Published in final edited form as:
Nat Struct Mol Biol. 2016 February ; 23(2): 116–124. doi:10.1038/nsmb.3151.

Neil DNA glycosylases promote substrate turnover by Tdg during DNA demethylation

Lars Schomacher[#1], **Dandan Han**[#1], **Michael U. Musheev**[#1], **Khelifa Arab**[1], **Sabine Kienhöfer**[1], **Annika von Seggern**[1], and **Christof Niehrs**[1,2]

[1]Institute of Molecular Biology (IMB), Mainz, Germany

[2]Division of Molecular Embryology, German Cancer Research Center-Zentrum für Molekulare Biologie der Universität Heidelberg (DKFZ-ZMBH) Alliance, Heidelberg, Germany

[#] These authors contributed equally to this work.

Abstract

DNA 5-methylcytosine is a dynamic epigenetic mark which plays important roles in development and disease. In the Tet-Tdg demethylation pathway, methylated cytosine is iteratively oxidized by Tet dioxygenases and unmodified cytosine is restored via thymine DNA glycosylase (Tdg). Here we show that human NEIL1 and NEIL2 DNA glycosylases coordinate abasic site processing during TET–TDG DNA demethylation. NEIL1 and NEIL2 cooperate with TDG during base excision: TDG occupies the abasic site and is displaced by NEILs, which further process the baseless sugar, thereby stimulating TDG substrate turnover. In early *Xenopus* embryos Neil2 cooperates with Tdg to remove oxidized methylcytosines and to specify neural crest development together with Tet3. Thus, Neils function as AP lyases in the coordinated AP site hand-over during oxidative DNA demethylation.

Introduction

DNA 5-methylcytosine (5mC) is an epigenetic mark which plays important regulatory roles in development and is involved in disease1,2. It has become clear that DNA methylation in animal cells is reversible by an enzymatic demethylation process involving oxidation of 5mC by the Ten-eleven translocation (Tet) family of dioxygenases3–7. Tet enzymes iteratively oxidize 5mC to form 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC). The oxidation products act as intermediates in DNA demethylation by both active and passive modes (reviewed in8,9). The active mode involves

Corresponding authors: Phone: 49-6131-39-21402 / 00, Fax: 49-6131-39-21421 l.schomacher@imb-mainz.de / c.niehrs@imb-mainz.de.

Author contributions

L.S. carried out biochemical assays. D.H. conducted *Xenopus* experiments. M.U.M. performed LC-MS/MS measurements. K.A. performed experiments on HNO387 cells. L.S. and S.K. carried out protein-protein interaction assays. All authors analyzed and discussed the data. L.S. and C.N. conceived the study, designed experiments and wrote the paper.

Competing financial interests

The authors declare no competing financial interests.

Europe PMC Funders Author Manuscripts

thymine DNA glycosylase (Tdg), which recognizes and excises higher oxidation products and initiates downstream base excision repair (BER) to restore unmodified cytosines5,10. This Tet–Tdg demethylation pathway is involved in gene specific, not global DNA demethylation. Tdg null mice die around embryonic day E12.5 and show a modest 5mC increase at CpG-rich gene promoters11,12. Base resolution sequencing of Tdg knockdown cells confirmed highly gene-specific hypermethylation13–15.

Early discovery of a role for Tdg in active DNA demethylation suggested that the enzyme might promote DNA demethylation through direct excision of 5mC16. However, Tdg processes 5fC and 5caC, but not 5hmC or 5mC5,10,11. Tdg acts as a monofunctional DNA glycosylase cleaving the N-glycosidic bond between the base and the sugar at 5fC and 5caC residues within double stranded DNA to generate abasic (apyrimidinic, AP) sites. It is assumed that the AP site intermediate is subsequently processed by the canonical BER pathway: AP endonuclease 1 (Apex1) cleaves the phosphodiester bond to generate a 3'-hydroxyl and 5'-deoxyribose phosphate; DNA polymerase β removes the sugar phosphate moiety and incorporates an unmodified cytosine; DNA ligase I or IIIα seals the nick17. However, which BER components indeed function in the Tet–Tdg pathway has remained unexplored.

DNA repair intermediates are often unstable and can cause genomic instability, cell cycle arrest, cell death, or cell transformation. It is thought that DNA repair pathways are build such that during each processing step the intermediates are sequestered and protected by the appropriate repair enzyme, thereby passing them along like a baton from one enzyme to the next in a coordinated, sequential fashion18,19. During BER, the AP site is a vulnerable intermediate, which if unprotected is unstable, mutagenic and cytotoxic. The cellular risks associated with AP site intermediates are accentuated in Tet–Tdg demethylation, where parallel processing of two abasic sites in a homomethylated mCpG dyad could generate double strand breaks and moreover where demethylation may occur in 5mC tandem arrays in CpG rich islands.

Besides Tdg, the nei endonuclease VIII-like family of DNA glycosylases (Neil1-3) has been implicated in Tet-mediated DNA demethylation. Neils are bifunctional enzymes, which excise the damaged base and introduce a DNA strand break via their AP lyase activity20,21. Neil1 and Neil3 were identified as potential binders for oxidized 5mC derivatives22. Overexpression experiments suggested that NEILs function as alternative DNA glycosylases to TDG in the excision of 5fC and 5caC23, apparently consistent with the reported preference of Neils for repairing oxidized bases, including 5-hydroxyuracil, thymine glycol and 8-oxoguanine20,21,24–29. However, their requirement for processing of Tet oxidation products and their mode of action in DNA demethylation are unknown.

We conducted a screen to search for novel components of DNA demethylation and identified NEIL1 and NEIL2. Unexpectedly, biochemical analysis reveals that NEIL enzymes neither bind nor process oxidized methylcytosines directly. Instead they cooperate with TDG in AP site processing. NEIL1 and NEIL2 displace TDG from the AP site and cleave the baseless sugar, thereby overcoming TDG product inhibition and accelerating 5fC and 5caC turnover. Knockdown experiments in early *Xenopus* embryos corroborate that Neil2 is required for

the removal of genomic 5fC and 5caC and for neural crest development together with Tdg and Tet3. Neils are, hence, involved as AP lyases in the coordinated substrate hand-over during oxidative DNA demethylation.

Results

NEIL1 and NEIL2 are required for removal of 5fC and 5caC

To identify novel components of DNA demethylation, we first analyzed the ability of HeLa cells to demodify oxidized methylcytosines. When HeLa whole cell extracts were incubated with 160 bp oligonucleotides containing a single modified cytosine residue, they replaced 5fC and 5caC efficiently with unmethylated cytosine. In a 'DNA demodification assay' we monitored replacement of cytosine derivatives by unmodified cytosine by gain of HpaII sensitivity, yielding a 79 nt product (Fig. 1a-b). In contrast, oligonucleotides containing 5mC remained unprocessed and 5hmC was only marginally converted. The demodification reaction was sequence-context independent, proceeding also with an oligonucleotide of unrelated sequence (Supplementary Fig. 1a-b). Whole cellular extracts might not mimic the *in vivo* repair scenario. Under these conditions, however, demodification of 5fC and 5caC oligos showed characteristics of short patch BER; i) in the reaction without HpaII digest ('demodification intermediate assay'), it was accompanied by occurrence of 79 and 80 nt products, corresponding to the expected cleaved AP site (79 nt) and incorporated single cytidine with an unligated 3'−OH (80 nt) intermediates, respectively (Fig. 1c), and ii) it was insensitive to the replicative DNA polymerase inhibitor Aphidicolin (Fig. 1d-e).

