we used a set of oligonucleotide substrates in which one strand contained a varied residue (C, mC or xC) at a unique CG site and the optional second strand contained mC (Figure 1(a)). Using HPLC-ESI/MS analysis of constituent 2′-deoxyribonucleosides we found that the consumption of mC and eC by nTET was comparable in both substrates, although the reactivity of the bulky unsaturated derivatives in double-stranded substrates was considerably lower as compared to that in single-stranded DNA (Figure 1(b)). Given that hyC was much less reactive than byC, one can presume that the size of the side chain rather than electronic factors was largely responsible for the observed effect.

In consideration of a general inhibitory effect of the bulkier groups, we also evaluated the oxidation of mC located on the opposite strand of the ds substrate. Remarkably, we observed that comparable amounts of the final oxidation product caC were produced in all cases (Figure 1(c)). This clearly indicates that bulky modifications on one strand of the CG site do not preclude nTET reaction on the other strand. Moreover, the best conversion of mC (Figure 1(c)) and the lowest conversion of xC (Figure 1(b)) observed with hyC/mC as compared to the other duplexes suggests that a bulky hydrophobic group in fact promotes the reaction on the opposite strand. Monomeric TET enzymes bind DNA duplexes targeting one of the two strands at a time, and thus the strand preference in asymmetrically modified duplexes can in part be dictated by acceptance of a target residue in the active site. The enhanced mC oxidation observed for the hyC/mC duplex most likely derives from impaired binding of the bulky 6-azidohexynyl group in the active site of the enzyme. This notion is consistent with crystallographic evidence suggesting that a bulky group can well be accommodated on the non-substrate C residue of the CG site.[33]

Although previous studies showed nTET to be similarly active on mC residues residing on either single-stranded and double-stranded DNA,[15] we found a much more favorable processing of extended groups in the ssDNA substrates. From the structural standpoint the contributing factors for this observation could be (i) the absence of competing binding to the opposite strand, and (ii) a lower energy barrier required for target base flipping in a ssDNA *versus* dsDNA.

TET-directed oxidation of extended 5-methylcytosine in ssDNA

During initial screening we observed a moderate mTET1 activity on eC and byC (Figure S2), whereas nTET was considerably more proficient with all of the examined substrates. Given a strongly impaired TET oxidation of the 5-modified cytosines (Figure 1(b)) in dsDNA, further experimental studies of the TET reactions were performed on the ssDNA substrates. Using

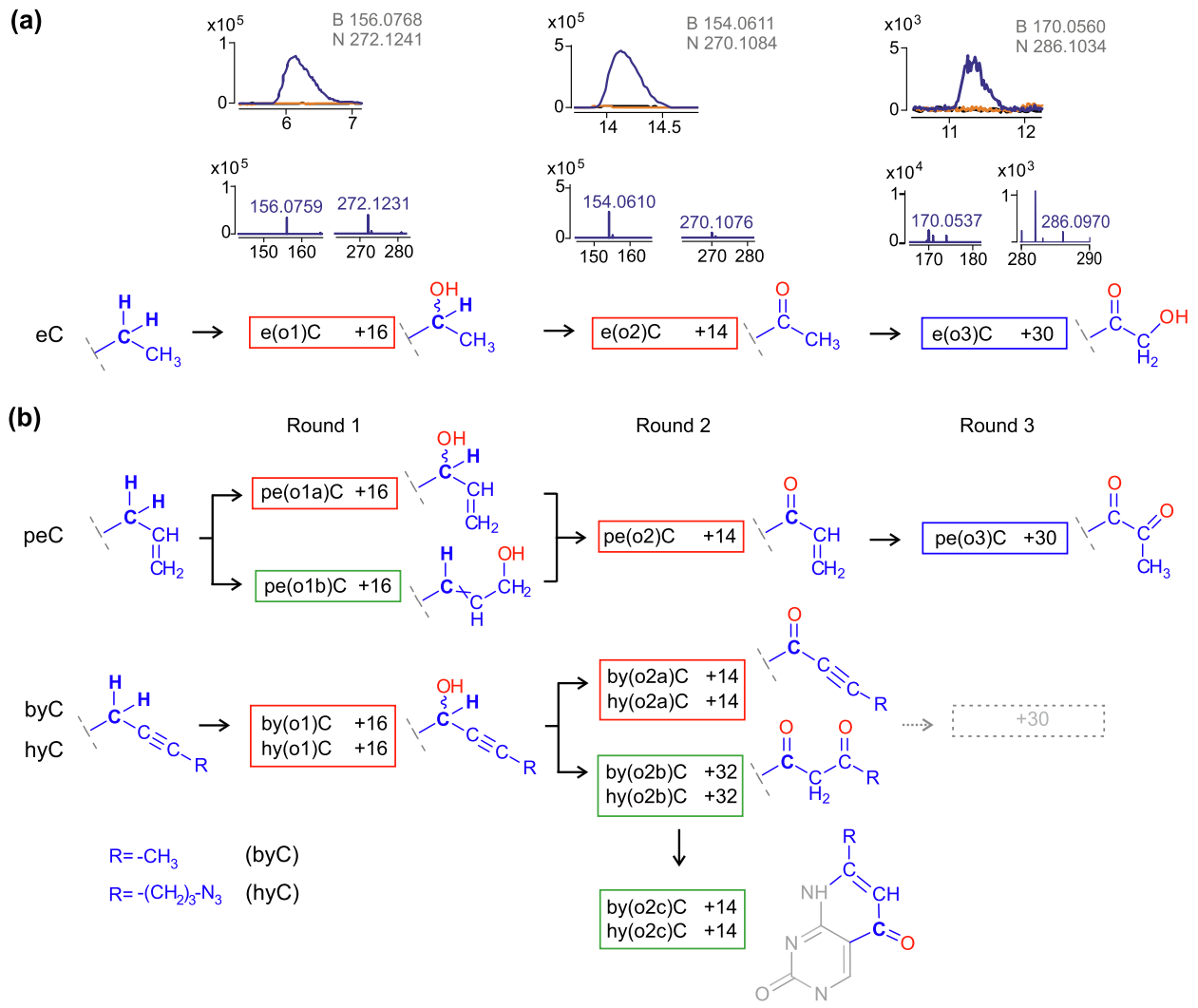


Figure 2. HPLC-MS analysis of nTET-dependent oxidation of extended mC derivatives in single-stranded DNA. (a) Representative analysis of eC DNA. 25-mer oligodeoxynucleotide substrates were incubated with nTET, digested to nucleosides and analyzed by HPLC-MS. Extracted-ion chromatograms (upper, blue trace) are shown along with observed ionic (H+) mass spectra (lower) of both the base (B) and nucleoside (N); respective theoretical mass values are shown in grey. Control samples contained a catalytic nTET mutant (orange) or reaction quenched at start (black). (b) Schematic representation of HPLC-MS analysis results for peC, byC and hyC substrates. Open boxes denote the presence of corresponding reaction products (see Figures S5 and S6 for details) where presumed alpha, beta and gamma hydroxylation products shown in red, blue and green, respectively; light dashed boxes denote scarcely abundant species detectable using HPLC-MS/MS analysis (see Figure S8).

