8592–8604 *Nucleic Acids Research, 2014, Vol. 42, No. 13* *Published online 19 June 2014*
doi: 10.1093/nar/gku552

TET-mediated oxidation of methylcytosine causes TDG or NEIL glycosylase dependent gene reactivation

Udo Müller, Christina Bauer, Michael Siegl, Andrea Rottach and Heinrich Leonhardt*

Department of Biology II, Ludwig-Maximilians University Munich and Center for Integrated Protein Science Munich (CIPSM), 82152 Planegg-Martinsried, Germany

Received March 28, 2014; Revised June 05, 2014; Accepted June 10, 2014

ABSTRACT

The discovery of hydroxymethyl-, formyl- and carboxylcytosine, generated through oxidation of methylcytosine by TET dioxygenases, raised the question how these modifications contribute to epigenetic regulation. As they are subjected to complex regulation *in vivo*, we dissected links to gene expression with *in vitro* modified reporter constructs. We used an Oct4 promoter-driven reporter gene and demonstrated that *in vitro* methylation causes gene silencing while subsequent oxidation with purified catalytic domain of TET1 leads to gene reactivation. To identify proteins involved in this pathway we screened for TET interacting factors and identified TDG, PARP1, XRCC1 and LIG3 that are involved in base-excision repair. Knockout and rescue experiments demonstrated that gene reactivation depended on the glycosylase TDG, but not MBD4, while NEIL1, 2 and 3 could partially rescue the loss of TDG. These results clearly show that oxidation of methylcytosine by TET dioxygenases and subsequent removal by TDG or NEIL glycosylases and the BER pathway results in reactivation of epigenetically silenced genes.

INTRODUCTION

DNA methylation at the C5-position of cytosine plays an essential role in a variety of fundamental processes, such as early embryonic development, X-chromosome inactivation, genome stability and imprinting (1,2). In vertebrates, this epigenetic modification is set by the three DNA methyltransferases DNMT1, DNMT3A and DNMT3B and the regulatory subunit DNMT3L (3–5).

Recently, it was discovered that the TET family of Fe(II)- and 2-oxoglutarate-dependent dioxygenases can successively convert 5-methylcytosine to 5-hydroxymethylcytosine (hmC), 5-formylcytosine (fC) and 5-carboxylcytosine (caC) *in vitro* and *in vivo* (6–8). Three different TET proteins (TET1, TET2 and TET3), each showing tissue-specific differential expression (9), have been identified in mouse and human (10). Functional studies indicate that they are involved in a variety of cellular processes including epigenetic reprogramming, differentiation, myelopoiesis and imprinting (11–13). Mutations of *TET2* correlating with lower hmC levels and altered gene expression patterns have been linked to various hematopoietic malignancies (14,15).

The discovery of TET proteins and their catalytic products hmC, fC and caC has raised the question about the functions of these oxidized cytosine variants. They might serve as independent epigenetic signals and have been shown to recruit a distinct and dynamic set of 'reader' proteins in embryonic stem cells (ESCs) and differentiated cells (16). It also has been described that cytosine oxidation affects the efficiency of transcription by RNA polymerase II (17). However, the low abundance of fC and caC suggests that these cytosine variants are quickly processed *in vivo* and have been proposed to be intermediates in active DNA demethylation (18,19).

Whereas the mechanism of setting the methylation mark is well understood, the process of its removal has long been elusive. DNA demethylation may either occur by a passive process via the inhibition of DNMT1 maintenance methylation after replication (20–22) or by an active enzymatic reaction. In principle, there are three possibilities: first, the direct removal of the methyl group, second, the excision of either the methylated cytosine or third, of the entire nucleotide. It is currently proposed that the additional oxidized cytosine derivatives hmC, fC and caC are intermediates in active DNA demethylation, thereby contributing to epigenetic plasticity and transcriptional regulation (23,24).

Several biochemical studies revealed that thymine DNA glycosylase (TDG) can specifically bind to and excise fC and caC, resulting in abasic sites, which might be subsequently processed by the base-excision repair (BER) machinery (18,25). In general, the BER pathway repairs damaged DNA sites through recognition and excision of base lesions by substrate-specific glycosylases. The generated abasic site is subsequently cleaved by the AP endonuclease 1 (APEX1), leading to a single-strand break, which is recognized by PARP1 through its N-terminal zinc fingers. PARP1

*To whom correspondence should be addressed: Tel: +49 89 2180 74229; Fax: +49 89 2180 74236; Email: h.leonhardt@lmu.de

then recruits XRCC1, LIG3 and DNA polymerase beta to complete the BER reactions (26–28). TDG depletion in mice causes embryonic lethality, and TDG deficient ESCs display prominent alterations of CpG modifications at a large number of gene regulatory regions (29,30).

Another discussed alternative for DNA demethylation is based on the initial deamination of hmC to hydroxymethyluracil (hmU) by members of the AID/APOBEC cytidine deaminase family (23). In the following step, hmU might be excised either by TDG, methyl-CpG-binding domain protein 4 (MBD4) or the single-strand-specific monofunctional uracil-DNA glycosylase 1 (SMUG1) (31–33). However, there is evidence that AID/APOBEC members are less active on modified cytosines *in vitro* or *in vivo*, challenging the prominence of the proposed deamination-linked demethylation pathway in living cells (34). Furthermore, a direct decarboxylation of caC to unmodified cytosine has been detected in ESC lysates, but no specific decarboxylase has been identified so far (35).

In addition to TDG, two members of the NEIL family of glycosylases (NEIL1 and NEIL3) have recently been identified as potential binders for oxidized cytosine derivatives (16). However, their function in TET-dependent demethylation has not been investigated to date.

To unravel the effects of TET-mediated cytosine oxidation on gene expression, we generated *in vitro* modified pOct4-reporter plasmids and monitored their *in vivo* expression in ESCs. Whereas methylation of the reporter DNA leads to silencing of gene expression, subsequent oxidation results in gene reactivation. We show that TET proteins interact with BER factors *in vivo* and propose that the observed oxidation-dependent gene reactivation requires the BER machinery. We demonstrate that initiation of this pathway is mainly dependent on TDG activity, but not on MBD4. Our results also indicate that the glycosylases NEIL1, NEIL2 and NEIL3 can contribute to an alternative BER pathway for DNA demethylation and cause gene reactivation.

MATERIALS AND METHODS

Cell culture and transfection

Human embryonic kidney 293T (HEK293T) and baby hamster kidney (BHK) cells containing a stably integrated lac operator array (36) were cultured in Dulbecco's modified Eagle's medium (Sigma) supplemented with 10% Fetal Calf Serum (FCS) (Biochrom) and 50 μg/μl gentamycine (PAA). HEK293T and BHK cells were transiently transfected using polyethylenimine pH 7.0 (Sigma) according to the manufacturer's instructions.

Mouse wild-type (wt) E14 as well as *Tdg*-/- and *Mbd4*-/- ESCs (29,37) were cultured on gelatin coated flasks or optical 96-well plates (Greiner) using 1000 U/ml LIF, 1μM PD032591 and 3 μM CHIR99021 (Axon Medchem, (38)). ESCs were transfected with Lipofectamine 2000 (Invitrogen) according to the manufacturer's instructions.

In vitro methylation and oxidation of plasmid DNA

In vitro methylation of pOct4-GFP plasmid DNA was performed using M.SssI methyltransferase (New England Biolabs) according to the manufacturer's instructions. The methylation status of the plasmid was tested by HpaII and MspI (Fermentas) digestion.

For the *in vitro* oxidation, GFP-TET1CD or GFP-TET1CD^{mut} (H1652Y, D1654A) was purified from mammalian cells. In detail, HEK293T cells were transfected with an expression construct for GFP-TET1CD/TET1CD^{mut} and immunoprecipitation was carried out using GBP-Ni-NTA beads. Proteins were eluted using imidazole. The *in vitro* methylated plasmid was diluted in TET reaction buffer (50 mM HEPES pH 8.0, 75 μM Fe(II), 2 mM Sodium-Ascorbate, 1 mM Di-Sodium-Ketoglutarate (39)) and added to the purified GFP-TET1CD.

Digestion of hydroxymethylated plasmid with PvuRts1I

A total of 200 ng oxidized plasmid DNA and 100 ng of reference DNA fragments containing exclusively unmodified C, mC or hmC were digested with PvuRts1I (150 mM NaCl, 20 mM Tris pH 8.0, 5 mM $MgCl_2$, 1 mM DTT) at 22°C for 20 min (40). The reaction was inactivated at 65°C for 10 min and digestion of the samples was analyzed by agarose-gelelectrophoresis.

Co-immunoprecipitation (Co-IP) using the GFP-Trap

Note that 36 h after transfection, whole cell lysates of HEK293T cells were prepared using RIPA-lysis buffer (50 mM Tris pH 7.0, 150 mM NaCl, 0.1% sodium dodecyl sulphate (SDS), 0,5% sodium deoxycholate, 5 mM ethylenediaminetetraacetic acid (EDTA), 2.5 mM $MgCl_2$, 0.5 mM $CaCl_2$, 2 mM PMSF, 1x Mammalian Protease Inhibitor Cocktail and 1 μg/μl DNaseI). After centrifugation, 10% of the supernatant was collected as input fraction and the remaining supernatant was diluted in IP-buffer (10 mM Tris/Cl pH 7.5, 150 mM NaCl and 0.5 mM EDTA) to 800 μl. Green Fluorescent Protein (GFP)-Trap beads (Chromotek, (41)) pre-equlibrated with IP-buffer were added to the supernatant dilution and rotated for 1.5 h at 4°C. The GFP-beads were centrifuged and 10% of the supernatant was collected as flowthrough fraction. For western blotting, input, flowthrough and bead fractions were boiled with Laemmli buffer at 95°C for 10 min, loaded on an SDS-polyacrylamide gel electrophoresis (SDS-PAGE) and transferred to a nitrocellulose membrane (Biorad). Immunodetection was performed using mouse monoclonal anti-GFP (Roche, 11814460001) or rat monoclonal anti-RFP antibodies (42) and Alexa488/Alexa594 coupled secondary antibodies (Jackson ImmunoResearch).