To test which of the eleven human DNA glycosylases that mediate base excision repair30 is involved in 5fC and 5caC processing, we down-regulated their expression with siRNAs and carried out DNA glycosylase assays (Fig. 2a-b, Supplementary Fig. 1c). Knockdown of *TDG* robustly inhibited 5fC and 5caC removal (Fig. 2c, Supplementary Fig. 2a), consistent with the notion that this DNA glycosylase plays a central role in DNA demethylation11,12. In addition, knockdown of the DNA glycosylases NEIL1 and NEIL2 inhibited 5fC and 5caC excision, notably after combined knockdown (Fig. 2d, Supplementary Fig. 2b). NEIL3 was negative in this assay and differs from NEIL1 and NEIL 2 in substrate specificity and structural features31 as well as by operating mainly as a monofunctional DNA glycosylase28. Expectedly, *TDG* siRNA and combined *NEIL1* and *NEIL2* siRNAs also inhibited the demodification of 5fC and 5caC modified oligonucleotides (Fig. 2e, Supplementary Fig. 2c).

Tdg-deficiency in mouse ES cells causes a 5- to 10-fold increase in genomic 5fC and 5caC abundancy5,13. If NEIL1 and NEIL2 function in 5fC and 5caC removal, these bases should accumulate when NEILs are depleted. Indeed, in HeLa cells *TDG*, *NEIL1* and *NEIL2* siRNAs increased the genomic levels of TET1-induced 5fC and, except for *NEIL2* depletion, also 5caC (Fig. 2f). Global 5hmC showed a moderate increase, while 5mC levels were unaffected, as expected, because TDG is involved in gene-specific, not global DNA demethylation.

NEIL2 is required for TCF21 DNA demethylation

Recently, we showed that transcriptional activation of the tumor suppressor *TCF21* is directed by the long non-coding RNA TARID and is accompanied by TET–TDG-mediated DNA demethylation32. *TCF21* is silenced and hypermethylated in the cancer cell line HNO387 but can be demethylated and activated by ectopic TARID expression (32 and Fig. 3b-c). To test if NEIL1 and NEIL2 are required for gene-specific demethylation of the *TCF21* locus we monitored the expression levels and promotor methylation of *TCF21* upon *NEIL1* and *NEIL2* depletion in HNO387 cells. *NEIL2* showed strong expression, while *NEIL1* was hardly expressed in HNO387 cells (Fig. 3a). Interestingly, TARID-triggered *TCF21* expression was blocked by *NEIL2*- but not by *NEIL1* siRNA (Fig. 3b). Moreover, TARID-induced demethylation of the promotor CpGs was impaired when *NEIL2* was down-regulated (Fig. 3c). Again, knockdown of *NEIL1* had no effect, as expected from its low expression level. Overexpression of TDG partially reversed the inhibition of *TCF21* induction and promoter hypermethylation by *NEIL2* siRNA (Fig. 3d-e). We conclude that NEIL2 is required for TET–TDG-mediated gene-specific demethylation of *TCF21*.

NEIL1 and NEIL2 promote TDG-mediated 5fC and 5caC excision

Since NEIL DNA glycosylases excise oxidized bases and were reported to bind 5mC derivatives22, our findings initially supported a model whereby they act redundantly with TDG as DNA glycosylases, to directly excise 5fC and 5caC from DNA23. To test this we purified recombinant TDG, NEIL1 and NEIL2 (Supplementary Fig. 3a) and carried out *in vitro* DNA glycosylase assays. As reported5,10, TDG but not the catalytically inactive point mutant TDG^{N140A} effectively excised 5fC and 5caC residues from oligonucleotides (Fig. 4a). Surprisingly, recombinant NEIL1 and NEIL2 showed no excision of 5fC and 5caC from double or single stranded DNA (Fig. 4b). Yet, they exhibited robust glycosylase activity towards their known substrate 5-hydroxyuracil21 and showed AP lyase activity (Supplementary Fig. 3b-d), indicating that the recombinant enzymes were active.

Furthermore, we carried out DNA binding assays with TDG, NEIL1 and NEIL2 using unmodified and hemi-modified oligonucleotides. TDG bound its substrates 5fC and 5caC, as well as its product, the AP site with several-fold higher affinity than unmodified DNA (Fig. 4c). Low affinity binding of TDG to unmodified DNA was previously interpreted as a lesion scanning mode33. In contrast, NEIL1 and NEIL2 did not discriminate between unmodified, 5fC and 5caC containing oligonucleotides (Fig. 4d-e). Interestingly, NEIL1 bound AP site containing oligonucleotides with a 3-fold higher affinity than unmodified DNA, whereas NEIL2 exhibited an overall high DNA binding affinity per se (apparent K_d of 27 nM on unmodified DNA). Binding of NEILs was cooperative towards all ligands, indicative of a switch-like mode of binding and release. We conclude that purified NEIL1 and NEIL2 neither process nor specifically bind 5fC or 5caC.

It seems paradoxical that NEIL DNA glycosylases with their known preference for oxidized lesions do not directly bind or process 5fC or 5caC, neither in double- nor single-stranded DNA. How then do NEILs function in 5fC and 5caC removal? We tested if NEILs cooperate with TDG in DNA glycosylase assays. Intriguingly, addition of NEIL1 or NEIL2 robustly stimulated 5fC and 5caC excision in the presence of wild type TDG but not TDG^{N140A} (Fig.

5a), arguing for a scenario in which NEILs boost base excision of TDG but not *vice versa*. Stimulation of TDG activity by NEILs was specific since recombinant SMUG1, a monofunctional DNA glycosylase, did not enhance TDG base excision (Supplementary Fig. 4a).

In multiple turnover kinetics of 5fC and 5caC excision with TDG alone, the reactions quickly levelled off (Fig. 5b). This rapid plateauing is characteristic for TDG, which is product inhibited after base release, binding and occupying the resulting AP site34–36. AP site binding is thought to shield this repair intermediate from further attack until the next general purpose BER enzyme, APEX1, is in place to proceed with strand cleavage. APEX1 is thought to displace TDG from AP sites and thereby enhances TDG turnover34,37. Interestingly, NEIL1 and NEIL2 could substitute APEX1 in this context: In the presence of NEIL1 or NEIL2, TDG processed 5fC and 5caC with enhanced steady state turnover, notably towards 5caC (Fig. 5b; Table 1).

Consistent with the ability of NEIL1 and NEIL2 to replace APEX1, TDG stimulation by recombinant APEX1 was slightly lower than for NEIL (Supplementary Fig. 4b; compare to Fig. 5a). Moreover, *APEX1* siRNA did not increase the genomic levels of TET1-induced 5fC and 5caC (Supplementary Fig. 4c), unlike *NEIL1* and *NEIL2* siRNAs. Finally, *APEX1* siRNA only marginally inhibited 5fC and 5caC-excision in a glycosylase assay, and 5fC and 5caC processing in a demodification assay (Supplementary Fig. 4d-f).