HPLC-MS analysis of constituent 2′-deoxynucleosides, we observed nTET-dependent oxidation of the alkylated cytosines to yield a series of oxidation products (Figure 2) as documented for control reactions involving mC (Figure S2). These observed mass shifts correspond to addition of an O atom (+16, Round 1), addition of a second O atom (+32) followed by dehydration (+14) (Round 2) and addition of a third O atom (+30, Round 3) (Figures 1(a), 2 and S3). Kinetic analyses (Figure S4) showed that the corresponding oxidation products (+16, +14 and +30) derived from mC or eC appeared (achieved peak concentration) in the same temporal order confirming that in both cases identical changes in chemical composition occurred at each round of the oxidation pathway.

All the oxidation products were nTET-dependent as the respective reactions with a catalytically impaired mutant (AKA)[33] rendered no similar products. A time course of the first oxidation round (measured by disappearance of the starting modification) (Figure 1(d)) showed that the initial reaction rates were rather similar for most substrates, although the extent of oxidation showed a general trend for reduced enzymatic activity as the alkyl

groups got longer. Although the oxidation of peC and hyC was rather fast during initial 10 min, a plateau was reached with a substantial fraction of the substrates left unconsumed. One explanation for this observation could be that a fraction of the substrates (xCs conformers) formed a productive and rapidly processed complex, whereas a larger fraction was stuck in an unproductive complex precluding its rebinding and turnover during the lifetime of the experiment.

TET-catalyzed oxidation of 5-ethylcytosine is regio- and stereoselective

Although the oxidation of mC can only occur at the α-carbon, the position of enzymatic hydroxylation in extended sidechains needs not be the same. Using a positionally deuterated eC substrate, Kavoosi *et al.* recently showed that the first round of TET2-catalyzed oxidation of eC occurs at the expected alpha position.[34] To determine the specificity of the nTET-directed oxidation of eC in ssDNA we took advantage of an atypical reaction of DNA cytosine-5-methyltransferases to couple exogenous aliphatic aldehydes at the 5-position of their target cytosine residues in DNA, yielding corresponding 5-α-hydroxyalkylcytosines.[39] The reaction of M.SssI and acetaldehyde in DNA produces diastereoisomeric 5-α-hydroxyethyl-dCyd, which proved a useful standard for our studies of the TET reactions. HPLC-MS analysis of the TET–treated eC DNA showed that both mTET1 and nTET generated the same one +16 Da product co-eluting with one of the standard diastereomers generated in the methyltransferase reaction (Figures 3 and S2(b)). In agreement with the previous findings, this observation suggests that the oxidation of eC occurs at the Cα atom. Furthermore, it suggests that the Round 1 enzymatic oxidation is stereo-selective yielding only one enantiomer at the new asymmetric centre. Alternatively, there was a theoretical possibility that the Round 1 reaction was not stereospecific, and the next step was - and what we observed was the remaining (unreactive) isomer. However, this possibility proved very unlikely due to the absence of the other diastereomeric (or any other) +16 species during initial stages of the reaction, even when no Round 2 or Round 3 product was yet detectable.

Regioselectivity of hydroxylation is broken in unsaturated side chains

In consideration of possible reaction products involving longer side chains, our general assumption was that the enzyme would preferentially abstract an H atom from the cognate α position, if one is available, as observed for eC. In the case of peC, two nTET-dependent +16 products were detected using HPLC-MS. The two isomers are unlikely to be diastereomers as they showed quite distinct HPLC mobilities. Due to conjugation with the adjoining sp^2 system, the unpaired electron of the intermediate allyl radical is expected to delocalize between the α and γ carbons, and thus could yield two distinct hydroxylation products (pe(o1a)C and pe(o1b)C in Figures 2(b) and S5). Although, theoretically, two isomers are possible for each α and γ hydroxylated sidechains, two species altogether were observed in line with the notion that the reaction is stereo specific. Further oxidation in Round 2 is predicted to ultimately yield a single +14 product (pe(o2)C), as observed (Figures 2(b), S3 and S5). This observation also excludes scenarios that envision hydroxylation of two distinct positions in the side chain (i.e. α and β), which would give a +32 oxidation species.

The TET reactivity of byC and hyC was not identical but largely yielded similar products. In both cases, only one nTET-dependent +16 product was detected using HPLC-MS (see by (o1)C and hy(o1)C in Figures 2(b), S3, S5). Although conjugation with the sp^1 system

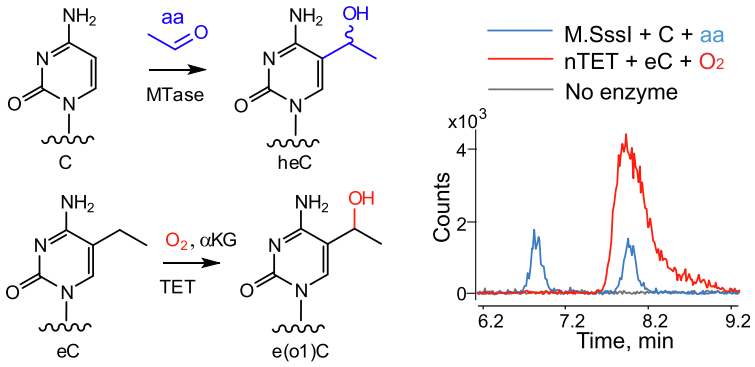


Figure 3. Stereo- and regio selectivity of 5-ethylcytosine (eC) oxidation by nTET. Diastereomeric 5-(α-hydroxyethyl)-2′-deoxycytidine (blue trace) produced *via* M.SssI-dependent α-hydroxyethylation of C in DNA with acetaldehyde (aa) was compared to the product (e(o1)C, mass shift +16) of eC oxidation by nTET (red trace). After DNA digestion to nucleosides and HPLC-MS analysis, EIC signals (156.0768) from both reactions and no enzyme control (black trace) are shown.