For mass spectrometry analysis, protein samples were denatured with GdnHCl, precipitated with acetone and digested with trypsin. Peptide mixtures were analyzed using electrospray tandem mass spectrometry in collaboration with the Core Facility of the Max-Planck-Institute for Biochemistry, Martinsried. Experiments were performed with an LTQ Orbitrap mass spectrometer (Thermo Scientific). Spectra were analyzed with MaxQuant (43).

Fluorescent-three-hybrid assay (F3H)

Transgenic BHK cells containing stably integrated lac-operator repeats (36) were grown to 60–70% confluence

on coverslips. The cells were transiently cotransfected with expression constructs encoding for LacI-GBP, murine RFP/mCherry- and GFP-fusion proteins (44). As controls the catalytically inactive mutants GFP-TET1$^{H1652Y\&D1654A}$, GFP-TET2$^{H1304Y\&D1306A}$ and GFP-TET3$^{H950Y\&D952A}$ were used. Note that 24 h after transfection cells were fixed with 4% formaldehyde in phosphate buffered saline (PBS), permeabilized with 0.5% Triton X-100/PBST, counterstained with 4',6-diamidino-2-phenylindole (DAPI) and mounted in Vectashield (Vector Laboratories). Imaging was performed using a Leica TCS SP5 confocal laser scanning microscope with a 63x/1.4 NA Plan-Apochromat oil immersion objective.

The Operetta automated imaging system (PerkinElmer) was used for F3H quantification (Harmony 3.6 software). After imaging, nuclei were detected based on DAPI signal. The lacO-spot was defined in the GFP channel and screened for enrichment at the RFP channel (intensity spot >1.2x mean intensity nucleus; see also Supplementary Figure S4).

High-throughput pOct4-reporter gene expression analysis

Wild type, *Tdg*-/- *and Mbd4*-/- ESCs were transiently transfected with unmodified, M.SssI methylated or *in vitro* oxidized pOct4-GFP plasmid DNA on coated optical 96-well plates (PerkinElmer). Note that 24 h after transfection, the cells were fixed with 4% formaldehyde/PBS, permeabilized with 0.5% Triton X-100/PBST and counterstained with DAPI.

Images were acquired using the Operetta automated imaging system with an 40x high NA objective and expression was quantified using Harmony 3.6 software (PerkinElmer). A total of 16 fields per well were imaged, cells were counted and segmented into nuclei and cytoplasm on the basis of DAPI and reporter mCherry/GFP signal. Prism software (GraphPad) was used for statistical analysis.

Generation of stable transgenic cell lines

Tdg-/- ESCs stably expressing GFP-fusions of wt TDG, TDG^{N151A}, TDG^{N168D} and TDG^{M280H} were generated by transfecting the respective plasmids in the presence of the selection marker blasticidine followed by repeated sorting for GFP expression with the fluorescence-activated cell sorting (FACS) AriaII (Becton-Dickinson) system. Single cell sorting was used to generate clonal transgenic cell lines. GFP-expression of the single cell clones was analyzed using the Operetta system or western blotting.

Activity of GFP-TDG *in vitro*

GFP-TDG and the different mutants were expressed in HEK293T cells and immunoprecipitated as described above. Equal amounts of GFP-tagged protein immobilized on GFP-Trap beads were incubated with 0.4 μM of DNA substrate in TDG reaction buffer (20 mM TrisHCl, pH 8.0 or pH 6.5, 100 mM NaCl, 1 mM $MgCl_2$, 0.2 mM EDTA, 1 mM DTT, 0.01 mM $ZnCl_2$). In detail, these DNA substrates were fluorescently (ATTO550) labeled 42 bp oligonucleotides (GGA TGA TGA CTC TTC TGG TCC GGA TGG TAG TTA AGT GTT GAG) (Eurofins MWG Operon) with a central modified CpG site: either hmCpG, fCpG or caCpG or harboring a G○T mismatch at this site. Incubation was performed in the presence of purified GFP-APEX1 for 2 h at room temperature. Following heat-inactivation of TDG (2 min, 95°C), fresh GFP-APEX1 was added and further incubated for 4 h at room temperature. An oligonucleotide containing a deoxyabasic site ('dSpacer', Eurofins MWG Operon) served as a positive control for APEX1 activity. DNA was analyzed on a denaturing 17% polyacrylamide gel with the Typhoon TRIO (GE Healthcare Life Sciences). Quantification was performed with ImageJ.

Genomic DNA extraction and slot blot analysis

Genomic DNA from ESCs was extracted using the Blood & Cell Culture Midi Kit (Qiagen) according to the manufacturer's instruction. Anti-oxidant BHT (200 μM, Sigma) and deaminase inhibitor THU (200 μM, Sigma) were added to the lysis buffer and elution buffer. The Bio-Rad slot blot system was used according to the manufacturer's instruction. Nitrocellulose membranes (Amersham) were crosslinked, blocked with 5% milk and immunostaining was performed using a mouse monoclonal antibody against mC (Eurogentec, 33D3) or rabbit polyclonal antibodies against hmC, fC and caC (Active motif: 39791, 61233, 61224). Alexa488-coupled secondary antibodies were used for detection and the membranes were scanned with the Typhoon TRIO (GE Healthcare Life Sciences). Quantification was performed with ImageJ.

Re-isolation of transfected plasmids

Note that 36 h after transfection, nuclei were extracted from the ESCs using the Blood & Cell Culture Midi Kit (Qiagen) according to the manufacturer's instructions. Plasmid DNA was re-isolated using the Qiaprep Plasmid Midi Kit (Qiagen). A total of 200 ng of re-isolated plasmid DNA was digested with 0.5 μl HpaII (Fermentas).

RESULTS

In vitro oxidation of Oct4 reporter plasmid DNA causes gene activation

Since the discovery of hmC, fC and caC, two major roles for these cytosine modifications have been proposed: first, as intermediates in active DNA demethylation, and second, as independent epigenetic marks. The latter has been investigated by identifying reader proteins for hmC, fC and caC in different murine tissues. These new DNA modifications recruit a dynamic set of readers including DNA repair factors and chromatin remodelers (16). TET-dependent cytosine oxidation has been shown to occur at a large number of gene regulatory elements and repetitive loci (30).

Here, we focus on the effects of hmC, fC and caC on gene expression. We transfected ESCs with Oct4 promoter-driven GFP and mCherry reporter plasmids (pOct4-GFP or pOct4-mCherry) in different modification states and analyzed expression of the reporter by quantitative imaging. To generate the different cytosine modifications *in*

vitro, the unmodified pOct4-reporter construct (CpG) was initially treated with the DNA-methyltransferase M.SssI, thereby creating fully methylated CpG sites (^{m}CpG), and subsequently incubated with the purified catalytic domain of TET1 (TET1CD) to create oxidized cytosine residues (^{ox}CpG; Figure 1a). Specific restriction enzymes were used to monitor the methylation and oxidation state of the plasmid DNA. The *in vitro* methylation of the reporter construct was confirmed by the mC-sensitive restriction endonuclease HpaII and the mC-insensitive enzyme MspI. Both MspI and HpaII fully cleave the unmethylated pOct4-reporter plasmid at CCGG sites, whereas the restriction pattern of the M.SssI-methylated reporter only appeared with MspI digestion (Figure 1b).

To follow the oxidation of mC to hmC by TET1CD, the reporter DNA was treated with the hmC-specific endonuclease PvuRts1I (40) (Supplementary Figure S1a). Treatment with TET1CD resulted in a gradual increase of hmC levels after 15, 45 and 90 min of incubation, visible as progressing fragmentation (Figure 1b). While using a catalytically inactive TET1CD mutant (TET1CD^{mut}) as a control, no hmC levels were detected (^{ox}*CpG), indicating specific enzymatic mC oxidation by TET1CD (Supplementary Figure S1b).

Consistent with the restriction digest results, methylation and oxidation of the reporter plasmid could also be shown by slot blot analysis (Figure 1c). Incubation of methylated plasmid DNA with TET1CD resulted in an increase of not only hmC but also of fC and caC, demonstrating that purified TET1CD did carry out the three oxidation steps *in vitro* (Figure 1c and d). The mC signal decreased over time as hmC, fC and caC appeared in the presence of active TET1CD, while remaining constantly high with TET1CD^{mut} (Supplementary Figure S1d).

Transfection of ESCs with the TET1CD-treated plasmid DNA allows to directly monitor the effect of the oxidized cytosines on gene expression, independent from *in vivo* TET activity. Reporter gene expression from either unmodified, *in vitro* methylated or oxidized pOct4-mCherry was visualized using confocal imaging or automated image acquisition for quantification. Transfection of ESCs with unmodified pOct4-mCherry resulted in a strong nuclear and cytoplasmic expression of the reporter, whereas expression drastically decreased when using the methylated construct. Interestingly, prominent reporter expression could be observed upon transfection of the oxidized plasmid DNA, but not with the ^{ox}*CpG reporter DNA (Figure 1e). This suggests that reactivation of gene expression requires oxidation of methylcytosine by TET proteins.

TET proteins interact with the BER machinery

Currently, three pathways for TET-mediated active DNA demethylation are discussed: TDG-dependent BER, deamination-dependent BER and direct decarboxylation of caC (18,23,35). Since the proteins responsible for gene reactivation in our assay might be physical interaction partners of the TET proteins, we performed an initial unbiased screen for interactors. We expressed GFP-TET1 in HEK293T cells, performed immunoprecipitation and analyzed the co-precipitated proteins by mass spectrometry. Prominently, we found PARP1, XRCC1 and LIG3, a subset of proteins involved in the BER pathway (Supplementary Figure S2a). These results point toward the two BER-dependent demethylation mechanisms.

To further investigate the interplay between TET proteins and BER, we systematically performed Co-IP and a recently established F3H assay of all three TET proteins with the following BER factors: TDG, MBD4, SMUG1, NEIL1, NEIL2, NEIL3, PARP1, LIG3 and XRCC1. F3H allows to directly visualize protein–protein interactions in living cells (44). We therefore co-expressed full-length GFP-TET fusion proteins together with potential mCherry/RFP-tagged interactors (Supplementary Figure S3a) and LacI-GBP in BHK cells containing a stably integrated lac-operator array (45). The GFP-tagged bait is enriched at the lac-operator array via LacI-GBP and is visible as a single spot inside the nucleus. If the mCherry/RFP-tagged prey protein interacts with the bait, it colocalizes at the same spot (Figure 2a).