Recombinant TDG and NEIL proteins directly bound each other and with moderate affinity in microscale thermophoresis assays (apparent K_d's 110-380 nM, Fig. 5c, Supplementary Fig. 5a-b). The affinity of NEIL1 and NEIL2 towards an AP site oligonucleotide was one order of magnitude higher (apparent K_d 18 nM and 24 nM, respectively, Fig 3d-e) and similar to that of TDG towards an AP site (apparent K_d 21 nM, Fig. 4c). The moderate NEIL–TDG interaction in combination with high affinity AP site binding may allow TDG to recruit NEILs, which then displace TDG from the AP site. Indeed, NEIL1 and NEIL2 could effectively displace TDG from an AP site oligonucleotide using near equimolar stoichiometries (Fig. 5d).

Collectively, these results support a model where TDG and NEILs act in the coordinated AP site hand-over during the processing of 5fC or 5caC (Fig. 5e). TDG hydrolyses the modified base and becomes displaced by NEIL1 or NEIL2, whose lyase activity cleaves the DNA backbone, making the displacement irreversible. The resulting single nucleotide gap if bearing 3'- and 5'-phosphate termini is likely processed by polynucleotide kinase, DNA polymerase β and DNA ligase IIIα.38 to yield demethylated cytosine.

Neil2 is required for neural crest development in Xenopus

Tet and Tdg are essential for vertebrate embryogenesis11,12,39,40. To study the physiological relevance of Neil cooperation with Tdg in this context, we turned to *Xenopus laevis* embryos, where Tet3 functions in early development, notably neural crest specification, as shown by antisense morpholino (MO) oligonucleotide injection40. In early *Xenopus* embryos *neil1* expression was low during early development and increased during organogenesis as assessed by qPCR analysis (Supplementary Fig. 5a). Conversely, *neil2* was

only expressed during early development, maternally and in neurulae (Supplementary Fig. 6a). Injection of MO oligonucleotides against *neil2*, *tdg* or *tet3*40 induced a very similar phenotype, with malformed heads, and several neural crest derivatives (pharyngeal pouches, dorsal- and tail fins) reduced or missing (Fig. 6a-b). Control and *neil1* morphants were normal and coinjection of *neil1* MO did not markedly enhance the *neil2* MO phenotype (data not shown). *neil2* morphants showed no enhanced apoptosis or reduced cell proliferation (Supplementary Fig. 6b-c). The *neil2*, *tdg* and *tet3* MO phenotypes were specific because i) they were rescued by mRNA coinjection of the respective human or *X. tropicalis* homologs (Fig. 6a-b); ii) a second MO against a non-overlapping *neil2* sequence (*neil2* MO2) yielded the same neural crest deficiency phenotype as *neil2* MO (Supplementary Fig. 6d); iii) in *neil2* morphants *neil2* expression was over 10-fold induced (Supplementary Fig. 6e), suggesting that a specific cellular mechanism senses Neil2 protein reduction and tries to compensate. Specificity of *tet3* morphants was established previously and we could confirm the phenotype40.

Consistent with a neural crest defect *neil2*, *tdg* and *tet3* morphants showed down regulation of the neural crest markers *sox10*, *twist* and *slug* on the injected side of unilaterally injected embryos, while *en2* (midbrain) and *krox20* (hindbrain) were unaffected (Fig. 6c-d), with exception of *krox20*, which was reduced by *tet3* MO. The requirement of Neil2, Tdg and Tet3 for neural crest specification was direct and not indirect via inducing mesoderm: Neural crest can be directly formed from *Xenopus* animal caps without mesoderm using combined *noggin* + *wnt8* mRNA injection41. In this regime *neil2*, *tdg* and *tet3* MO inhibited *sox10*, *slug* and *twist* expression, but not *N-cam* (pan-neural) (Fig. 6e). We conclude that Neil2, Tdg and Tet3 are required for *Xenopus* neural crest specification.

Neil2 is required for 5fC and 5caC removal in Xenopus

We tested if 5fC and 5caC residues accumulate concomitant with a *neil* depletion *in vivo*. We first established quantitative monitoring of genomic 5mC, 5hmC, 5fC and 5caC by LC-MS/MS in *Xenopus* embryos. Tadpole stage embryos had a high genomic level of 5mC (~6%), two-fold higher than HEK293T cells and mouse ES cells (mESCs) (Supplementary Fig. 7a). 5hmC levels were intermediate between HEK293T cells and mouse ES cells, while 5fC and 5caC levels, which are characteristically elevated in mESCs6, both were even higher in *Xenopus* embryos.

Importantly, the levels of total genomic 5fC and 5caC in *Xenopus* animal caps were elevated by *neil2* MO, further increased by combined *neil1* and *neil2* MOs, and unaffected by *neil1* MO alone (Fig. 7a). The levels of 5mC, 5hmC (Fig. 7a) and the oxidative lesion 8 oxoguanine (8oxoG, Supplementary Fig. 7b) were unaffected, supporting the specificity of the MO effect. Moreover, the increase of 5fC and 5caC levels by *neil2* MO was effectively reduced by *TDG* mRNA coinjection (Fig. 7b). Conversely, knockdown of *tdg* expectedly increased 5fC and 5caC and this could be partially compensated by human *NEIL2* mRNA coinjection, suggesting that residual Tdg activity is enhanced by excess NEIL2. In line with incomplete *tdg* knockdown, combined *tdg* and *neil2* MOs further increased 5fC and 5caC (Fig. 7c). This functional cooperation between *neil2* and *tdg* also manifested phenotypically, where injection of subthreshold MO doses, which by themselves yielded normal embryos,

induced malformations when *neil2* was combined with *tdg* or *tet3* MOs (Fig. 7d-e). This experiment is analogous to synthetic lethality in combined mutants, an indicator of functional interaction. In summary, in *Xenopus* embryos Neil2 cooperates with Tdg to process 5fC and 5caC residues and to specify neural crest development together with Tet3. These results, however, do not rule out a function of Tdg and Neil2 in two parallel pathways for 5fC and 5caC processing *in vivo*23.

Discussion

Neils function in the coordinated processing of 5fC and 5caC

Propositions of repair-based demethylation mechanisms have been controversial because of the perceived risk of genomic instability, notably for homomethylated mCpGs in tandem arrays. Our study reveals a cooperation of human NEIL1 and NEIL2 with TDG, which functions in the coordinated processing of 5fC and 5caC. TDG is a product inhibited enzyme whose turnover is slow. This may protect against the mutagenic and cytotoxic properties of AP sites, which inhibit certain DNA polymerases, lack base-pairing information during replication, and can lead to strand breaks42. APEX1 is thought to cooperate with TDG by stimulating its turnover for a variety of mismatches and lesions34,37. Here we show that NEIL1 and NEIL2 can also promote TDG product release. NEIL1 and NEIL2 bind with high affinity to DNA containing AP sites and therefore can compete with TDG for AP site binding. Displacement of TDG by NEIL1 and NEIL2 gains directionality by the NEIL lyase activity. Thus, our study shows that Neil DNA glycosylases act in the coordinated substrate hand-over of the vulnerable AP site intermediate from Tdg. A similar Apex1 bypass reaction was demonstrated in cell free assays where NEIL1 can enhance the activity of OGG1 DNA glycosylase in the removal of 8oxoG38,43. Our data do not exclude involvement of Apex1 in Tet–Tdg DNA demethylation, for example Apex1 may process the intermediate of the Neil AP lyase reaction, the 3'-phospho-α,β-unsaturated aldehyde44.