potentially creates a propargyl radical with two resonance positions of the unpaired electron, no products corresponding to a second radical were evident in HPLC-MS or HPLC-MS/MS chromatograms. In Round 2, two +14 products were detectable for both byC and hyC (a second peak much less evident for the latter), one of which co-eluted with a broad +32 peak. A +2O species could appear if a Round 2 hydroxylation occurred at another than Cα position of the extended group. For example, an α ↔ γ radical migration *via* propargyl-allenyl rearrangement (see Figure S5) would yield a 1,3-dioxo compound (by(o2b)C, hy(o2b)C). The latter is likely to be in equilibrium with a cyclic structure, which would be prone to further dehydration into an aromatic ring. Therefore, the co-eluting +14 peak (by(o2c)C, hy(o2c)C) is likely a fragment (dehydration product −18) of +32 that is readily formed in solution or appears in the gas phase during MS analysis. Since all the extended alkyl groups contain only two Hα atoms, the formation of a +30 species should occur in Round 3 *via* deposition of an O atom at another (Cβ or further) position in the side chain! Therefore we presume that, for eC and likely peC, the H abstraction and subsequent hydroxylation in the α-oxo derivative (+14) occurs at the Cβ position (Figures 2(b) and S5). An even more peculiar example is presented with byC and hyC, which contain no Hβ or Hγ atoms in their side chains. We detected no +30 or +48 species using HPLC-ESI/MS (Q-TOF) analysis, although a scarcely abundant +30 signals appeared in a highly sensitive MS/MS MRM experiment (Figure S6). We have difficulty predicting how such species could originate from the α-oxo derivative by(o2a)C, hy(o2a)C, which contains no Hß or Hγ atoms; alternative candidate precursors would be by(o2b)C, hy(o2b)C, by(o2c)C, hy(o2c)C (see Figures 2(b) and S5).

TDG and NEIL1 activity on 5-alkylcytosines and their oxidation products

To further explore the complex interplay between the multiple functional roles of the TET oxygenases and DNA glycosylases, we finally tested if the extended mC modifications and their oxidation products generated by nTET are excised by the mammalian TDG and NEIL1 glycosylases. For this, the nTET reaction products obtained with ssDNA substrates (see above) were converted to duplexes by annealing with an unmodified complementary strand. These dsDNA substrates were then interrogated with NEIL1 or TDG and analyzed by PAGE or digested to nucleosides for HPLC-MS/MS analysis (Figure 4(a)). Using 5′-P-labeled 25-mer hemialkylated duplexes (xC/C), we found no base excision activity with the original alkylated substrates. After treatment with nTET, a moderate, yet clearly detectable activity of NEIL1 glycosylase on both strands containing triple C—C bonds in the extended side chain (byC and hyC) (Figures 4(b) and S7) was observed. HPLC-MS/MS chromatograms of the same duplexes, showed disappearance of co-eluting +32 and +14 species after the NEIL1 treatment (Figures 4(c) and S8). This means that although several oxidation products were present after nTET treatment, only one of the Round 2 oxidation products was a substrate of NEIL1 (which also explains the observed partial strand cleavage activity in Figure 4(b)). In light of the documented capacity of NEIL1 to excise bulky substrates,[28] the presumed bicyclic structures of these compounds (by(o2b,c)C and hy(o2b,c)C) may be a distinctive factor conferring susceptibility to the NEIL1 activity.

Discussion

5-alkylcytosines with aliphatic carbon chains (ethyl-, propenyl-, butynyl- or azidohexynyl-) can undergo (partial) oxidation by nTET and mTET1 oxygenases *in vitro* to corresponding alpha-hydroxyl- and oxo- compounds. The efficiency of the reaction inversely correlates with the length and bulkiness of the alkylgroup. The enzymes most probably have a lower acceptance of bulky alkylcytosines in the active site, as our experiments show preferential reactivity toward mC over alkylcytosines in substrates containing both modifications on different strands of the same CG site. Structural perturbations in the modified duplex substrates are less likely to account for the observed differences in enzymatic oxidation rates since cytosine-5 substitutions point to the major groove of B-DNA, and even large non-polar moieties exert minor effects on duplex stability as compared to 5-methylcytosine.[40] This would imply that the presence of a bulkier group (such as alkyl or carboxyl) on one strand could enhance the nTET oxidation of mC on the opposite strand (mC/caC duplex) as compared to that in a fully methylated or hemimethylated CG site (mC/mC or mC/C duplexes). Furthermore, the activity of nTET appeared to be even higher with ssDNA as compared to analogous duplexes (Figure 1), suggesting that it may preferentially target ss-regions of the genome such as replication forks or R-loops.

On mC as a substrate, the TET oxygenases can perform three rounds of oxidation, one for each H atom present at the α carbon. This is the first study that looked at full three rounds of TET-dependent oxidation of extended mC probes to yield a series of hydroxylated products. We obtained evidence of a high stereoselectivity for eC and other Cα oxidation reactions in Round 1 for both the mouse and *Naegleria* enzymes. From the structural point, this implies that either the orientation of the ß-carbon is strictly defined in the active center of both TET enzymes, or, alternatively, two or more binding modes are possible but only one of them is catalytically competent. The latter possibility could