Consistent with the mass spectrometry results, interactions of TET1, TET2 and TET3 were observed with PARP1, LIG3 and XRCC1 both in Co-IP and F3H. Interestingly, several glycosylases also showed a clear interaction in both assays, among them TDG, MBD4, NEIL1, NEIL2 and NEIL3, but not SMUG1 (Figure 2b and c; Supplementary Figures S2b and S3b, c). Automated high-throughput image analysis was used to quantify the F3H results (Supplementary Figure S4). The interaction of all three TET proteins with TDG was the most robust and detectable in more than 80% of all analyzed cells. For the other factors, numbers vary between 40% and 75% (Figure 2d).

To exclude that the recruitment to the lacO-spot is dependent on the locally enriched cytosine oxidation products generated by TET proteins, we repeated the F3H quantification with catalytically inactive TET mutants (Figure 2e). The percentage of cells showing an interaction did hardly differ compared to the assay with active TET proteins. The only exception is XRCC1 where only half as many cells displayed a colocalization at the spot, suggesting a potential cooperative effect.

Taken together, these results suggest that TET proteins physically interact with the BER machinery and are therefore able to recruit these factors to the site of cytosine oxidation for immediate removal of the modified base *in vivo*.

TDG but not MBD4 mediates oxidation-dependent gene reactivation in ESCs

Since DNA glycosylases catalyze the first step of the BER pathway and are therefore the initiators of TET-dependent cytosine demethylation, we investigated their role in gene reactivation with the reporter gene assay. Besides TDG, also MBD4 has been implicated in DNA demethylation via excision of hmU, the deamination product of hmC (46). To investigate the role of these two glycosylases, we transfected *Tdg*-/- and *Mbd4*-/- ESCs (29,37) with either unmodified, *in vitro* methylated or oxidized pOct4-mCherry plasmid DNA. In contrast to wt E14 ESCs, *Tdg*-/- ESCs showed no reporter gene expression from the oxidized plasmid. However, *Mbd4* knockouts were able to fully reactivate gene expression from the oxidized reporter construct.

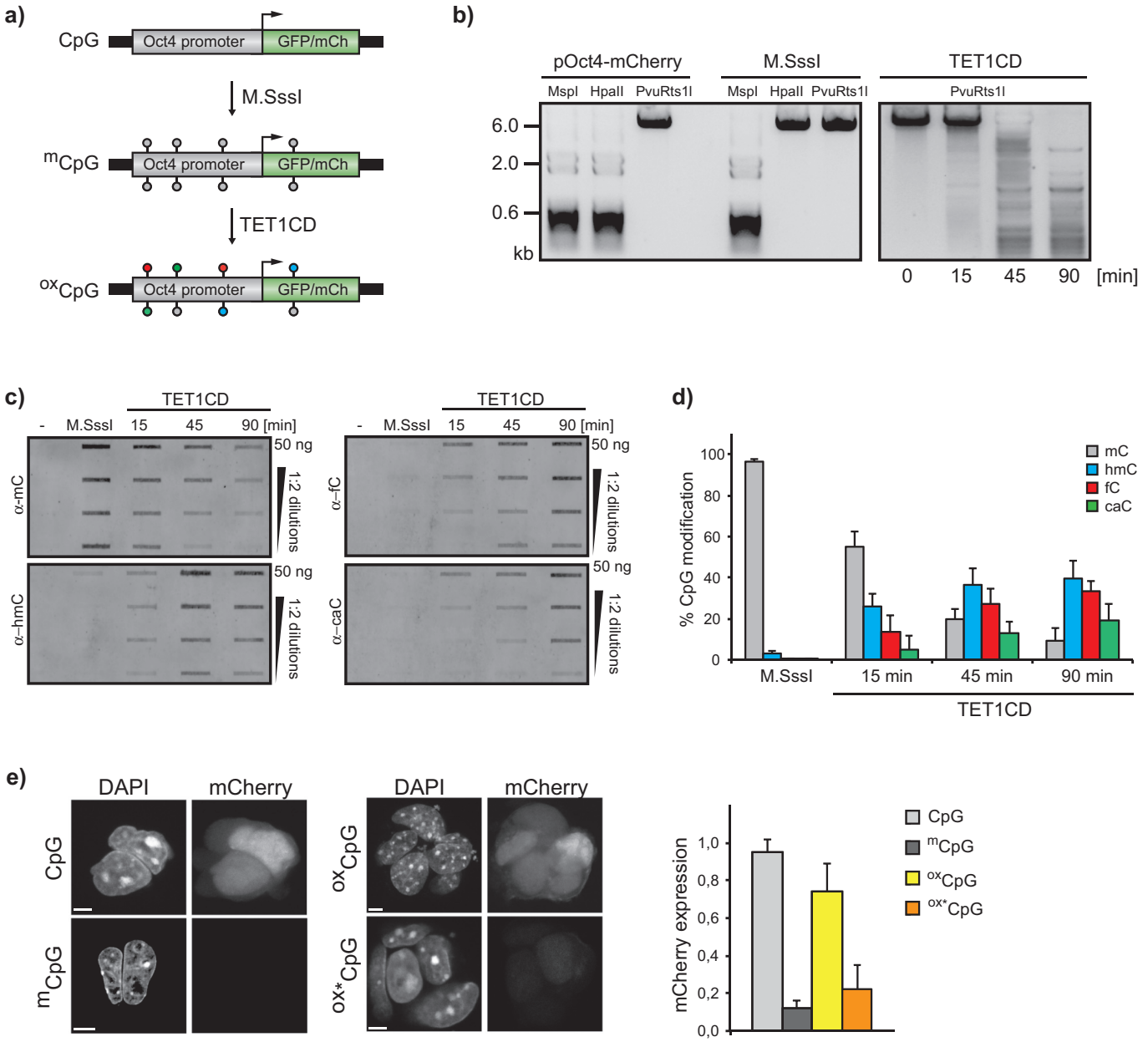

Figure 1. *In vitro* oxidation of mC causes gene reactivation in ESCs. (**a**) Schematic representation of *in vitro* reporter DNA modification: Unmethylated pOct4-reporter DNA was methylated using the CpG methyltransferase M.SssI. Incubation with purified TET1CD results in oxidation of mC sites to hmC, fC and caC. (**b**) M.SssI treatment of pOct4-mCherry results in full methylation as shown after restriction with the methylation sensitive enzyme HpaII. MspI cuts irrespective of the methylation state. The hmC-specific restriction endonuclease PvuRts1I detects increasing hmC levels during incubation of methylated pOct4-mCherry with TET1CD. (**c**) Cytosine modification states of untreated, methylated and TET1CD oxidized pOct4-mCherry plasmid DNA were detected by slot blot. A 2-fold serial dilution of the plasmid DNA was loaded and detected using antibodies against mC, hmC, fC and caC. A gradual increase of hmC, fC and caC signals was obtained with longer incubation time with TET1CD while the mC signal decreases accordingly. (**d**) Quantification of the slot blot signals of pOct4-mCherry after treatment with TET1CD shows increasing oxidation of mC to hmC, fC and caC. The sum of all CpG modification signals was set to 100%. Error bars indicate standard deviation ($n = 3$). (**e**) ESCs were transfected with pOct4-mCherry plasmids containing either unmodified (CpG), methylated (mCpG), TET1CD-oxidized (^{ox}CpG) or TET1CD^{mut}-treated (^{ox}*CpG) cytosines. Confocal imaging and quantification show reporter gene silencing upon methylation and reactivation upon oxidation. Cells were fixed with formaldehyde and counterstained with DAPI. Scale bar: 5 μm. (Right: $n = 200\,000$; error bars indicate standard deviation).