What may be the relevance of this reaction? Neils may act preferentially on actively transcribed genes together with Tet and Tdg. NEIL2 associates with RNA polymerase II and binds to transcribed but not transcriptionally silent genes45. Hence, this preference may not only reflect the involvement of Neils in removing transcription-blocking lesions, but possibly in processing AP sites during Tet–Tdg gene activation. Along this line, NEILs bind to TET proteins23, consistent with a recruitment of Neils to sites of DNA demethylation. Hence, Neils may act specifically in context of epigenetic gene activation, but not in universal lesion processing by TDG34,37. In Arabidopsis, for example, the Apex1 homolog APE1L functions in DNA demethylation downstream of the DNA glycosylases ROS1 and DME46,47.

TDG stalling at AP sites can also be overcome by sumoylation, which potentiates the stimulatory effect of APEX1 (and presumably also NEIL1 and NEIL2) on TDG33,36. Further studies have to explore how these mechanisms are integrated in Tet–Tdg-mediated DNA demethylation.

Europe PMC Funders Author Manuscripts

A role for the Tet–Tdg–Neil module in neural crest formation

Our analysis in *Xenopus* embryos provides evidence that Tet, Tdg and Neil2 interplay is of physiological relevance for neural crest formation during early development. In addition, Tdg and Neil2 are required and cooperate to maintain normal levels of 5fC and 5caC *in vivo*. An unanswered question is if these two biological roles of Tdg and Neil2 - neural crest formation and 5fC and 5caC processing - are linked. It is generally difficult to pinpoint the role of genes with multiple functions and functional redundancies *in vivo*. For example, Tet proteins function not only enzymatically but can also act non-enzymatically in transcriptional regulation48–50. Moreover, 5hmC, 5fC and 5caC may not only act as demethylation intermediates but also as epigenetic marks in their own right since they are recognized by specific readers22. Thus, it is unclear whether the Tet mutant phenotypes are due to DNA hypermethylation, failure to set oxidized methylcytosine epigenetic marks, or defects in transcription. In addition, *Tet1* and *Tet2* single-mutant mice are viable, while a fraction of *Tet1* and *Tet2* double knockout mice die perinatally with developmental abnormalities51, indicative of functional redundancy. *Tet3* single mutants die perinatally52 and triple mutants have not been reported. Similarly, Tdg not only functions in Tet–Tdg demethylation but also as a generic mismatch DNA repair enzyme. Moreover, like Tets, TDG can also act non-enzymatically as a scaffold protein to recruit the transcriptional coactivator CBP/p300 to numerous transcription factors53. Thus, while Tdg knock out mice die in utero, the etiology is unclear.

Given similar complexities in *neil2* morphants it is difficult to know whether the observed neural crest defects are due to reduced processing of 5fC and 5caC by Tdg. For example, we cannot exclude that Neil2 functions primarily as a lesion repair DNA glycosylase and that the neural crest may be a particularly sensitive organ in this respect. However, the fact that i) *tet3, tdg* and *neil2* morphants show a similar phenotype affecting the neural crest, ii) that *neil2* morphants did not exhibit significant cell cycle arrest or apoptosis, and iii) that combination of subthreshold MO doses cooperate in producing malformed embryos suggests an at least partially shared role for these enzymes in development, which is related to gene specific 5fC and 5caC processing.

We note that in mouse, *Neil1* and *Neil3* mutants also show neural defects, albeit in adults54–56, while *Neil2* mutants have not been reported. The observed mouse phenotypes were attributed to lesion repair deficits. Our results will provide a mechanistic and physiological framework to revisit mouse *Neil* mutants to study the role of the enzymes in the context of DNA demethylation, with its many exciting biological facets, ranging from embryonic stem cell differentiation and reprogramming to cancer.

Online Methods

Cell culture and transfection

HeLa, HEK293T (ATCC numbers CCL-2 and CRL-11268) and head and neck cancer (HNO387) cells32 were cultured in Dulbecco’s modified Eagle’s medium (DMEM, Lonza) supplemented with 10% Fetal Bovine Serum Gold (FBS, PAA), 2 mM L-Glutamine

(Lonza), 100 U/ml penicillin-streptomycin (PEN-STREP, Lonza) at 37 °C in 5% CO_2. All cell lines have been tested for mycoplasma contamination and were negative.

Cells have been transiently transfected with siRNAs using Lipofectamine 2000 (Life technologies) and with plasmid DNAs using Turbofect (Thermo Scientific) according to the manufacturer's instructions. siRNA knockdown efficiencies were routinely assessed by qPCR and were at least 60%. Transfection of TDG shown in Figure 3d-e yielded ~10-fold overexpression as tested by qPCR (data not shown).

Mouse E14TG2a (ATCC number CRL-1821) embryonic stem cells (mESCs) were cultured on plates coated with 0.1% Gelatin (Millipore) in DMEM supplemented with 15% PANSera ES FBS (PAN Biotech), 2 mM L-Glutamine, 100 μM non-essential amino acids (NEAA, Gibco), 1 mM sodium pyruvate (Gibco), 100 μM 2-mercaptoethanol (Sigma), 1000 U/ml Leukemia inhibitory factor (LIF, Millipore), 100 U/ml PEN-STREP at 37 °C in 5% CO_2 and 5% O_2.

Reverse transcriptase coupled quantitative real time PCR (RT qPCR)

Total RNA was prepared by RNeasy mini kit (Qiagen) including an on-column DNase digestion according to the manufacturer's instructions. Complementary DNA (cDNA) was synthesized using SuperScript II reverse transcriptase (Life technologies). Quantitative real time PCR was performed on a LightCycler 480 (Roche) in technical duplicates using the Universal ProbeLibrary technology (Roche) including the supplier's LightCycler 480 Probes Master. Quantitative analysis was performed with LightCycler 480 software (Roche). Primer sequences and hydrolysis probe numbers are listed in Supplementary Table 1.

Expression constructs

Human *NEIL1*, *NEIL2* and *APEX1* cDNAs (BC010876.1, BC013964.2 and BC008145.1, respectively) were from the ORFeome clone collection. Human *TDG* and mutant TDG^{N140A} cDNAs (plasmids pRS210 and pRS211, respectively, coding both for N-terminally HA-tagged proteins) were a kind gift from P. Schär, Department of Biomedicine, University of Basel, Basel, Switzerland. *TDG*, TDG^{N140A}, *NEIL1*, *NEIL2* and *APEX1* cDNAs were inserted into pET-24b(+) and/or pET-28a(+) (both from Novagen) encoding C-terminally (TDG, TDG^{N140A}, NEIL1, NEIL2) and N-terminally (NEIL1, NEIL2, APEX1) hexahistidin-tagged proteins for gene expression in *E. coli*. *TDG*, *NEIL1* and *NEIL2* cDNAs were inserted into pCS2FLAG (Addgene plasmid #16331) encoding for N-terminally (TDG) and C-terminally (NEIL1, NEIL2) 2x FLAG-tagged proteins. Additionally, *NEIL1* and *NEIL2* cDNAs were inserted into pCS2HA (Addgene #16330) as C-terminal HA-tag expression constructs. Plasmid pXC1010 containing N-terminally hexahistidin-SUMO-tagged Naegleria gruberi Tet157 was a kind gift of Y. Zheng (New England Biolabs, Ipswich, Massachusetts, USA). Human TET1 catalytic domain (aa 1418-2136) was cloned from FH-TET1-pEF4 containing full length human *TET1* cDNA (kind gift from A. Rao, Division of Signaling and Gene Expression, La Jolla Institute for Allergy and Immunology, La Jolla, CA, USA, Addgene plasmid #49792) into pCS2FLAG as an N-terminal 2x FLAG-tag expression construct. *X. tropicalis tet3* cDNA was from Open Biosystems (MXT1765-99235490).