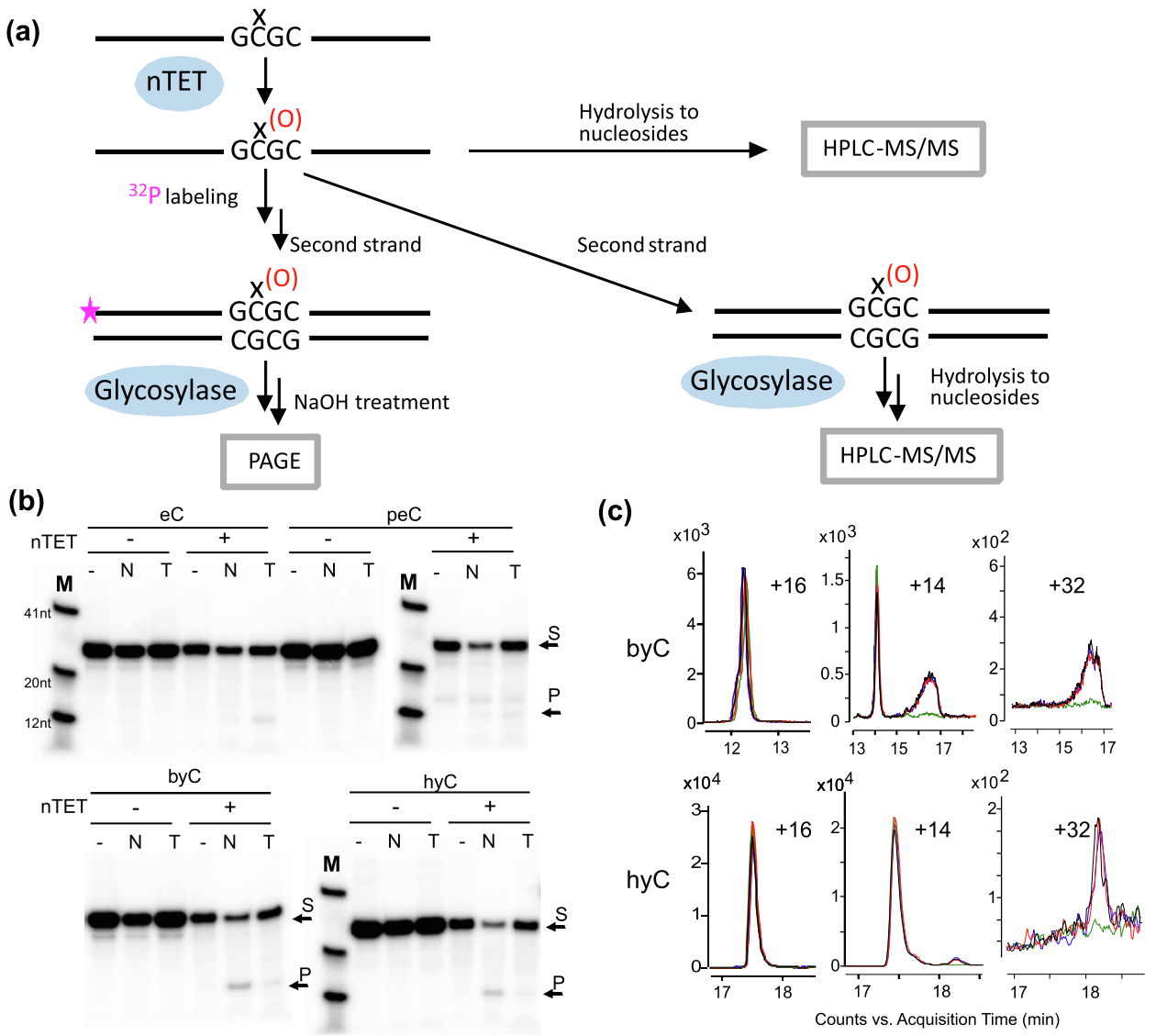


Figure 4. Activity of the TDG and NEIL1 glycosylases on extended 5-methylcytosines and their TET oxidation products. (a) Single stranded xC oligonucleotides were preincubated with nTET, and the produced x(O)C strands were annealed to a complementary unmodified oligonucleotide. The resulting duplexes were incubated with NEIL1 or TDG and analyzed using PAGE or HPLC-MS/MS. (b) Electrophoretic analysis of the glycosylase activity on oxidised 5-alkylcytosines. 5′-P-labeled 25-mer hemialkylated dsDNA substrates (S) obtained after treatment of xC strand with nTET and annealing of a complementary unmodified oligonucleotide were incubated with TDG (T) or NEIL1 (N) for 3 h, followed by treatment with NaOH to confer strand cleavage at excision sites (P). See Figure S7 for activity control reactions. (c) HPLC-MS/MS analysis of NEIL1 glycosylase activity on TET-oxidized byC and hyC. 25-mer hemialkylated DNA duplexes obtained after treatment of xC strand with nTET and annealing of a complementary unmodified oligonucleotide were incubated with NEIL1 followed by HPLC-MS/MS of constituent nucleosides. Shown are MRM traces corresponding to +16, +14 and +32 oxidation species for wt NEIL1 (green), catalytic mutant variants P2T (red), E3Q (blue) and no enzyme (grey).

explain the formation of stable unproductive complexes leading to incomplete consumption of the substrate observed in kinetic experiments (Figure 1(d)).

Here we show for the first time that after all α C—H bonds are consumed, nTET affords deposition of O atoms at other (ß or further) positions in the side chain in Round 3 (Figure 2). An open question remains whether the H abstraction by the active Fe(IV) centre occurs directly at the adjacent C—H bond of the extended residue or involves a radical migration *via* an intermediatory position in the enzyme? We also unexpectedly found that the regio-selectivity

of the reaction is broken when the reactive carbon is adjoined with an unsaturated system. This affords the conjugation of the unpaired orbital with the flanking π-electron system and radical delocalization between α and γ positions thereby yielding two hydroxylation products. Altogether, this for the first time demonstrates that the selectivity of the TET hydroxylation toward the α-position is not absolute and can also involve ß- or γ-carbons in extended substrates. Different consequences from steric and electronic effects suggest that the regio-specificity of this enzymatic reaction can be controlled by varying the chemical environment within adjoining moieties.

Lately, TET dioxygenases have found roles as molecular tools for manipulations and analysis of epigenetic modifications in DNA.[15,34,35,41] The current work extends the repertoire of extended mC analogues by exploiting the promiscuity of TETs towards these bases. Another natural way to expand the variety of modified cytosine bases has recently been discovered in protists.[42] CMD1, a TET-related enzyme, which uses vitamin C as a co-substrate, resolves the nucleobase radical not by hydroxyl insertion, but by glyceryl transfer. It will be interesting to see whether yet other nucleobase variants can be made by CMD1 action on the bulky mC analogues in this work.

mC oxidation products fC and caC are efficiently excised by TDG.[21,22] Modest TDG activity on a perruthenate oxidation product of 5-(α-hydroxyethyl)C (heC), presumably 5-acetylC or e(o2)C, has recently been reported.[34] Consistent with this comes a faint TDG excision band detectable upon treatment of the aggregate eC oxidation products (Figure 4(b)). NEIL1 is a repair enzyme involved in removing oxidative DNA lesions.[43] NEIL1 can cooperate with TDG, but on its own has a much lower caC glycosylase activity than TDG.[25] Notably, we were able to detect excision of TET-oxidized 5-alkylcytosines byC and hyC by NEIL1, but not of byC and hyC themselves (Figures 4(b), S7 and S8). The higher activity of NEIL1 (compared to TDG) towards the sterically demanding oxidized species appears in agreement with its enlarged catalytic site and reported ability to process substrates that are unusually bulky for the base excision repair pathway.[44]

In the context of their general suitability for applications in chemical and synthetic biology, the extended mC derivatives largely avoided removal by the TET-TDG-NEIL1 system as compared to the cognate 5-methylcytosine. This suggests that such artificial modifications installed in DNA would likely remain functionally stable inside mammalian cells. This paves the way to a range of strategies for direct chemical probing of or physical reporting on epigenetic events that occur during homeostasis, development and disease.