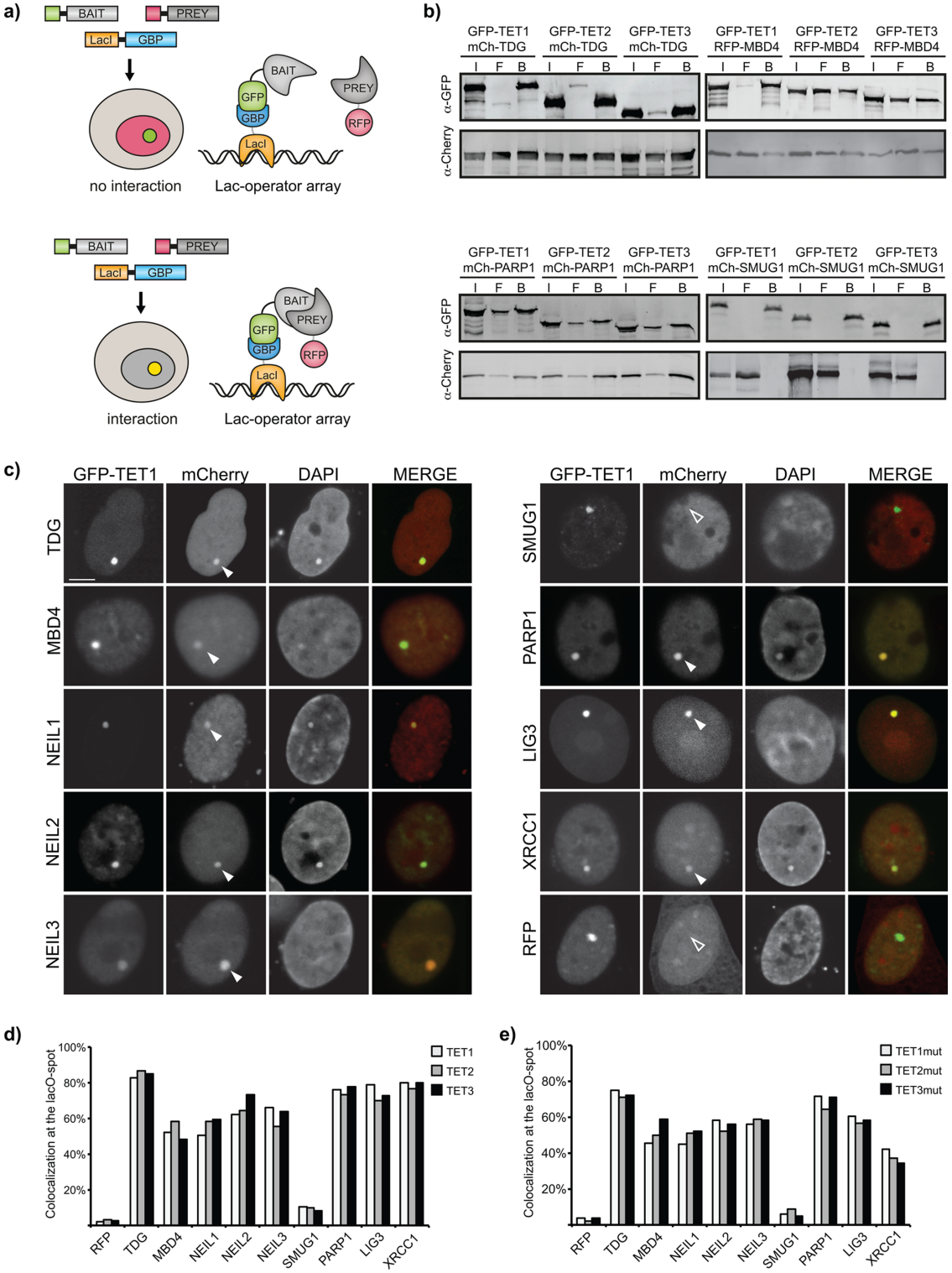

Figure 2. All three TET proteins interact with a variety of BER factors. (**a**) Scheme depicting the F3H assay for *in vivo* visualization of protein interactions: BHK cells containing a stably integrated lac-operator array were transfected with plasmids expressing a lac-repressor-GBP fusion protein, GFP-BAIT and mCherry/RFP-PREY. The Lac-repressor binds to the lac-operator array and recruits the GFP-BAIT through GBP. Proteins interacting with the BAIT are consequently enriched at the lac-operator array. (**b**) Co-IP with subsequent SDS-PAGE and western blotting shows interaction of the three TET proteins with TDG, PARP1, MBD4 but not SMUG1. GFP-tagged TET proteins and the respective mCherry-fusions were expressed in HEK293T cells and immunoprecipitated with the GFP-Trap (I: Input; F: Flowthrough; B: Bound). (**c**) F3H was used to confirm TET1 interactions with different factors involved in BER. GFP-TET1 is enriched at the lac-operator array and mCherry-tagged interacting factors are recruited to the same spot (solid triangle). SMUG1 shows no co-localization at the lacO-spot (empty triangle). Scale bar: 5 μm. (**d**) Quantification of the F3H assay of all three TET proteins with the indicated BER factors. Bars represent percentage of cells that show colocalization at the lacO-spot (n>200). (**e**) As in (**d**) but with catalytically inactive TET mutants.

The rescue experiment with wt GFP-TDG re-established the ability of *Tdg*-/- to express mCherry from the ^{ox}CpG plasmid, whereas the transient rescue of *Mbd4*-/- cells with GFP-MBD4 led to no significant differences with regard to reporter gene expression (Figure 3a).

High-throughput image analysis of 200 000 cells revealed that the signal from the oxidized plasmid in wt E14 ESCs is about 80% of the signal from unmodified reporter DNA. In *Tdg*-/- cells, the ^{ox}CpG signal drops almost to the level of the fully methylated reporter. Thus, the ^{ox}CpG construct remains silent in *Tdg*-/- ESCs. The stable rescue with wt GFP-TDG led to a recovery of mCherry expression to about 70% of the signal from the unmodified reporter. Knockout of *Mbd4* and also the corresponding rescue did not alter expression levels compared to wt E14 ESCs (Figure 3b), indicating that oxidation-dependent gene reactivation requires TDG but not MBD4 expression.

To gain insight into the mechanism by which TDG mediates gene activation, we recovered the transfected reporter plasmid DNA from wt E14 ESCs, *Tdg*-/- and *Mbd4*-/- ESCs. We analyzed the modification status by digestion with HpaII, which specifically cuts unmodified cytosines in a CCGG context (Supplementary Figure S1c). We observed a reappearance of the HpaII restriction pattern on the ^{ox}CpG plasmids isolated from wt E14 and *Mbd4*-/- ESCs indicating that TET-dependent demethylation occurred *in vivo*. HpaII only displayed minor activity on ^{ox}CpG reporter DNA re-isolated from *Tdg*-/- cells (Figure 3c). This provides strong evidence that the substitution of oxidized cytosine by unmodified cytosine is the major mechanism for the observed gene reactivation and that this substitution depends on TDG.

TDG activity is required for reporter gene reactivation

To investigate whether the glycosylase activity of TDG is responsible for the observed gene reactivation, we generated *Tdg*-/- rescue cell lines, which stably express GFP-fusions of either wt, catalytically inactive (N151A), DNA binding deficient (M280H) or caC-specific (N168D) TDG (31,47) at equal levels (Supplementary Figure S5a). The murine caC-specific mutant corresponds to a published human TDG mutant (48). Expression of the oxidized mCherry-reporter plasmid in the stable rescue with wt GFP-TDG increased almost to the levels of wt E14 ESCs in line with the results from the transient rescues. In contrast, the catalytically inactive mutant (N151A) or DNA binding deficient (M280H) TDG was not able to recover reporter expression. Cells stably expressing the caC-specific TDG (N168D) were only capable of partially restoring reporter expression (Figure 4a and b; Supplementary Figure S5b). This demonstrates that base excision by active TDG is essential for gene reactivation.

To further characterize the activity of wt TDG as well as the TDG mutants, we established an *in vitro* assay based on a defined DNA substrate with a single modification site. Base excision by TDG generates an abasic site, which can specifically be converted into a single-strand break by purified APEX1 and can be detected on a denaturing gel. Thus, this assay mimics the first two steps of the BER reactions and is also applicable on fluorescently labeled DNA substrates, in contrast to the previously described 'nicking assay', in which alkaline treatment and subsequent boiling is used to create single-strand breaks (49,50).

TDG has long been known to repair G○T mismatches (51) and was recently found to excise fC and caC (18,25). We could show that wt TDG is more active on fC and caC than on G○T mismatches, its eponymous substrate. No activity was detected on hmC-containing DNA. Although the caC-specific TDG mutant (N168D) was able to partially reactivate reporter gene expression *in vivo*, we could only detect basal activity on the caC substrate *in vitro* (Figure 4c, Supplementary Figure S5d). Since the activity of TDG *in vitro* is pH-dependent (48), we repeated the assay at pH = 6.5. Under these conditions, the preference of the caC-specific TDG mutant (N168D) toward caC could be confirmed (Supplementary Figure S5d). Taken together, these data suggest that the excision of fC and caC by TDG is essential for TET-mediated demethylation and causes reactivation of gene expression.

NEIL1, NEIL2 and NEIL3 glycosylases can partially compensate for loss of TDG

Besides TDG, we identified the family of NEIL glycosylases as interactors of TET proteins (Figure 2c). Interestingly, NEIL1 and NEIL3 have also been described as binders of hmC, fC or caC cytosines in ESCs in a proteome wide screen (16). However, their function in this context has not been investigated so far.

To elucidate whether the NEIL glycosylase family contributes to gene reactivation, we measured the expression of the modified reporter plasmids in *Tdg*-/- cells transiently overexpressing NEIL1, NEIL2 or NEIL3 at similar levels (Figure 5a). Interestingly, we observed a significant increase of pOct4-GFP expression in *Tdg*-/- ESCs rescued with mCherry-NEIL1, NEIL2 or NEIL3 in comparison to *Tdg*-/- ESCs. However, lower expression levels as in rescues with wt TDG were detected. RFP-MBD4 was not able to rescue the *Tdg*-/- phenotype (Figure 5b).

Additionally, we isolated genomic DNA from wt E14 and *Tdg*-/- ESCs as well as from the transient rescues with wt TDG and NEIL1, 2 and 3. Slot blot analyses were carried out for relative hmC, fC and caC quantifications. Genomic hmC was present at comparable levels in all tested cell lines and was not affected by *Tdg* knockout or NEIL overexpression. Since TDG is able to recognize and excise fC and caC, accumulation of these oxidized bases was observed in *Tdg*-/- ESCs, consistent with previous reports (30,52). Rescues of *Tdg*-/- cells with transiently expressed wt TDG, NEIL1, NEIL2 or NEIL3 resulted in decreased genomic fC, and for wt TDG and NEIL1, also in decreased caC levels (Figure 5c). These findings support the role of the glycosylases TDG, NEIL1, NEIL2 and NEIL3 in active DNA demethylation and subsequent reactivation of gene expression via excision of fC and caC followed by BER (Figure 6).