Protein expression and purification

E. coli BL21-CodonPlus(DE3)-RIL cells (Stratagene) were transformed with the respective expression construct. Cultures were grown at 37 °C to an OD_{600} of 0.6 and induced with 1 mM isopropyl β-D-thiogalactopyranoside (IPTG), followed by cultivation at 16 °C overnight. Pelleted cells were resuspended in lysis buffer (25 mM HEPES-KOH pH 7.6, 0.5 M NaCl, 40 mM Imidazol) and lysed by passage through a Constant Systems LTD cell disrupter (1.8 kbar, constant run). Cell debris were pelleted by centrifugation at 38,400 × g for 30 min and supernatant subjected to a Ni^{2+}-charged chelating sepharose Fast Flow column (GE Healthcare). Bound proteins were eluted by a stepwise gradient of imidazol (60–500 mM) in lysis buffer. Fractions containing enriched target proteins were united, desalted with HP buffer (20 mM HEPES-KOH pH 7.6, 5 mM β-mercaptoethanol) and subjected to a HiTrap Heparin HP column (GE Healthcare). Bound proteins were eluted with a linear gradient of 0 - 1.5 M NaCl in HP buffer. Fractions containing pure protein were concentrated using Vivaspin ultrafiltration devices (GE Healthare). Concentrations of proteins were determined by Bradford assay (Bio-Rad) using BSA as standard.

DNA substrates

160 bp ds oligonucleotides, where one strand contained both, a single modified CpG dinucleotide and a 5'-end fluorescent lable, were individually prepared, purified and annealed. 160 nt ss oligonucleotides were prepared by ligation of three oligonucleotides (73mer, 12mer and 75mer, for sequences see Supplementary Table 1) with T4 DNA ligase (NEB) according to manufacturer's instructions in presence of a complementary splint oligonucleotide. Ligation products were PAGE purified on a 14% denaturing polyacrylamide gel (7 M urea, 1× TBE) followed by ethanol precipitation. Purified DNA strands were hybridized to 160 bp substrates using 500 nM fluorescently-labeled strand and 750 nM complementary strand in 1x SSC. Substrates for multiple turnover kinetics were hybridized using 10 μM fluorescently-labeled 20mer and 15 μM complementary strand in 1x SSC. For oligonucleotide sequences see Supplementary Table 1.

Activity assays

Demodification assay—HeLa whole cell extracts were prepared by lysis in 10 mM HEPES-KOH pH 7.9, 1.5 mM $MgCl_2$, 10 mM KCl, 1% Nonidet P 40 (Sigma) and clearance at 600 × g for 5 min. Concentrations of cell extracts were determined by Bradford assay (Bio-Rad) using BSA as standard. Demodification assays were perfomed in 50 μl containing 50 μg cell extract and 20 nM double stranded 160 bp modified oligonucleotide substrate in 40 mM HEPES-KOH pH 7.8, 70 mM KCl, 5mM $MgCl_2$, 10 μM $ZnCl_2$, 0.5 mM DTT, 2mM ATP, 20 μM dNTPs, 20 mM phosphocreatine, 2.5 μg creatine phosphokinase (Type I, Sigma C3755), 5 μg salmon sperm DNA for 3 h at 37 °C. Reactions were stopped by phenol/chloroform extraction and ethanol precipitation of the DNA. Reaction products were in incubated with 5 units HpaII (Promega) in 7.5 μl for 1h at 37°C, or were mock treated (demodification intermediate assay). Reaction was stopped by adding 2.5 μl loading dye (95% formamide, 20 mM EDTA pH 8.0, 10 mg/ml Dextran-Blue), reaction products were heat denatured and subjected to denaturing 4% PAGE analysis using an Applied Biosystems

ABI Prism 377 DNA Sequencer according to the manufacturer's instructions. Peak areas were quantified using the software GeneScan Analysis (Applied Biosystems).

Small scale siRNA screen—HeLa cells were harvested 48h after siRNA transfection and whole cell extracts were prepared as described above. DNA glycosylase assays were performed with 50 μg cell extract and 20 nM 160 bp DNA substrate in 10 mM HEPES-KOH pH 7.4, 100 mM KCl, 10 mM EDTA, 0.5 mM DTT in a total volume of 50 μl for 60 min at 37 °C. Reactions were stopped and DNA was isolated as described above. To generate strand breaks at AP sites reaction products were incubated with 5 u recombinant APEX1 (M0282, NEB) for 30 min at 37 °C in a total volume of 10 μl. Enzymatic strand cleavage by APEX1 instead of alkali treatment of the reaction products10 proved necessary since DNA substrates containing 5-formyl- or 5-carboxylcytidine residues showed significant strand cleavage after boiling in 150 mM NaOH even in absence of any prior enzymatic incubation (data not shown). After strand cleavage, 2.5 μl loading dye were added, reaction products heat denatured and subjected to electrophoresis as described above. Reactions were stopped by adding 2.5 μl loading dye, reaction products heat denatured and subjected to gel electrophoresis as described above.

Glycosylase assays with purified proteins—Single turnover reactions with TDG, NEIL1 and NEIL2 were performed with 200 nM enzyme and 20 nM 160 bp substrate in glycosylase buffer (10 mM HEPES-KOH pH 7.4, 1 M trimethylamine-N-oxide21, 100 mM KCl, 10 mM EDTA, 0.5 mM DTT, 100 μg/ml BSA) for 20 min at 37 °C. Stimulation of TDG excision was carried out with 6 nM TDG alone or in combination with 30 nM NEIL1 or NEIL2 and 100 nM of 160 bp substrate in glycosylase buffer for 2h at 37 °C. Control reactions with mutant TDG^{N140A}, SMUG1 and APEX1 were performed essentially as depicted above. Multiple turnover kinetics were conducted with 150 nM TDG alone or in combination with 20 nM NEIL1 or 75 nM NEIL2 and 1 μM 20 bp DNA substrate (sequences see Supplementary Table 1) in glycosylase buffer at 37 °C for up to 3h. For stimulation assays the active enzyme fraction was determined and used. TDG active fraction was derived from multiple turnover kinetics43 and NEIL1 and NEIL2 active fractions were determined by AP trapping assay using cyanoborohydride28 (data not shown). Based on this assay only a minor fraction of NEIL1 and NEIL2 protein appeared active and some contribution of the 'inactive' Neil fraction in the TDG stimulation assay cannot be ruled out. Downstream processing of reaction products and analysis was essentially as described above. Data for multiple turnover kinetics were fitted by non-linear regression to the following equation43 using GraFit 7 (Erithacus Software):

$$A_t = A_\infty(1 - e^{-kt}) + vt$$

A_∞, amplitude of the burst; *k*, 1st order rate constant; *t*, time; *v*, reaction rate in the linear phase.