Experimental

DNA substrates

DNA oligonucleotides containing one CG site modified at the 5th position with ethyl-, prop-2-enyl-, but-2-ynyl- or 6-azidohex-2-ynyl groups were prepared (see Figure S1) using an engineered version of the DNA methyltransferase HhaI (eM.HhaI)[38] and corresponding synthetic cofactor analogs.[37,45] Briefly, complementary oligonucleotides (the HhaI recognition sequence is underlined and the target cytosine in bold) "upper" GTTGTGGGTCAG**C**GCCTGATACTGT and "lower" TACAGTATCAGG[mC]GCTGACCCACAA were annealed at a ratio 1:1.1. The double-stranded substrate in concentration 25 μM was incubated with 2.5 μM eM.HhaI and 300 μM cofactor in TE buffer (50 mM Tris-HCl pH 7.4, 5 mM EDTA) at 37 °C overnight. In the case of ethylation both the substrate and enzyme concentrations were 11 μM, while AdoEth cofactor was 600 μM. Further, the enzyme was inactivated by heating at 65 °C for 20 min and the DNA was column-purified (Zymo Research). For preparing of single-stranded modified oligonucleotides, the lower (methylated) strand also contained 5'-biotin, which enabled subsequent strand separation and purification by HPLC.

Oligonucleotides were then separated and analyzed on an integrated HPLC/ESI-MS system (Agilent 1290 Infinity) equipped with a Zorbax Eclipse Plus C18 column (15 cm × 2. 1 mm) by elution with a linear gradient of solvents A (5 mM ammonium acetate pH 7.0 in water) and B (5 mM ammonium acetate pH 7.0 in methanol) at a flow of 0.15 ml/min at 55 °C as follows: 0–10 min, 5–14.5% B; 10–18 min, 14.5% B; 18–22 min, 14.5–100% B. High-resolution mass spectra of analytes were acquired on an Agilent Q-TOF 6520 mass analyzer (100–3200 m/z range, negative ionization mode) and analyzed with Agilent MassHunter Qualitative Analysis software. HPLC-MS analysis of the resulting double or single stranded DNA oligonucleotides confirmed a nearly quantitative modification of the upper strand with an ethyl-, propen-, butyn and hexynazide- moiety, respectively.

Protein production and purification

His-tagged mTET1 CD KRAK protein and nTET wildtype and AKA variant (H249A/D251A double mutant) proteins were expressed in *Escherichia coli* BL21 (DE3) CodonPlus RIL cells and purified as previously described.[33,46]

pET28a (+) vector containing the synthetic gene encoding His6-nTET wild type, AKA catalytic mutant or His6-mTET1 CD KRAK were transformed into *E. coli* BL21 (DE3) CodonPlus RIL (Novagen). Single colonies were selected and inoculated into 20 ml of Luria Bertani (LB) media containing the antibiotics kanamycin (25 μg/ml) and grown at 37 °C with constant shaking over-night. Afterwards, the pre-culture was transferred into fresh flask containing one litre of LB media (1:1000 dilution) with kanamycin (25 μg/ml) and 1X trace metals solution. The cultures were grown at 37 °C until reached an OD_{600} of 0.6–0.8, subsequently transferred to 28 °C and the protein expression was induced by adding isopropyl ß-D-1-thiogalactpyranoside (IPTG) at a final concentration of 0.5 mM. The cells were further grown at 28 °C for 3 h. Afterwards, the cells were harvested by centrifugation (Sorvall Lynx 6000, Thermo Scientific) and washed once with STE buffer containing 10 mM Tris-HCl pH 8.0, 1 mM EDTA and 100 mM NaCl. The harvested cells were stored at −20 °C.

For purification, the cells containing His-tagged mTET, nTET wild-type and nTET AKA catalytic mutant proteins were thawed on ice and suspended in sonication buffer (50 mM HEPES pH 6.8, 35 mM imidazole, 1 mM DTT, 500 mM NaCl, 10% glycerol) freshly supplemented with protease inhibitor cocktail (1 mM of AEBSF, 0.8 μM of aprotinin, 50 μM of bestatin, 15 μM of E-64, 5 mM of EDTA, 20 μM of leupeptin, 10 μM of pepstatin A). The

cells were lysed with sonication (25–30 cycles, 40% power, 15 s ON and 45 s OFF, Bandelin Sonoplus sonicator) and the lysate was cleared by centrifugation at ~ 38,800x g for 75 minutes using Sorvall Lynx 6000 centrifuge (Thermo Scientific). The cleared lysate was loaded onto a pre-equilibrated column containing the Ni-NTA agarose beads (Genaxxon, Germany). Afterwards the beads containing the protein were washed with 150–200 ml of sonication buffer and eluted (50 mM HEPES pH 6.8, 300 mM imidazole, 1 mM DTT, 500 mM NaCl, 10% glycerol). Highly concentrated fractions of the eluted protein were pooled together and dialysed against dialysis buffer I for 3 hours. The dialysed protein was aliquoted, flash frozen in liquid nitrogen and stored at −80 °C.
eM.HhaI and glycosylases TDG and NEIL1 were prepared as previously described in Lukinavičius *et al.* and Daujotytė *et al.*[38,47] and Slyvka *et al.*[25], respectively.

Mammalian TET activity assay

Modified double-stranded DNA oligonucleotides (70 pmol, 2.3 μM) were incubated with murine TET1 KRAK (134 pmol, 4.5 μM), or *Naegleria's* TET (140 pmol, 2.3 μM). Synthetic oligonucleotides with mC on the lower or both strands served as controls. Reactions were incubated for 3 h for at 37 °C in appropriate buffers: 50 mM HEPES, pH 6.8 (for TET1 KRAK, TET3 CD) or 50 mM Bis-Tris pH, 6.0 (for nTET); 150 mM NaCl (for TET1 KRAK) or 100 mM NaCl (for nTET); 1 mM α-ketoglutarate; 1 mM ascorbic acid; 75 μM $Fe(NH_4)_2(SO_4)_2$ and stopped by adding EDTA to 16 mM final concentration. Reaction mixtures were further treated and analyzed by HPLC-MS as described below.