DISCUSSION

In this study, we investigated the effects of the oxidized cytosine variants hmC, fC and caC on gene expression. By carrying out the enzymatic oxidation of a methylated

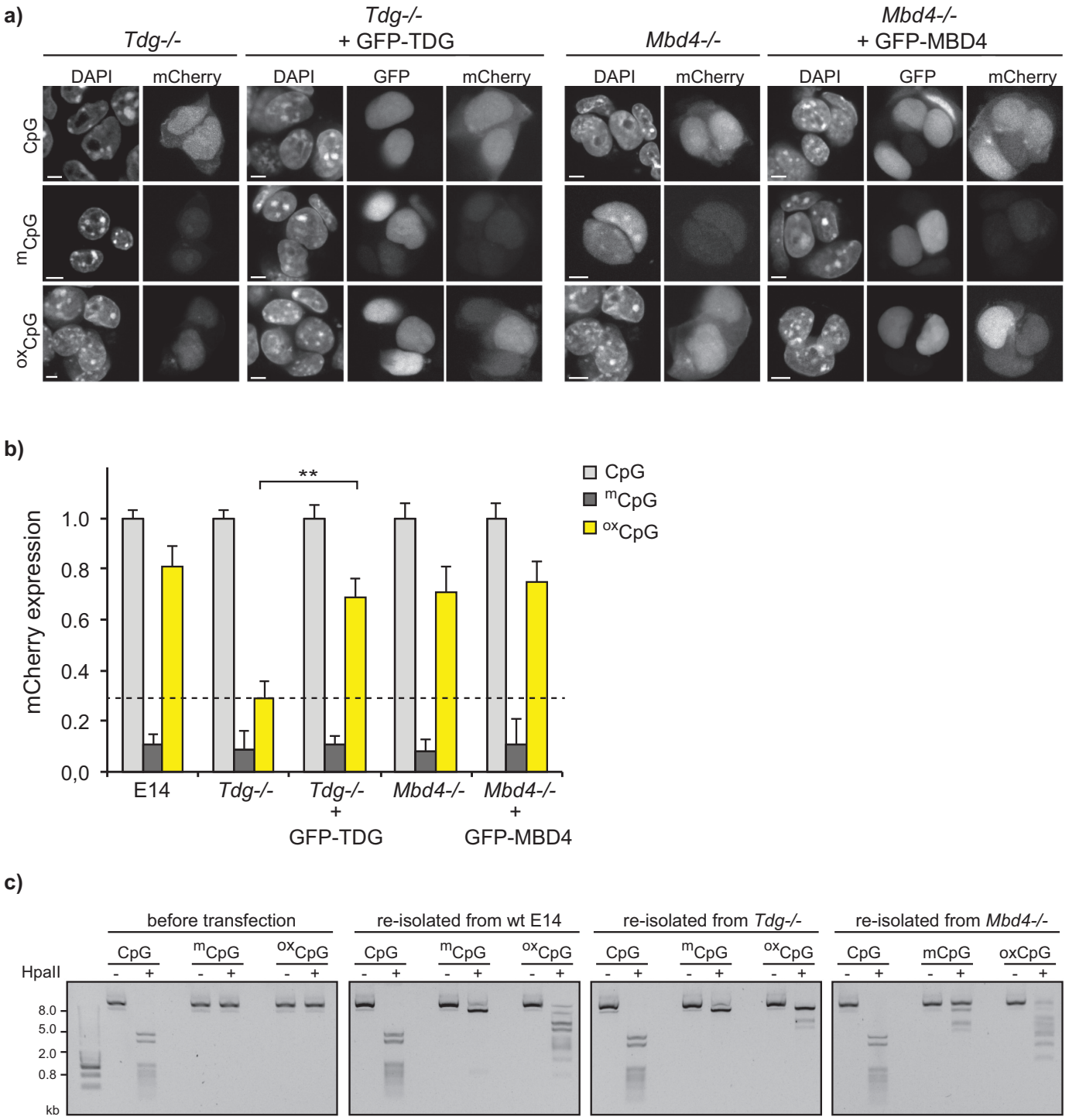

Figure 3. Oxidation of ^{m}CpG plasmid DNA leads to TDG-dependent gene reactivation. (**a**) *Tdg*-/- ESCs were transfected with pOct4-mCherry plasmids containing either unmodified, methylated or oxidized CpGs. Confocal images show a defect of ^{ox}CpG gene reactivation in *Tdg*-/- ESCs but not in *Mbd4*-/- ESCs. Transient rescue of *Tdg*-/- ESCs with GFP-TDG re-establishes ^{ox}CpG reporter gene expression. Cells were fixed with formaldehyde and counterstained with DAPI. Scale bar: 5 μm. (**b**) High-throughput image acquisition and quantification of pOct4-mCherry expression shows that oxidation of ^{m}CpG sites in the pOct4-reporter results in reactivation of mCherry-expression in wt E14 ESCs and *Mbd4*-/- ESCs, but not in *Tdg*-/- ESCs. Expression of GFP-TDG rescues the phenotype (student's *t*-test, $^{**}P < 0.025$, $n = 200\,000$; error bars indicate standard deviation). (**c**) Analytical digest with HpaII of differentially modified reporter plasmid DNA before and after transfection confirms substitution of ^{ox}CpG with CpG in wt E14 and *Mbd4*-/- ESCs.

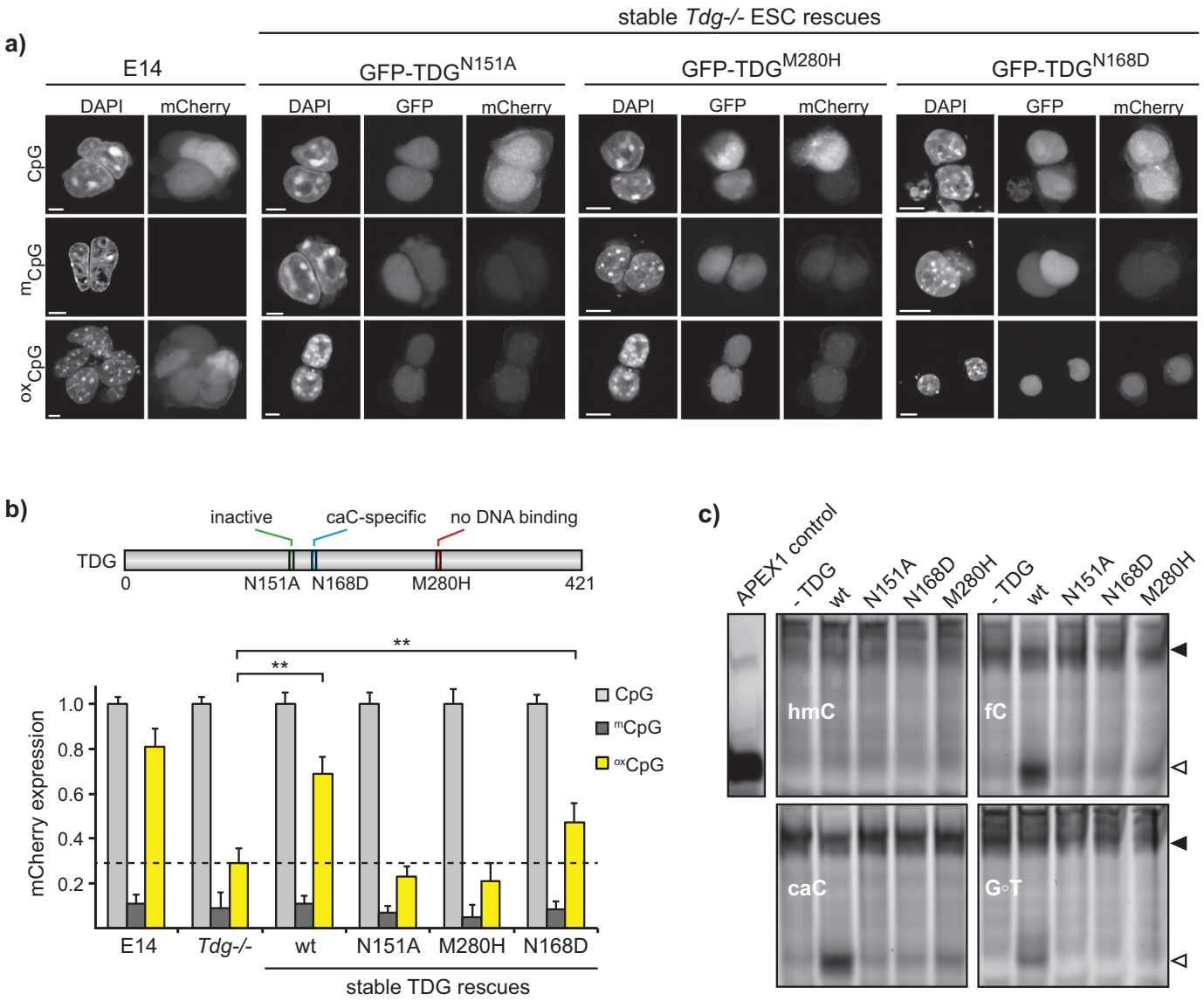

Figure 4. TDG activity is essential for gene reactivation. (**a**) Confocal images depicting expression levels of oxidized pOct4-mCherry expression in *Tdg*-/- ESCs stably rescued with GFP-TDG^{N151A}, GFP-TDG^{M280H} and GFP-TDG^{N168D} in comparison to wt E14 ESCs. Scale bar: 5 μm. (**b**) High-throughput image acquisition and quantification of pOct4-mCherry expression in wt E14, *Tdg*-/- and *Tdg*-/- ESCs stably expressing wt, catalytically inactive, DNA binding deficient and caC-specific TDG mutants. Methylation of the pOct4-mCherry reporter leads to a 5-fold lower expression compared to unmodified plasmid. Oxidation of ^{m}CpG sites in the pOct4-reporter results in reactivation of mCherry-expression in wt ESCs but not in *Tdg*-/- ESCs. This re-increase was also obtained in *Tdg*-/- ESCs rescued with wt or caC-specific TDG, while the latter was not as efficient (student's *t*-test, $^{**}P < 0.025$, $n = 200\,000$; error bars indicate standard deviation). (**c**) TDG activity was monitored using an *in vitro* assay based on the ability of APEX1 to create a single strand break out of an abasic site, detectable as distinct band on the gel (empty triangle). Full length DNA is indicated by a solid triangle. The DNA substrates contain one defined modification site as indicated in the figure. Wt TDG is highly active on fC or caC and to a much lesser extent on a G◦T mismatch.

pOct4-reporter construct *in vitro*, we separated the generation of modified cytosines from their further processing *in vivo*. This allowed us to directly investigate the cellular factors responsible for gene reactivation independent of endogenous TET activity and regulation. In wt ESCs, we observed strong reporter expression from oxidized but not from methylated plasmids, suggesting oxidation-dependent gene reactivation.