AP lyase assay—To test NEIL1 and NEIL2 for their ability to incise at AP sites 20 nM of a 160 bp substrate bearing a single T:G mismatch was incubated with 200 nM TDG in

glycosylase buffer for 20 min at 37 °C followed by phenol/chloroform extraction and ethanol precipitation of the DNA. After resuspension DNA was incubated with 200 nM NEIL1 or NEIL2 (C- or N-terminally hexahistidin tagged) or 20 nM APEX1 in a total volume of 10 µl for 20 min at 37 °C. Reactions were stopped by adding 2.5 µl loading dye, reaction products heat denatured and subjected to gel electrophoresis as described above.

Electrophoretic mobility shift assays (EMSAs)

End-labeling of DNA was performed with 200 nM unmodified 20mer oligonucleotide using [γ-^{32}P] ATP (Perkin Elmer) and T4 polynucleotide kinase (NEB) according to manufacturer's instructions. 100 nM labeled DNA was hybridized in SSC buffer (150 mM NaCl, 15 mM trisodium citrate) with 100 nM complementary oligonucleotides bearing either no modification or a single 5fC, 5caC, or a tetrahydrofuran (AP site), respectively (sequences see Supplementary Table 1). Unincorporated [γ-^{32}P] ATP was removed by G-25 Quick Spin Columns (Roche). EMSAs were performed with 20 pM duplex DNA and different amounts of enzyme in binding buffer (20 mM Tris-HCl pH 7.6, 100 mM NaCl, 1mM EDTA, 1 mM DTT, 10% glycerol, 100 µg/ml BSA) for 20 min at 4 °C or at room temperature. AP site binding competition assay was performed with 20 pM of an AP site containing double stranded oligonucleotide, preincubated with 50 µg/ml BSA (Control) or 500 nM TDG in binding buffer for 10 min at room temperature before adding 500 µg/ml BSA (Control), 250 nM NEIL1 or 750 nM NEIL2 for another 10 min. PAGE was done as described previously58 but with 10% native polyacrylamide gels. Phosphorimaging was done on a Typhoon FLA 9500 (GE Healthcare) and quantification using ImageQuant TL software (GE Healthcare). Determination of dissociation constants (K_d's) was performed by fitting data for bound duplex (%) over log [enzyme] with non-linear regression to the equations for 1-site (TDG on modified DNA) or cooperative ligand binding (TDG on unmodified DNA, NEIL1 and NEIL2 throughout) using GraFit7 (Erithacus Software). For K_d calculation and AP site competition the active enzyme fractions were used (see above).

Microscale thermophoresis

For protein/protein interactions purified TDG, NEIL1 or NEIL2 were labeled using Monolith NT Protein labeling Kit RED-NHS (NanoTemper technologies) essentially as described by the manufacturer. Binding reactions were composed of 50 nM labeled protein and increasing amounts of unlabeled recombinant NEIL1, NEIL2, TDG or SMUG1 (NEB) in MST buffer (50 mM Tris/HCl pH 7.6, 150 mM NaCl, 10 mM MgCl2, 0.05% TWEEN® 20) and incubated for 5 min at room temperature (RT). Binding was monitored in hydrophilic capillaries on a Monolith NT.115 (NanoTemper technologies) using 80% Led power, 60% MST power, 30 sec laser-on time and 5 sec laser-off time. To test microscale thermophoresis for high affinity binding 30 nM purified EGFP (STA-201, Cell Biolabs) was incubated with increasing amounts of αGFP antibody (DLN-07227, Dianova, validated for specificity prior to microscale thermophoresis) in MST buffer for 5 min at RT. Binding was conducted as described above with the exception of using standard treated capillaries and 100% Led power (Supplementary Fig 4b). Curve fittings and calculation of the dissociation constants (Kd Fit) were performed with NT Analysis software (NanoTemper technologies).

X. laevis embryo manipulation and staining

Animal experiments with *X. laevis* were approved by state authorities (Landesuntersuchungsamt Rheinland-Pfalz). No blinding or randomization was performed. No statistical method was used to predetermine sample size. Embryos were obtained by *in vitro* fertilization as described59. Human and *X. tropicalis* expression constructs as depicted above were used as templates to generate mRNAs with the MEGAscript SP6 Transcription kit (Life technologies) according to the manufacturer's instructions. Morpholino antisense oligonucleotides (sequences are depicted in Supplementary Table 1) were designed to target the start codon of the respective gene. Morpholinos and mRNAs were injected 4 times into animal blastomeres at 2-4 cell stage with a total volume of 10 nl per embryo. Total amounts of single morpholinos injected per embryo were as follows: *neil1* MO, 60ng; *neil2* MO and *neil2* MO2, 40ng; *tdg* MO, 20ng and *tet3* MO, 60ng (30 ng per isoform-specific MO). The synergy experiment was performed with reduced doses of morpholinos per embryo: *neil2* MO, 10ng; *tdg* MO, 10ng and *tet3* MO, 20ng. For *neil1* and *neil2* double knockdown, embryos were injected with a mixture of 20ng *neil1* MO and 40ng *neil2* MO. Control mRNA used for injections was *preprolactin*. Animal cap assays were performed as described60. In brief, the indicated MOs or mRNAs were animally injected in both blastomeres of two-cell stage embryos, the animal caps were dissected from the embryos at stage 8.5, cultivated further and harvested at the equivalent of stage 16 for qPCR analysis or stage 32 for mass spectrometry analysis. Whole mount *in situ* hybridization was performed as described61. *In situ* hybridization probes were generated from cDNAs for *X. laevis* neuronal marker genes using the Dig RNA labeling kit (Roche). For lineage tracing, *lacZ* mRNA was co-injected (250 pg/blastomere) and β-gal staining was performed as described62 using X-gal as substrate. TUNEL assay and whole mount immunostaining of phospho histone H3 were carried out as described63,64.

Mass spectrometry analysis of genomic DNA

Generation and purification of ^{15}N-labeled 5hmC, 5fC and 5caC—Stable isotope labeling of DNA modifications was performed by a series of enzymatic -reactions. Firstly, a DNA strand of 83 bp length was PCR amplified using ^{15}N-labeled dCTP (Silantes) instead of unlabeled dCTP in the reaction mixture. Forward primer: CTCCTCTGACTGTAACCACG, reverse primer: AGGCTTCTGGACTACCTATGC, template: CTCCTCTGACTGTAACCACGCCGGTACGTTACGATACGATTACGTAATACGATTTCGAACCGGCATAGGTAGTCCAGAAGCCT. Purified PCR product was MspI (NEB) digested to remove unlabeled primer sequences and PAGE purified on a non-denaturing 20% polyacrylamide gel. Subsequently, DNA was *in vitro* methylated with M.SssI (NEB), phenol/chloroform-purified and incubated with purified NgTet1 for 30 min as described previously57. Finally, DNA was degraded to nucleosides with nuclease P1 (Roche), snake venom phosphodiesterase (Worthington) and alkaline phosphatase (Fermentas) essentially as described65. Individual nucleosides were separated on an Agilent 1290 Infinity Binary LC system (Agilent technologies) using a ReproSil 100 C18 column (Jasco). Isotopically labeled 5hmC, 5fC and 5caC were identified by analytical HPLC in tandem with triple quadruple mass spectrometry (Agilent 6490, Agilent Technologies) and purified by

preparative HPLC. An aliquot was mixed with known concentrations of corresponding unlabeled 5hmC, 5fC and 5caC nucleosides (Berry & Associates) and the concentrations of ^{15}N-labeled nucleosides were determined after LC-MS/MS by comparing the areas of the labeled and unlabeled compounds.