nTET activity assay

The modified double- or single-stranded DNA oligonucleotides (2.3 μM) were incubated with 4.6 μM nTET or inactive mutant nTET AKA in a reaction buffer: 50 mM Bis-Tris pH 6.0; 100 mM NaCl; 1 mM α-ketoglutarate; 1 mM ascorbic acid; 75 μM $Fe(NH_4)_2(SO_4)_2$ at 34 °C for 3 h. The reaction was stopped by adding EDTA to 16 mM final concentration. For the time course assay, 1.5 μM modified ssDNA was incubated with 3 μM nTET enzyme for the specified time and quenched by heating at 95 °C for 3 min. The reaction mixtures were then supplemented with 0.1% SDS, 3 mM $CaCl_2$ and incubated for 1 h at 55 °C with proteinase K (Thermo Fisher Scientific). Further, DNA was column-purified (Zymo Research) and digested with nuclease P1 (0.01 u/μl final concentration, Sigma) for 2 h at 55 °C and with addition of phosphatase FastAP (0.01 u/μl final concentration, Thermo Fisher Scientific) overnight at 37 °C.
The resulting nucleosides were analyzed on an integrated HPLC-ESI/MS system (Agilent 1290 Infinity) equipped with a Discovery HS C18 column (7.5 cm × 2.1 mm) by elution with a linear gradient of solvents A (0.0075% formic acid in water) and B (0.0075% formic acid in acetonitrile) at a flow of 0.3 ml/min at 30 °C as follows: 0–6 min, 0% B; 6–18 min, 0–10% B; 18–20 min, 10–100% B. High-resolution mass spectra were acquired on an Agilent Q-TOF 6520 mass analyzer (100–3200 m/z range, positive ionization mode) and analyzed with Agilent MassHunter Qualitative Analysis software. Chromatograms are shown as Extracted Ion Chromatogram (EIC). Assays for quantitative measurements were done in duplicate at least, and standard deviation is shown as error bars.

Glycosylase activity assay

Modified single-stranded DNA oligonucleotides were incubated with nTET for 1 h at 34 °C as described previously (see above) and the oxidized products were confirmed by HPLC-MS/MS (see below). The unmodified second strand was then annealed at a ratio 1:1.2 and the resulting double-stranded DNA was incubated for 1 h at 37 °C with a 20-fold molar excess of wild type human NEIL1 or its catalytically impaired variants NEIL-P2T and NEIL-E3Q (controls)[48] in a reaction buffer: 70 mM MOPS pH 7.5, 5% glycerol, 0.1 mM EDTA, 1 mM DTT. The reaction was stopped by heating at 95 °C for 5 min, and treated with proteinase K, DNA was column-purified (Zymo Research). For MS composition analysis, DNA was digested to nucleosides as described in the TET activity assay (see above). The resulting samples were analyzed on an integrated HPLC/ESI-MS/MS system (Agilent 1290 Infinity/ 6410 Triple Quad LC/MS) equipped with a Supelco Discovery®HS C18 column (7.5 cm × 2.1 mm, 3 μm) by elution with a linear gradient of solvents A (0.0075% formic acid in water) and B (0.0075% formic acid in acetonitrile) at a flow of 0.3 ml/min at 30 °C as follows: 0–5 min, 0% B; 5–15 min, 10% B; 15–20 min, 100% B. Mass spectrometer was operating in the positive ion mode at a capillary voltage of 1800 V, drying gas temperature 300 °C and flow rate 10 l/min. Multiple reaction monitoring (MRM) was used and nucleoside-specific ion transitions were recorded: byC 280.1 → 164.1, oxidation products of byC 294.1 → 178.1, 296.1 → 180.1, 310.1 → 194.1, 312.1 → 196.1; hyC 349.2 → 233.1, oxidation products of hyC 363.1 → 247.1, 365.2 → 249.1, 379.1 → 263.1, 381.2 → 265.1.
For PAGE analysis, nTET treated and untreated modified single-stranded DNA oligonucleotides were labelled at 5′-end with ^{32}P using T4 PNK (Thermo). As positive controls for NEIL1 and TDGcd (Figure S7), 41-mer DNA oligonucleotides containing a single 5-hydroxyl-2′-deoxycytidine or 5-formyl-2′-deoxycytidine residue (K) were used (5′-GCTACCTACCTAGCAGGGGK CAGCTGTCCCAC TGCTCGGAA-3′). Labelled oligonucleotides were then annealed to the unmodified unlabeled complementary strand in 1:1.2 ratio. Resulting duplexes were purified using Mini Quick Spin DNA columns (Roche). 10 nM of each duplex were mixed with 200 nM of NEIL1 (wild-type) or TDG catalytic domain (wild-type) in the reaction buffer: 70 mM MOPS pH 7.5, 5% glycerol, 1 mM EDTA, 1 mM DTT. The reaction mixtures were incubated for 3 hours at 37 °C and subsequently treated with 0.1 M NaOH for 5 min at 95 °C and neutralized with 0.1 M HCl. Samples were then supplemented with an equal volume of 2x loading solution (96% formamide, 20 mM EDTA and 0.1% bromophenol blue) heated for 5 min at 95 °C and resolved on 20% PAGE (19:1 acrylamide: bis-acrylamide) that contained 8 M urea. Radioactively labelled DNA was detected using Typhoon Trio (GE Healthcare) phosphorimager and the obtained reads were processed using ImageQuant software (GE Healthcare).

Preparation of diastereomeric 5-α-hydroxyethyl-2′-deoxycytidine standard

The procedure was adapted from.[39] Briefly, unmethylated double stranded substrate (13 μM) was incubated with 13 μM M.SssI and 800 mM acetaldehyde in reaction buffer (10 mM Tris-HCl pH 7.4; 50 mM NaCl, 0.1 mg/mL BSA; 2 mM 2-mercaptoethanol) at 25 °C for 3 h. After the incubation DNA was column-purified (Zymo Research), digested to nucleosides and analyzed by HPLC-MS as described in the TET activity assay (see above).

DECLARATION OF COMPETING INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We thank Giedrė Urbanavičiūtė for eM.HhaI, Gra vydas Lukinavičius for the ethyl-, propenyl- and butynyl-cofactor analogues and Viktoras Masevičius for the azidohexynyl cofactor analogue and fruitful discussions. This work was supported by the European Research Council (grant ERC-2016-AdG/742654 to S. K), by the Foundation for Polish Science/EU Regional Development Fund (FNP, POIR.04.04.00-00-5D81/17-00 to M.B. and START 79.2020 to A.S.), by the Polish National Academic Exchange agency (NAWA, PPI/APM/2018/1/00034) and by the Deutsche Forschungsgemeinschaft (JU 2773-2 to T.P.J.).

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jmb.2020.10.011.