To investigate which pathway is responsible for the observed gene reactivation, we searched for potential TET interaction partners. So far, it has been shown that MBD3 colocalizes with TET1 regulating hmC-marked gene expression (53) and that TET1, TET2 and TET3 interact with OGT controlling protein stability, localization and histone modification (54–57). Also, several chromatin-binding factors, such as HDAC1, EZH2 and MeCP2, have been described to associate with TET1 (58). However, none of these factors is likely to be involved in the process of DNA demethylation. Therefore, we performed a mass spectrometry-based pull-down approach in which several BER factors co-precipitated with TET1. To confirm these results, we used Co-IP and a recently described F3H as-

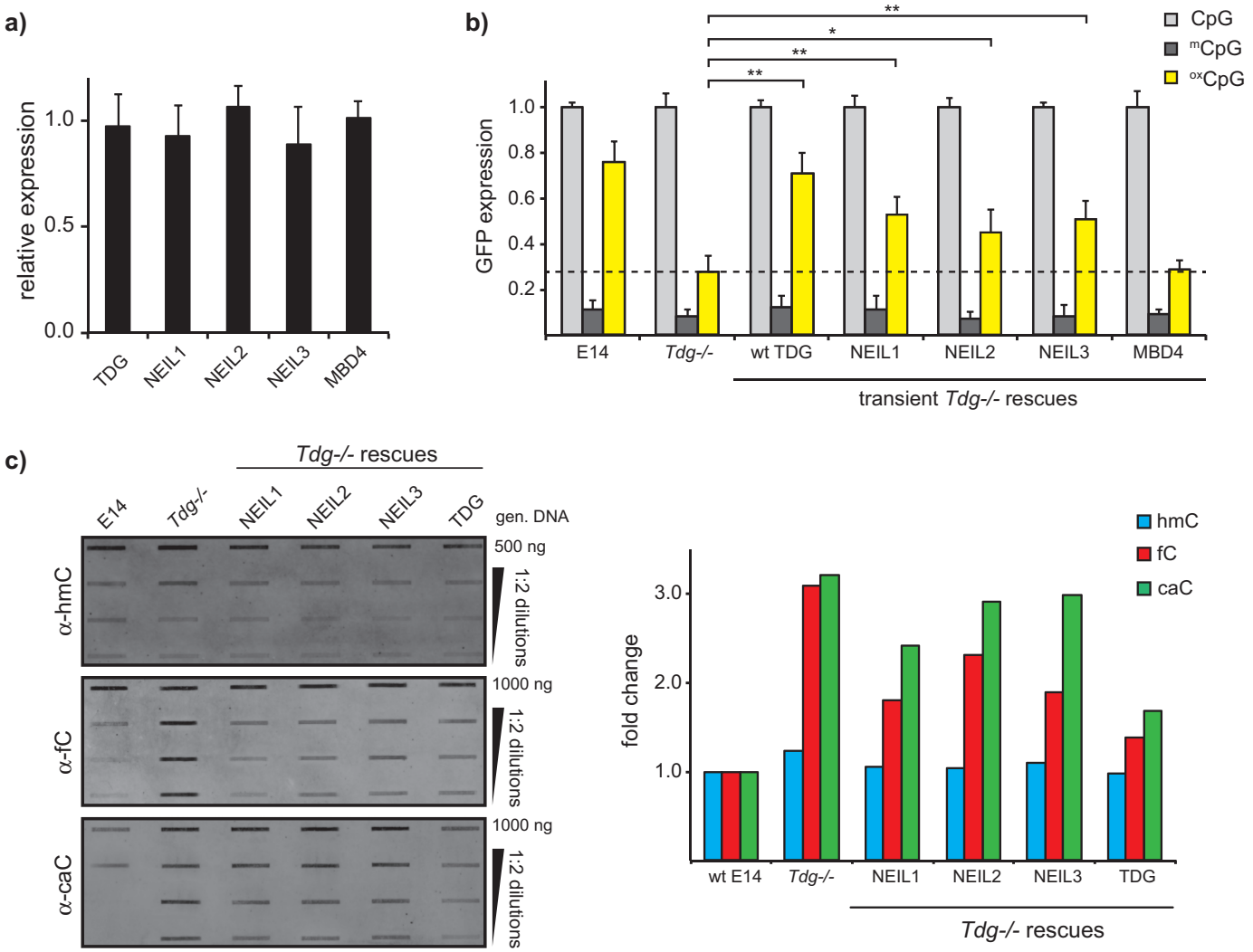

Figure 5. The NEIL glycosylase family can partially compensate for TDG. (**a**) Quantification of mCherry intensities with high-throughput imaging shows that transient *Tdg*-/- rescue ESCs express mCherry-tagged TDG, NEIL1/2/3 and MBD4 glycosylases at comparable levels ($n = 100\,000$; error bars indicate standard deviation). (**b**) The ability of the NEIL family of glycosylases to substitute TDG *in vivo* was monitored by expression of differentially modified pOct4-GFP. With ectopic expression of mCherry-NEIL1, NEIL2 and NEIL3, reporter gene signal was significantly higher than in *Tdg*-/- cells, although not reaching the levels of the wt mCherry-TDG rescue. MBD4 overexpression could not compensate for loss of TDG (student's *t*-test, $*P < 0.05$, $**P<0.025$, $n = 200\,000$; error bars indicate standard deviation). (**c**) Slot blot analysis of genomic DNA isolated from wt E14, *Tdg*-/- and the indicated rescues shows constant levels of hmC and accumulation of fC and caC in *Tdg*-/- cells. Overexpression of wt TDG or NEIL1, 2 or 3 leads to a decrease in genomic fC amounts but does not reach wt E14 levels. Expression of wt TDG or NEIL1 also reduces caC accumulation.

say to test interaction of all three TETs with different BER factors *in vitro* and *in vivo*. We were able to detect interactions of TET proteins with the DNA glycosylases TDG, MBD4 as well as NEIL1, NEIL2 and NEIL3, which excise damaged or oxidized DNA bases (18,25,32,59,60). Furthermore, interactions of all three TET proteins were observed with PARP1, which detects single-strand breaks and modifies repair factors by PolyADP-ribosylation (61). Finally, we showed TET interactions with LIG3 and XRCC1, which are recruited by PARP1 to the site of DNA damage and ligate the DNA strand after the insertion of cytosine (62–64). The observed interactions were largely independent of TET activity, indicating direct protein–protein interactions.

Our findings suggest that both TET-dependent oxidation of mC and subsequent excision of oxidized cytosines by the BER machinery take place in one large protein complex in a spatially and temporally coordinated manner. This close association enables highly efficient replacement of oxidized cytosines. In accordance with these results, fC and caC, in contrast to hmC, are detected at very low genomic levels and are proposed to be immediately removed after generation (18,19).

Initial hypotheses proposed that hmC might be deaminated to hmU by AID/APOBEC deaminases prior to base excision by DNA glycosylases. Suggested candidates were TDG and MBD4, which both have been shown to recognize G○T and G○hmU mismatches (23,65) and have been identified as TET protein interactors in this study. TDG has also been described to be active on fC and caC (18,25). In contrast to wt ESCs, no gene reactivation on the oxidized

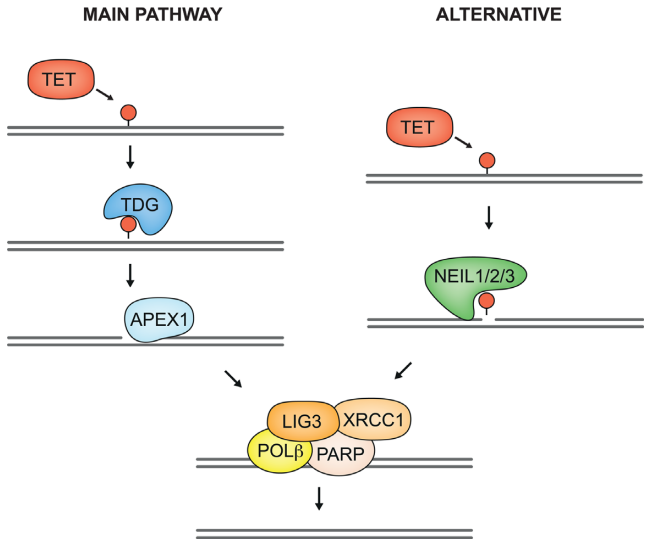

Figure 6. Two alternative pathways for TET-mediated active demethylation. TDG is the major glycosylase that removes fC or caC, generated by TET proteins. Alternatively, the NEIL glycosylases can excise oxidized cytosines, but less efficiently. Both pathways are completed by the BER machinery.

reporter plasmid was observed in *Tdg*-/- cells. Knockout of *Mbd4,* however, had no effect on reporter gene expression, although MBD4 interacts with TET1, TET2 and TET3. Apparently, this association does not contribute to gene activation and may be involved in a different regulatory mechanism.

Our data suggest that conversion of hmC to hmU and subsequent excision by MBD4 does not play a major role in ESCs and are in line with previous studies showing that AID is unable to operate on double-stranded DNA and no detectable deamination of hmC by AID/APOBEC *in vitro* or *in vivo* could be observed (34,66). However, we cannot rule out that AID/APOBEC is involved in the TDG-dependent demethylation pathway or that this pathway contributes to TET3-dependent active demethylation in zygotes (67). Restriction enzyme-based analysis of the oxidized plasmid DNA recovered from wt E14 ESCs provided evidence that conversion of oxidized cytosine to unmodified cytosine led to gene activation. Again, this effect was dependent on TDG but not on MBD4.

Additional *in vivo* experiments with different TDG mutants showed that TDG activity and not the recruitment of unknown factors through TDG is essential for the recovery of gene expression. The specificity of TDG toward fC and caC, but not hmC, was confirmed using a newly established assay based on the ability of APEX1 to recognize glycosylase-generated abasic sites and convert them into single-strand breaks. The results also revealed that TDG activity is much higher on fC or caC than on a G○T mismatch, arguing that oxidized cytosines are the major substrate for TDG and that deamination is not necessary for gene reactivation.

Besides TDG and MBD4, we also investigated the role of the NEIL glycosylase family in TET-mediated demethylation. NEIL1, NEIL2 and NEIL3 have been shown to excise several lesions resulting from DNA oxidations, such as 5-hydroxyuracil or thymine glycol (59,68). NEIL glycosylases are bifunctional, i.e. are also able to convert abasic sites into single-strand breaks (69). *Neil3*-/- mice have been reported to be viable as well as fertile, and the expression of NEIL3 is elevated in hematopoietic tissues, suggesting a function in the immune system or hematopoiesis (70).