Genomic DNA preparation and LC-MS/MS analysis—Genomic DNA was isolated after RNase A (Fermentas) treatment of lysed cells by DNeasy blood & tissue kit (Qiagen) according to the manufacturer's instructions followed by ethanol precipitation using ammonium acetate as salt. For the assessment of genomic 8oxoG levels DNA was isolated in the presence of desferrioxamine essentially as described66. About 1 μg of DNA was degraded to nucleosides with nuclease P1 (Roche), snake venom phosphodiesterase (Worthington) and alkaline phosphatase (Fermentas)65. An equal volume of isotopic standard mixture ($^{15}N_3$-C (Silantes), $^{2}H_3$-5mC (TRC) and self-synthesized $^{15}N_3$-5hmC, $^{15}N_3$-5fC and $^{15}N_3$-5caC, (see above) was added to the DNA and about 100 ng of total DNA was injected for LC-MS/MS analysis. Quantitative analysis was performed on an Agilent 1290 Infinity Binary LC system (Agilent technologies) using a ReproSil 100 C18 column (Jasco) coupled to an Agilent 6490 triple quadrupole mass spectrometer (Agilent technologies). Running buffers were 5 mM ammonium acetate pH 6.9 (A) and Acetonitrile (B). Separations were performed at a flow rate of 0.5 ml/min using the following gradient: 0% of solvent B from 0 min to 8 min, linear increase to 15% solvent B for the next 16 min. Washing and reconditioning of the column was performed with a flow rate of 1.0 ml/min with 15% solvent B for one minute and 100% buffer A for additional 5 min. During the last minute the flow rate was linearly decreased to the initial value of 0.5 ml/min. The detailed mass spectrometer settings as well as the multiple reaction monitoring (MRM) transitions are listed in Supplementary Table 2.

Quantification of highly abundant C and 5mC was performed using 100x diluted samples. The data were analyzed with the Agilent MassHunter Quantitative Analysis software version B.05.02 (Agilent technologies) using isotopic standards to confirm the peak identity. Areas of the integrated peaks were exported into Microsoft Excel with which the areas were normalized to the area of the corresponding isotopic standard. Absolute values for the nucleosides were calculated using linear interpolation from a standard curve. Linear interpolation was performed using the two closely matching data points from the standard curve. Standards were spiked into the mixture of isotopic standards to normalize for ionization variability. The standard curve for every nucleoside was prepared to cover the amount of the corresponding nucleoside in the DNA sample analyzed. The linearity of standard curves over each region was monitored after every run and confirmed to be between 1-0.996 (R^2-values) within a concentration range of at least two orders of magnitude. The technical standard deviation was < 7%. Standard curves were newly prepared with every new dataset.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgement

We thank U. Stapf (IMB) for technical assistance, A. Rao (Division of Signaling and Gene Expression, La Jolla Institute for Allergy and Immunology, La Jolla, CA, USA), P. Schär (Department of Biomedicine, University of Basel, Basel, Switzerland), Y. Zheng (New England Biolabs, Ipswich, Massachusetts, USA) for reagents and the IMB core facilities for technical support. This work was supported by an ERC senior investigator grant to C.N. ("DNAdemethylase"). M.U.M was supported by a Natural Sciences and Engineering Research Council of Canada Postdoctoral Fellowship (NSERC-PDF 403829-2011).

References

1. Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nat Rev Genet. 2012; 13:484–92. [PubMed: 22641018]
2. Schübeler D. Function and information content of DNA methylation. Nature. 2015; 517:321–6. [PubMed: 25592537]
3. Kriaucionis S, Heintz N. The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. Science. 2009; 324:929–30. [PubMed: 19372393]
4. Tahiliani M, et al. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. Science. 2009; 324:930–5. [PubMed: 19372391]
5. He YF, et al. Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. Science. 2011; 333:1303–7. [PubMed: 21817016]
6. Ito S, et al. Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. Science. 2011; 333:1300–3. [PubMed: 21778364]
7. Pfaffeneder T, et al. The discovery of 5-formylcytosine in embryonic stem cell DNA. Angew Chem Int Ed Engl. 2011; 50:7008–12. [PubMed: 21721093]
8. Kohli RM, Zhang Y. TET enzymes, TDG and the dynamics of DNA demethylation. Nature. 2013; 502:472–9. [PubMed: 24153300]
9. Wu SC, Zhang Y. Active DNA demethylation: many roads lead to Rome. Nat Rev Mol Cell Biol. 2010; 11:607–20. [PubMed: 20683471]
10. Maiti A, Drohat AC. Thymine DNA glycosylase can rapidly excise 5-formylcytosine and 5-carboxylcytosine: potential implications for active demethylation of CpG sites. J Biol Chem. 2011; 286:35334–8. [PubMed: 21862836]
11. Cortazar D, et al. Embryonic lethal phenotype reveals a function of TDG in maintaining epigenetic stability. Nature. 2011; 470:419–23. [PubMed: 21278727]
12. Cortellino S, et al. Thymine DNA glycosylase is essential for active DNA demethylation by linked deamination-base excision repair. Cell. 2011; 146:67–79. [PubMed: 21722948]
13. Shen L, et al. Genome-wide analysis reveals TET- and TDG-dependent 5-methylcytosine oxidation dynamics. Cell. 2013; 153:692–706. [PubMed: 23602152]
14. Song CX, et al. Genome-wide profiling of 5-formylcytosine reveals its roles in epigenetic priming. Cell. 2013; 153:678–91. [PubMed: 23602153]
15. Raiber EA, et al. Genome-wide distribution of 5-formylcytosine in embryonic stem cells is associated with transcription and depends on thymine DNA glycosylase. Genome Biol. 2012; 13:R69. [PubMed: 22902005]
16. Zhu B, et al. 5-methylcytosine-DNA glycosylase activity is present in a cloned G/T mismatch DNA glycosylase associated with the chicken embryo DNA demethylation complex. Proc Natl Acad Sci U S A. 2000; 97:5135–9. [PubMed: 10779566]
17. Fromme JC, Verdine GL. Base excision repair. Adv Protein Chem. 2004; 69:1–41. [PubMed: 15588838]
18. Hosfield DJ, et al. DNA damage recognition and repair pathway coordination revealed by the structural biochemistry of DNA repair enzymes. Prog Nucleic Acid Res Mol Biol. 2001; 68:315–47. [PubMed: 11554309]
19. Prasad R, Shock DD, Beard WA, Wilson SH. Substrate channeling in mammalian base excision repair pathways: passing the baton. J Biol Chem. 2010; 285:40479–88. [PubMed: 20952393]