Received 24 June 2020;
Accepted 8 October 2020;
Available online 14 October 2020

Keywords:
DNA cytosine-5 methylation;
TET dioxygenase;
DNA glycosylase;
epigenetic regulation;
radical intermediate

References

1. Carell, T., Kurz, M.Q., Muller, M., Rossa, M., Spada, F., (2018). Non-canonical bases in the genome: the regulatory information layer in DNA. *Angew. Chem. Int. Ed. Engl.*, **57**, 4296–4312.
2. Casadesus, J., (2016). Bacterial DNA Methylation and Methylomes. *Adv. Exp. Med. Biol.*, **945**, 35–61.
3. Schubeler, D., (2015). Function and information content of DNA methylation. *Nature*, **517**, 321–326.
4. Jurkowska, R.Z., Jeltsch, A., (2016). Mechanisms and Biological Roles of DNA Methyltransferases and DNA Methylation: From Past Achievements to Future Challenges. *Adv. Exp. Med. Biol.*, **945**, 1–17.
5. Ravichandran, M., Jurkowska, R.Z., Jurkowski, T.P., (2018). Target specificity of mammalian DNA methylation and demethylation machinery. *Org. Biomol. Chem.*, **16**, 1419–1435.
6. Tahiliani, M., Koh, K.P., Shen, Y., Pastor, W.A., Bandukwala, H., Brudno, Y., Agarwal, S., Iyer, L.M., et al., (2009). Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science*, **324**, 930–935.
7. Ito, S., D'Alessio, A.C., Taranova, O.V., Hong, K., Sowers, L. C., Zhang, Y., (2010). Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature*, **466**, 1129–1133.
8. Ito, S., Shen, L., Dai, Q., Wu, S.C., Collins, L.B., Swenberg, J.A., He, C., Zhang, Y., (2011). Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science*, **333**, 1300–1303.
9. Kriukienė, E., Liutkevičiūtė, Z., Klimašauskas, S., (2012). 5-Hydroxymethylcytosine – the elusive epigenetic mark in mammalian DNA. *Chem. Soc. Rev.*, **41**, 6916–6930.
10. Gu, T.-P., Guo, F., Yang, H., Wu, H.-P., Xu, G.-F., Liu, W., Xie, Z.-G., Shi, L., et al., (2011). The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. *Nature*, **477**, 606–610.
11. Pastor, W.A., Aravind, L., Rao, A., (2013). TETonic shift: biological roles of TET proteins in DNA demethylation and transcription. *Nat. Rev. Mol. Cell Biol.*, **14**, 341–356.
12. Ko, M., Bandukwala, H.S., An, J., Lamperti, E.D., Thompson, E.C., Hastie, R., Tsangaratou, A., Rajewsky, K., et al., (2011). Ten-Eleven-Translocation 2 (TET2) negatively regulates homeostasis and differentiation of hematopoietic stem cells in mice. *Proc. Natl. Acad. Sci. USA*, **108**, 14566–14571.
13. Kaas, Garrett A., Zhong, C., Eason, Dawn E., Ross, Daniel L., Vachhani, Raj V., Ming, G.-L., King, Jennifer R., Song, H., et al., (2013). TET1 controls CNS 5-methylcytosine hydroxylation, active DNA demethylation, gene transcription, and memory formation. *Neuron*, **79**, 1086–1093.
14. Iyer, L.M., Tahiliani, M., Rao, A., Aravind, L., (2009). Prediction of novel families of enzymes involved in oxidative and other complex modifications of bases in nucleic acids. *Cell Cycle*, **8**, 1698–1710.
15. Pais, J.E., Dai, N., Tamanaha, E., Vaisvila, R., Fomenkov, A.I., Bitinaite, J., Sun, Z., Guan, S., et al., (2015). Biochemical characterization of a Naegleria TET-like oxygenase and its application in single molecule sequencing of 5-methylcytosine. *Proc. Natl. Acad. Sci. USA*, **112**, 4316–4321.
16. Hashimoto, H., Pais, J.E., Dai, N., Correa Jr, I.R., Zhang, X., Zheng, Y., Cheng, X., (2015). Structure of Naegleria Tet-like dioxygenase (NgTet1) in complexes with a reaction intermediate 5-hydroxymethylcytosine DNA. *Nucleic Acids Res.*, **43**, 10713–10721.
17. Hashimoto, H., Liu, Y., Upadhyay, A.K., Chang, Y., Howerton, S.B., Vertino, P.M., Zhang, X., Cheng, X., (2012). Recognition and potential mechanisms for replication and erasure of cytosine hydroxymethylation. *Nucleic Acids Res.*, **40**, 4841–4849.
18. Branco, M.R., Ficz, G., Reik, W., (2011). Uncovering the role of 5-hydroxymethylcytosine in the epigenome. *Nat. Rev. Genet.*, **13**, 7–13.
19. Jiang, D., Wei, S., Chen, F., Zhang, Y., Li, J., (2017). TET3-mediated DNA oxidation promotes ATR-dependent DNA damage response. *EMBO Rep.*, **18**, 781–796.
20. Kafer, Georgia R., Li, X., Horii, T., Suetake, I., Tajima, S., Hatada, I., Carlton, Peter M., (2016). 5-Hydroxymethylcytosine marks sites of DNA damage and promotes genome stability. *Cell Rep*, **14**, 1283–1292.
21. He, Y.F., Li, B.Z., Li, Z., Liu, P., Wang, Y., Tang, Q., Ding, J., Jia, Y., Chen, Z., Li, L., Sun, Y., Li, X., Dai, Q., Song, C.X., Zhang, K., He, C., Xu, G.L., (2011). Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science*, **333**, 1303–1307.
22. Maiti, A., Drohat, A.C., (2011). Thymine DNA glycosylase can rapidly excise 5-formylcytosine and 5-carboxylcytosine: potential implications for active demethylation of CpG sites. *J. Biol. Chem.*, **286**, 35334–35338.
23. Zhang, L., Lu, X., Lu, J., Liang, H., Dai, Q., Xu, G.L., Luo, C., Jiang, H., He, C., (2012). Thymine DNA glycosylase