Since the NEIL glycosylases have been described as potential binders of oxidized cytosines (16), we tried to compensate for the loss of TDG in *Tdg*-/- cells with each of the three NEIL proteins. Indeed, we could detect a significant reactivation of reporter gene expression, although not reaching the levels of the rescue with wt TDG. We conclude that the NEIL glycosylases can also initiate BER after TET-mediated cytosine oxidation. This hypothesis was further confirmed by the fact that the accumulation of genomic fC and caC in *Tdg*-/- ESCs was far less prominent when NEIL1, 2 or 3 was overexpressed. These data clearly show that the NEIL proteins are not only capable of reactivating the oxidized reporter gene, but also of excising formylated and carboxylated cytosine in its chromatin context. Thus, the NEIL glycosylases may constitute an alternative pathway for active demethylation and reactivation of epigenetically silenced genes (Figure 6).

In summary, we show that the TET proteins interact with a set of factors involved in catalyzing the multiple steps of BER. Furthermore, we demonstrate that TDG is the main glycosylase in TET-mediated reactivation of the epigenetically silenced Oct4 promoter via the BER pathway. It would be of interest whether the activity of the TET–BER machinery differs on other promoter types, such as CpG island containing promoters. Our results also indicate that the NEIL family of glycosylases can functionally replace TDG. It remains to be elucidated to which extent the NEIL glycosylases contribute to TET-mediated demethylation and gene reactivation and how the usage of different glycosylases is regulated *in vivo*.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We would like to thank Prof. P. Schär (University of Basel) for providing *Tdg*-/- ESC and N. Nagaraj at the core facility of the MPI for Biochemistry for mass spectrometry analysis. Furthermore, we want to thank E. Schmidtmann for help with establishing the TDG/APEX1 activity assay, C. Trummer for F3H quantifications, D. Meilinger for FACS sorting and C. Mulholland for suggestions on the manuscript.

FUNDING

Deutsche Forschungsgemeinschaft [Collaborative Research Centers SFB 646/B10 and SFB 1064/A17]. C.B. is a member of IMPRS-LS; H.L. is a member of the Nanosystems Initiative Munich (NIM).
Conflict of interest. None declared.

REFERENCES

1. Wutz,A. and Barlow,D.P. (1998) Imprinting of the mouse Igf2r gene depends on an intronic CpG island. *Mol. Cell. Endocrinol.*, **140**, 9–14.
2. Bird,A. (2002) DNA methylation patterns and epigenetic memory. *Genes Dev.*, **16**, 6–21.
3. Bestor,T., Laudano,A., Mattaliano,R. and Ingram,V. (1988) Cloning and sequencing of a cDNA encoding DNA methyltransferase of mouse cells. The carboxyl-terminal domain of the mammalian enzymes is related to bacterial restriction methyltransferases. *J. Mol. Biol.*, **203**, 971–983.
4. Okano,M., Bell,D.W., Haber,D.A. and Li,E. (1999) DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell*, **99**, 247–257.
5. Bourc'his,D., Xu,G.L., Lin,C.S., Bollman,B. and Bestor,T.H. (2001) Dnmt3L and the establishment of maternal genomic imprints. *Science*, **294**, 2536–2539.
6. Ito,S., Shen,L., Dai,Q., Wu,S.C., Collins,L.B., Swenberg,J.A., He,C. and Zhang,Y. (2011) Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science*, **333**, 1300–1303.
7. Kriaucionis,S. and Heintz,N. (2009) The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. *Science*, **324**, 929–930.
8. Tahiliani,M., Koh,K.P., Shen,Y., Pastor,W.A., Bandukwala,H., Brudno,Y., Agarwal,S., Iyer,L.M., Liu,D.R., Aravind,L. *et al.* (2009) Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science*, **324**, 930–935.
9. Szwagierczak,A., Bultmann,S., Schmidt,C.S., Spada,F. and Leonhardt,H. (2010) Sensitive enzymatic quantification of 5-hydroxymethylcytosine in genomic DNA. *Nucleic Acids Res.*, **38**, e181.
10. Iyer,L.M., Tahiliani,M., Rao,A. and 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.
11. Costa,Y., Ding,J., Theunissen,T.W., Faiola,F., Hore,T.A., Shliaha,P.V., Fidalgo,M., Saunders,A., Lawrence,M., Dietmann,S. *et al.* (2013) NANOG-dependent function of TET1 and TET2 in establishment of pluripotency. *Nature*, **495**, 370–374.
12. Koh,K.P., Yabuuchi,A., Rao,S., Huang,Y., Cunniff,K., Nardone,J., Laiho,A., Tahiliani,M., Sommer,C.A., Mostoslavsky,G. *et al.* (2011) Tet1 and Tet2 regulate 5-hydroxymethylcytosine production and cell lineage specification in mouse embryonic stem cells. *Cell Stem Cell*, **8**, 200–213.
13. Piccolo,F.M., Bagci,H., Brown,K.E., Landeira,D., Soza-Ried,J., Feytout,A., Mooijman,D., Hajkova,P., Leitch,H.G., Tada,T. *et al.* (2013) Different roles for Tet1 and Tet2 proteins in reprogramming-mediated erasure of imprints induced by EGC fusion. *Mol. Cell*, **49**, 1023–1033.
14. Konstandin,N., Bultmann,S., Szwagierczak,A., Dufour,A., Ksienzyk,B., Schneider,F., Herold,T., Mulaw,M., Kakadia,P.M., Schneider,S. *et al.* (2011) Genomic 5-hydroxymethylcytosine levels correlate with TET2 mutations and a distinct global gene expression pattern in secondary acute myeloid leukemia. *Leukemia*, **25**, 1649–1652.
15. Ko,M., Huang,Y., Jankowska,A.M., Pape,U.J., Tahiliani,M., Bandukwala,H.S., An,J., Lamperti,E.D., Koh,K.P., Ganetzky,R. *et al.* (2010) Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2. *Nature*, **468**, 839–843.
16. Spruijt,C.G., Gnerlich,F., Smits,A.H., Pfaffeneder,T., Jansen,P.W., Bauer,C., Munzel,M., Wagner,M., Muller,M., Khan,F. *et al.* (2013) Dynamic readers for 5-(hydroxy)methylcytosine and its oxidized derivatives. *Cell*, **152**, 1146–1159.
17. Kellinger,M.W., Song,C.X., Chong,J., Lu,X.Y., He,C. and Wang,D. (2012) 5-formylcytosine and 5-carboxylcytosine reduce the rate and substrate specificity of RNA polymerase II transcription. *Nat. Struct. Mol. Biol.*, **19**, 831–833.
18. Maiti,A. and 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.
19. Pfaffeneder,T., Hackner,B., Truss,M., Munzel,M., Muller,M., Deiml,C.A., Hagemeier,C. and Carell,T. (2011) The discovery of 5-formylcytosine in embryonic stem cell DNA. *Angewandte Chemie*, **50**, 7008–7012.
20. Inoue,A. and Zhang,Y. (2011) Replication-dependent loss of 5-hydroxymethylcytosine in mouse preimplantation embryos. *Science*, **334**, 194.
21. Iqbal,K., Jin,S.G., Pfeifer,G.P. and Szabo,P.E. (2011) Reprogramming of the paternal genome upon fertilization involves genome-wide oxidation of 5-methylcytosine. *Proc. Natl. Acad. Sci. U.S.A.*, **108**, 3642–3647.
22. Cardoso,M.C. and Leonhardt,H. (1999) DNA methyltransferase is actively retained in the cytoplasm during early development. *J. Cell Biol.*, **147**, 25–32.
23. Guo,J.U., Su,Y., Zhong,C., Ming,G.L. and Song,H. (2011) Hydroxylation of 5-methylcytosine by TET1 promotes active DNA demethylation in the adult brain. *Cell*, **145**, 423–434.
24. Wu,S.C. and Zhang,Y. (2010) Active DNA demethylation: many roads lead to Rome. *Nat. Rev. Mol. Cell Biol.*, **11**, 607–620.
25. He,Y.F., Li,B.Z., Li,Z., Liu,P., Wang,Y., Tang,Q., Ding,J., Jia,Y., Chen,Z., Li,L. *et al.* (2011) Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science*, **333**, 1303–1307.
26. Campalans,A., Kortulewski,T., Amouroux,R., Menoni,H., Vermeulen,W. and Radicella,J.P. (2013) Distinct spatiotemporal patterns and PARP dependence of XRCC1 recruitment to single-strand break and base excision repair. *Nucleic Acids Res.*, **41**, 3115–3129.
27. Dianova,II, Sleeth,K.M., Allinson,S.L., Parsons,J.L., Breslin,C., Caldecott,K.W. and Dianov,G.L. (2004) XRCC1-DNA polymerase beta interaction is required for efficient base excision repair. *Nucleic Acids Res.*, **32**, 2550–2555.
28. Mortusewicz,O., Rothbauer,U., Cardoso,M.C. and Leonhardt,H. (2006) Differential recruitment of DNA Ligase I and III to DNA repair sites. *Nucleic Acids Res.*, **34**, 3523–3532.
29. Cortazar,D., Kunz,C., Selfridge,J., Lettieri,T., Saito,Y., MacDougall,E., Wirz,A., Schuermann,D., Jacobs,A.L., Siegrist,F. *et al.* (2011) Embryonic lethal phenotype reveals a function of TDG in maintaining epigenetic stability. *Nature*, **470**, 419–423.
30. Shen,L., Wu,H., Diep,D., Yamaguchi,S., D'Alessio,A.C., Fung,H.L., Zhang,K. and Zhang,Y. (2013) Genome-wide analysis reveals TET- and TDG-dependent 5-methylcytosine oxidation dynamics. *Cell*, **153**, 692–706.
31. Cortellino,S., Xu,J., Sannai,M., Moore,R., Caretti,E., Cigliano,A., Le Coz,M., Devarajan,K., Wessels,A., Soprano,D. *et al.* (2011) Thymine DNA glycosylase is essential for active DNA demethylation by linked deamination-base excision repair. *Cell*, **146**, 67–79.
32. Hashimoto,H., Zhang,X. and Cheng,X. (2012) Excision of thymine and 5-hydroxymethyluracil by the MBD4 DNA glycosylase domain: structural basis and implications for active DNA demethylation. *Nucleic Acids Res.*, **40**, 8276–8284.
33. Kemmerich,K., Dingler,F.A., Rada,C. and Neuberger,M.S. (2012) Germline ablation of SMUG1 DNA glycosylase causes loss of 5-hydroxymethyluracil- and UNG-backup uracil-excision activities and increases cancer predisposition of Ung-/-Msh2-/- mice. *Nucleic Acids Res.*, **40**, 6016–6025.
34. Nabel,C.S., Jia,H., Ye,Y., Shen,L., Goldschmidt,H.L., Stivers,J.T., Zhang,Y. and Kohli,R.M. (2012) AID/APOBEC deaminases disfavor modified cytosines implicated in DNA demethylation. *Nat. Chem. Biol.*, **8**, 751–758.
35. Schiesser,S., Hackner,B., Pfaffeneder,T., Muller,M., Hagemeier,C., Truss,M. and Carell,T. (2012) Mechanism and stem-cell activity of 5-carboxycytosine decarboxylation determined by isotope tracing. *Angewandte Chemie*, **51**, 6516–6520.
36. Tsukamoto,T., Hashiguchi,N., Janicki,S.M., Tumbar,T., Belmont,A.S. and Spector,D.L. (2000) Visualization of gene activity in living cells. *Nat. Cell Biol.*, **2**, 871–878.
37. Millar,C.B., Guy,J., Sansom,O.J., Selfridge,J., MacDougall,E., Hendrich,B., Keightley,P.D., Bishop,S.M., Clarke,A.R. and Bird,A. (2002) Enhanced CpG mutability and tumorigenesis in MBD4-deficient mice. *Science*, **297**, 403–405.
38. Ying,Q.L., Wray,J., Nichols,J., Batlle-Morera,L., Doble,B., Woodgett,J., Cohen,P. and Smith,A. (2008) The ground state of embryonic stem cell self-renewal. *Nature*, **453**, 519–523.
39. Ito,S., D'Alessio,A.C., Taranova,O.V., Hong,K., Sowers,L.C. and 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.
40. Szwagierczak,A., Brachmann,A., Schmidt,C.S., Bultmann,S., Leonhardt,H. and Spada,F. (2011) Characterization of PvuRts1I endonuclease as a tool to investigate genomic 5-hydroxymethylcytosine. *Nucleic Acids Res.*, **39**, 5149–5156.
41. Rothbauer,U., Zolghadr,K., Muyldermans,S., Schepers,A., Cardoso,M.C. and Leonhardt,H. (2008) A versatile nanotrap for biochemical and functional studies with fluorescent fusion proteins. *Mol. Cell. Proteomics*, **7**, 282–289.
42. Rottach,A., Kremmer,E., Nowak,D., Leonhardt,H. and Cardoso,M.C. (2008) Generation and characterization of a rat monoclonal antibody specific for multiple red fluorescent proteins. *Hybridoma*, **27**, 337–343.
43. Cox,J. and Mann,M. (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.*, **26**, 1367–1372.
44. Herce,H.D., Deng,W., Helma,J., Leonhardt,H. and Cardoso,M.C. (2013) Visualization and targeted disruption of protein interactions in living cells. *Nat. Commun.*, **4**, 2660.
45. Zolghadr,K., Rothbauer,U. and Leonhardt,H. (2012) The fluorescent two-hybrid (F2H) assay for direct analysis of protein-protein interactions in living cells. *Methods Mol. Biol.*, **812**, 275–282.
46. Hashimoto,H., Liu,Y., Upadhyay,A.K., Chang,Y., Howerton,S.B., Vertino,P.M., Zhang,X. and Cheng,X. (2012) Recognition and potential mechanisms for replication and erasure of cytosine hydroxymethylation. *Nucleic Acids Res.*, **40**, 4841–4849.
47. Hardeland,U., Bentele,M., Jiricny,J. and Schar,P. (2000) Separating substrate recognition from base hydrolysis in human thymine DNA glycosylase by mutational analysis. *J. Biol. Chem.*, **275**, 33449–33456.
48. Hashimoto,H., Zhang,X. and Cheng,X. (2013) Selective excision of 5-carboxylcytosine by a thymine DNA glycosylase mutant. *J. Mol. Biol.*, **425**, 971–976.
49. Hashimoto,H., Hong,S., Bhagwat,A.S., Zhang,X. and Cheng,X. (2012) Excision of 5-hydroxymethyluracil and 5-carboxylcytosine by the thymine DNA glycosylase domain: its structural basis and implications for active DNA demethylation. *Nucleic Acids Res.*, **40**, 10203–10214.
50. Neddermann,P., Gallinari,P., Lettieri,T., Schmid,D., Truong,O., Hsuan,J.J., Wiebauer,K. and Jiricny,J. (1996) Cloning and expression of human G/T mismatch-specific thymine-DNA glycosylase. *J. Biol. Chem.*, **271**, 12767–12774.
51. Wiebauer,K. and Jiricny,J. (1990) Mismatch-specific thymine DNA glycosylase and DNA polymerase beta mediate the correction of G.T mispairs in nuclear extracts from human cells. *Proc. Natl. Acad. Sci. U.S.A.*, **87**, 5842–5845.
52. Raiber,E.A., Beraldi,D., Ficz,G., Burgess,H.E., Branco,M.R., Murat,P., Oxley,D., Booth,M.J., Reik,W. and Balasubramanian,S. (2012) Genome-wide distribution of 5-formylcytosine in embryonic stem cells is associated with transcription and depends on thymine DNA glycosylase. *Genome Biol.*, **13**, R69.
53. Yildirim,O., Li,R., Hung,J.H., Chen,P.B., Dong,X., Ee,L.S., Weng,Z., Rando,O.J. and Fazzio,T.G. (2011) Mbd3/NURD complex regulates expression of 5-hydroxymethylcytosine marked genes in embryonic stem cells. *Cell*, **147**, 1498–1510.
54. Deplus,R., Delatte,B., Schwinn,M.K., Defrance,M., Mendez,J., Murphy,N., Dawson,M.A., Volkmar,M., Putmans,P., Calonne,E. *et al.* (2013) TET2 and TET3 regulate GlcNAcylation and H3K4 methylation through OGT and SET1/COMPASS. *EMBO J.*, **32**, 645–655.
55. Chen,Q., Chen,Y., Bian,C., Fujiki,R. and Yu,X. (2013) TET2 promotes histone O-GlcNAcylation during gene transcription. *Nature*, **493**, 561–564.
56. Shi,F.T., Kim,H., Lu,W., He,Q., Liu,D., Goodell,M.A., Wan,M. and Songyang,Z. (2013) Ten-eleven translocation 1 (Tet1) is regulated by O-linked N-acetylglucosamine transferase (Ogt) for target gene repression in mouse embryonic stem cells. *J. Biol. Chem.*, **288**, 20776–20784.
57. Zhang,Q., Liu,X., Gao,W., Li,P., Hou,J., Li,J. and Wong,J. (2014) Differential regulation of the ten-eleven translocation (TET) family of dioxygenases by O-linked beta-N-acetylglucosamine transferase (OGT). *J. Biol. Chem.*, **289**, 5986–5996.
58. Cartron,P.F., Nadaradjane,A., Lepape,F., Lalier,L., Gardie,B. and Vallette,F.M. (2013) Identification of TET1 Partners That Control Its DNA-Demethylating Function. *Genes Cancer*, **4**, 235–241.
59. Dou,H., Mitra,S. and 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.
60. Hazra,T.K., Kow,Y.W., Hatahet,Z., Imhoff,B., Boldogh,I., Mokkapati,S.K., Mitra,S. and Izumi,T. (2002) Identification and characterization of a novel human DNA glycosylase for repair of cytosine-derived lesions. *J. Biol. Chem.*, **277**, 30417–30420.
61. Gibson,B.A. and Kraus,W.L. (2012) New insights into the molecular and cellular functions of poly(ADP-ribose) and PARPs. *Nat. Rev. Mol. Cell Biol.*, **13**, 411–424.
62. Kubota,Y., Nash,R.A., Klungland,A., Schar,P., Barnes,D.E. and Lindahl,T. (1996) Reconstitution of DNA base excision-repair with purified human proteins: interaction between DNA polymerase beta and the XRCC1 protein. *EMBO J.*, **15**, 6662–6670.
63. Rice,P.A. (1999) Holding damaged DNA together. *Nat. Struct. Biol.*, **6**, 805–806.
64. Mortusewicz,O., Ame,J.C., Schreiber,V. and Leonhardt,H. (2007) Feedback-regulated poly(ADP-ribosyl)ation by PARP-1 is required for rapid response to DNA damage in living cells. *Nucleic Acids Res.*, **35**, 7665–7675.
65. Niehrs,C. and Schafer,A. (2012) Active DNA demethylation by Gadd45 and DNA repair. *Trends Cell Biol.*, **22**, 220–227.
66. Bransteitter,R., Pham,P., Scharff,M.D. and Goodman,M.F. (2003) Activation-induced cytidine deaminase deaminates deoxycytidine on single-stranded DNA but requires the action of RNase. *Proc. Natl. Acad. Sci. U.S.A.*, **100**, 4102–4107.
67. Santos,F., Peat,J., Burgess,H., Rada,C., Reik,W. and Dean,W. (2013) Active demethylation in mouse zygotes involves cytosine deamination and base excision repair. *Epigenet. Chromat.*, **6**, 39.
68. Hazra,T.K., Izumi,T., Boldogh,I., Imhoff,B., Kow,Y.W., Jaruga,P., Dizdaroglu,M. and Mitra,S. (2002) Identification and characterization of a human DNA glycosylase for repair of modified bases in oxidatively damaged DNA. *Proc. Natl. Acad. Sci. U.S.A.*, **99**, 3523–3528.
69. Jacobs,A.L. and Schar,P. (2012) DNA glycosylases: in DNA repair and beyond. *Chromosoma*, **121**, 1–20.
70. Torisu,K., Tsuchimoto,D., Ohnishi,Y. and Nakabeppu,Y. (2005) Hematopoietic tissue-specific expression of mouse Neil3 for endonuclease VIII-like protein. *J. Biochem.*, **138**, 763–772.