specifically recognizes 5-carboxylcytosine-modified DNA. *Nat. Chem. Biol.*, **8**, 328–330.
24. Schomacher, L., Han, D., Musheev, M.U., Arab, K., Kienhofer, S., von Seggern, A., Niehrs, C., (2016). Neil DNA glycosylases promote substrate turnover by Tdg during DNA demethylation. *Nat. Struct. Mol. Biol.*, **23**, 116–124.
25. Slyvka, A., Mierzejewska, K., Bochtler, M., (2017). Nei-like 1 (NEIL1) excises 5-carboxylcytosine directly and stimulates TDG-mediated 5-formyl and 5-carboxylcytosine excision. *Sci. Rep.*, **7**, 9001.
26. Muller, U., Bauer, C., Siegl, M., Rottach, A., Leonhardt, H., (2014). TET-mediated oxidation of methylcytosine causes TDG or NEIL glycosylase dependent gene reactivation. *Nucleic Acids Res.*, **42**, 8592–8604.
27. Hardeland, U., Bentele, M., Jiricny, J., Schar, P., (2003). The versatile thymine DNA-glycosylase: a comparative characterization of the human, Drosophila and fission yeast orthologs. *Nucleic Acids Res.*, **31**, 2261–2271.
28. Minko, I.G., Christov, P.P., Li, L., Stone, M.P., McCullough, A.K., Lloyd, R.S., (2019). Processing of N(5)-substituted formamidopyrimidine DNA adducts by DNA glycosylases NEIL1 and NEIL3. *DNA Repair*, **73**, 49–54.
29. Sjolund, A.B., Senejani, A.G., Sweasy, J.B., (2013). MBD4 and TDG: multifaceted DNA glycosylases with ever expanding biological roles. *Mutat. Res.*, **743–744**, 12–25.
30. Nemec, A.A., Wallace, S.S., Sweasy, J.B., (2010). Variant base excision repair proteins: contributors to genomic instability. *Semin. Cancer Biol.*, **20**, 320–328.
31. Alexeeva, M., Guragain, P., Tesfahun, A.N., Tomkuviene, M., Arshad, A., Gerasimaite, R., Ruksenaite, A., Urbanaviciute, G., et al., (2018). Excision of the doubly methylated base N(4),5-dimethylcytosine from DNA by Escherichia coli Nei and Fpg proteins. *Philos. Trans. Roy. Soc. B Biol. Sci.*, **373**, 20170337.
32. Hu, L., Lu, J., Cheng, J., Rao, Q., Li, Z., Hou, H., Lou, Z., Zhang, L., et al., (2015). Structural insight into substrate preference for TET-mediated oxidation. *Nature*, **527**, 118–122.
33. Hashimoto, H., Pais, J.E., Zhang, X., Saleh, L., Fu, Z.-Q., Dai, N., Corrêa, I.R., Zheng, Y., et al., (2014). Structure of a Naegleria Tet-like dioxygenase in complex with 5-methylcytosine DNA. *Nature*, **506**, 391–395.
34. Kavoosi, S., Sudhamalla, B., Dey, D., Shriver, K., Arora, S., Sappa, S., Islam, K., (2019). Site- and degree-specific C-H oxidation on 5-methylcytosine homologues for probing active DNA demethylation. *Chem. Sci.*, **10**, 10550–10555.
35. Ghanty, U., DeNizio, J.E., Liu, M.Y., Kohli, R.M., (2018). Exploiting substrate promiscuity to develop activity-based probes for ten-eleven translocation family enzymes. *J. Am. Chem. Soc.*, **140**, 17329–17332.
36. Lukinavičius, G., Lapienė, V., Staševskij, Z., Dalhoff, C., Weinhold, E., Klimašauskas, S., (2007). Targeted labeling of DNA by methyltransferase-directed transfer of activated groups (mTAG). *J. Am. Chem. Soc.*, **129**, 2758–2759.
37. Lukinavičius, G., Tomkuvienė, M., Masevičius, V., Klimašauskas, S., (2013). Enhanced chemical stability of adomet analogues for improved methyltransferase-directed labeling of DNA. *ACS Chem. Biol.*, **8**, 1134–1139.
38. Lukinavičius, G., Lapinaitė, A., Urbanavičiūtė, G., Gerasimaitė, R., Klimašauskas, S., (2012). Engineering the DNA cytosine-5 methyltransferase reaction for sequence-specific labeling of DNA. *Nucleic Acids Res.*, **40**, 11594–11602.
39. Liutkevičiūtė, Z., Lukinavičius, G., Masevičius, V., Daujotytė, D., Klimašauskas, S., (2009). Cytosine-5-methyltransferases add aldehydes to DNA. *Nat. Chem. Biol.*, **5**, 400–402.
40. Guza, R., Kotandeniya, D., Murphy, K., Dissanayake, T., Lin, C., Giambasu, G.M., Lad, R.R., Wojciechowski, F., et al., (2011). Influence of C-5 substituted cytosine and related nucleoside analogs on the formation of benzoapyrene diol epoxide-dG adducts at CG base pairs of DNA. *Nucleic Acids Res.*, **39**, 3988–4006.
41. Parker, M.J., Weigele, P.R., Saleh, L., (2019). Insights into the biochemistry, evolution, and biotechnological applications of the ten-eleven translocation (TET) enzymes. *Biochemistry*, **58**, 450–467.
42. Xue, J.-H., Chen, G.-D., Hao, F., Chen, H., Fang, Z., Chen, F.-F., Pang, B., Yang, Q.-L., et al., (2019). A vitamin-C-derived DNA modification catalysed by an algal TET homologue. *Nature*, **569**, 581–585.
43. Dou, H., Mitra, S., Hazra, T.K., (2003). Repair of oxidized bases in DNA bubble structures by human DNA glycosylases NEIL1 and NEIL2. *J. Biol. Chem.*, **278**, 49679–49684.
44. Vartanian, V., Minko, I.G., Chawanthayatham, S., Egner, P. A., Lin, Y.-C., Earley, L.F., Makar, R., Eng, J.R., et al., (2017). NEIL1 protects against aflatoxin-induced hepatocellular carcinoma in mice. *Proc. Natl. Acad. Sci. USA*, **114**, 4207–4212.
45. Dalhoff, C., Lukinavičius, G., Klimašauskas, S., Weinhold, E., (2006). Synthesis of S-adenosyl-L-methionine analogs and their use for sequence-specific transalkylation of DNA by methyltransferases. *Nat. Protoc.*, **1**, 1879–1886.
46. Hore, T.A., von Meyenn, F., Ravichandran, M., Bachman, M., Ficz, G., Oxley, D., Santos, F., Balasubramanian, S., et al., (2016). Retinol and ascorbate drive erasure of epigenetic memory and enhance reprogramming to naïve pluripotency by complementary mechanisms. *Proc. Natl. Acad. Sci. USA*, **113**, 12202–12207.
47. Daujotytė, D., Vilkaitis, G., Manelytė, L., Skalicky, J., Szyperski, T., Klimašauskas, S., (2003). Solubility engineering of the HhaI methyltransferase. *Protein Eng.*, **16**, 295–301.
48. Bandaru, V., Sunkara, S., Wallace, S.S., Bond, J.P., (2002). A novel human DNA glycosylase that removes oxidative DNA damage and is homologous to Escherichia coli endonuclease VIII. *DNA Repair*, **1**, 517–529.