20. Hazra TK, et al. Identification and characterization of a human DNA glycosylase for repair of modified bases in oxidatively damaged DNA. Proc Natl Acad Sci U S A. 2002; 99:3523–8. [PubMed: 11904416]
21. Hazra TK, et al. Identification and characterization of a novel human DNA glycosylase for repair of cytosine-derived lesions. J Biol Chem. 2002; 277:30417–20. [PubMed: 12097317]
22. Spruijt CG, et al. Dynamic readers for 5-(hydroxy)methylcytosine and its oxidized derivatives. Cell. 2013; 152:1146–59. [PubMed: 23434322]
23. Müller U, Bauer C, Siegl M, Rottach A, Leonhardt H. TET-mediated oxidation of methylcytosine causes TDG or NEIL glycosylase dependent gene reactivation. Nucleic Acids Res. 2014; 42:8592–604. [PubMed: 24948610]
24. Bandaru V, Sunkara S, Wallace SS, Bond JP. A novel human DNA glycosylase that removes oxidative DNA damage and is homologous to Escherichia coli endonuclease VIII. DNA Repair (Amst). 2002; 1:517–29. [PubMed: 12509226]
25. Liu M, et al. The mouse ortholog of NEIL3 is a functional DNA glycosylase in vitro and in vivo. Proc Natl Acad Sci U S A. 2010; 107:4925–30. [PubMed: 20185759]
26. Morland I, et al. Human DNA glycosylases of the bacterial Fpg/MutM superfamily: an alternative pathway for the repair of 8-oxoguanine and other oxidation products in DNA. Nucleic Acids Res. 2002; 30:4926–36. [PubMed: 12433996]
27. Takao M, et al. A back-up glycosylase in Nth1 knock-out mice is a functional Nei (endonuclease VIII) homologue. J Biol Chem. 2002; 277:42205–13. [PubMed: 12200441]
28. Krokeide SZ, et al. Human NEIL3 is mainly a monofunctional DNA glycosylase removing spiroimindiohydantoin and guanidinohydantoin. DNA Repair (Amst). 2013; 12:1159–64. [PubMed: 23755964]
29. Dou H, Mitra S, Hazra TK. Repair of oxidized bases in DNA bubble structures by human DNA glycosylases NEIL1 and NEIL2. J Biol Chem. 2003; 278:49679–84. [PubMed: 14522990]
30. Krokan HE, Bjoras M. Base excision repair. Cold Spring Harb Perspect Biol. 2013; 5:a012583. [PubMed: 23545420]
31. Liu M, Doublie S, Wallace SS. Neil3, the final frontier for the DNA glycosylases that recognize oxidative damage. Mutat Res. 2013; 743–744:4–11.
32. Arab K, et al. Long noncoding RNA TARID directs demethylation and activation of the tumor suppressor TCF21 via GADD45A. Mol Cell. 2014; 55:604–14. [PubMed: 25087872]
33. Steinacher R, Schär P. Functionality of human thymine DNA glycosylase requires SUMO-regulated changes in protein conformation. Curr Biol. 2005; 15:616–23. [PubMed: 15823533]
34. Waters TR, Gallinari P, Jiricny J, Swann PF. Human thymine DNA glycosylase binds to apurinic sites in DNA but is displaced by human apurinic endonuclease 1. J Biol Chem. 1999; 274:67–74. [PubMed: 9867812]
35. Hardeland U, Bentele M, Jiricny J, Schar P. Separating substrate recognition from base hydrolysis in human thymine DNA glycosylase by mutational analysis. J Biol Chem. 2000; 275:33449–56. [PubMed: 10938281]
36. Hardeland U, Steinacher R, Jiricny J, Schär P. Modification of the human thymine-DNA glycosylase by ubiquitin-like proteins facilitates enzymatic turnover. EMBO J. 2002; 21:1456–64. [PubMed: 11889051]
37. Fitzgerald ME, Drohat AC. Coordinating the initial steps of base excision repair. Apurinic/apyrimidinic endonuclease 1 actively stimulates thymine DNA glycosylase by disrupting the product complex. J Biol Chem. 2008; 283:32680–90. [PubMed: 18805789]
38. Wiederhold L, et al. AP endonuclease-independent DNA base excision repair in human cells. Mol Cell. 2004; 15:209–20. [PubMed: 15260972]
39. Dawlaty MM, et al. Loss of Tet enzymes compromises proper differentiation of embryonic stem cells. Dev Cell. 2014; 29:102–11. [PubMed: 24735881]
40. Xu Y, et al. Tet3 CXXC domain and dioxygenase activity cooperatively regulate key genes for Xenopus eye and neural development. Cell. 2012; 151:1200–13. [PubMed: 23217707]
41. LaBonne C, Bronner-Fraser M. Neural crest induction in Xenopus: evidence for a two-signal model. Development. 1998; 125:2403–14. [PubMed: 9609823]

42. Wilson DM 3rd, Barsky D. The major human abasic endonuclease: formation, consequences and repair of abasic lesions in DNA. Mutat Res. 2001; 485:283–307. [PubMed: 11585362]
43. Mokkapati SK, Wiederhold L, Hazra TK, Mitra S. Stimulation of DNA glycosylase activity of OGG1 by NEIL1: functional collaboration between two human DNA glycosylases. Biochemistry. 2004; 43:11596–604. [PubMed: 15350146]
44. Pascucci B, et al. Reconstitution of the base excision repair pathway for 7,8-dihydro-8-oxoguanine with purified human proteins. Nucleic Acids Res. 2002; 30:2124–30. [PubMed: 12000832]
45. Banerjee D, et al. Preferential repair of oxidized base damage in the transcribed genes of mammalian cells. J Biol Chem. 2011; 286:6006–16. [PubMed: 21169365]
46. Lee J, et al. AP endonucleases process 5-methylcytosine excision intermediates during active DNA demethylation in Arabidopsis. Nucleic Acids Res. 2014; 42:11408–18. [PubMed: 25228464]
47. Li Y, et al. An AP endonuclease functions in active DNA dimethylation and gene imprinting in arabidopsis. PLoS Genet. 2015; 11:e1004905. [PubMed: 25569774]
48. Williams K, Christensen J, Helin K. DNA methylation: TET proteins-guardians of CpG islands? EMBO Rep. 2012; 13:28–35. [PubMed: 22157888]
49. Wu H, et al. Dual functions of Tet1 in transcriptional regulation in mouse embryonic stem cells. Nature. 2011; 473:389–93. [PubMed: 21451524]
50. Vella P, et al. Tet proteins connect the O-linked N-acetylglucosamine transferase Ogt to chromatin in embryonic stem cells. Mol Cell. 2013; 49:645–56. [PubMed: 23352454]
51. Dawlaty MM, et al. Combined deficiency of Tet1 and Tet2 causes epigenetic abnormalities but is compatible with postnatal development. Dev Cell. 2013; 24:310–23. [PubMed: 23352810]
52. Gu TP, et al. The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. Nature. 2011; 477:606–10. [PubMed: 21892189]
53. Tini M, et al. Association of CBP/p300 acetylase and thymine DNA glycosylase links DNA repair and transcription. Mol Cell. 2002; 9:265–77. [PubMed: 11864601]
54. Canugovi C, et al. Endonuclease VIII-like 1 (NEIL1) promotes short-term spatial memory retention and protects from ischemic stroke-induced brain dysfunction and death in mice. Proc Natl Acad Sci U S A. 2012; 109:14948–53. [PubMed: 22927410]
55. Regnell CE, et al. Hippocampal adult neurogenesis is maintained by Neil3-dependent repair of oxidative DNA lesions in neural progenitor cells. Cell Rep. 2012; 2:503–10. [PubMed: 22959434]
56. Sejersted Y, et al. Endonuclease VIII-like 3 (Neil3) DNA glycosylase promotes neurogenesis induced by hypoxia-ischemia. Proc Natl Acad Sci U S A. 2011; 108:18802–7. [PubMed: 22065741]
57. Hashimoto H, et al. Structure of a Naegleria Tet-like dioxygenase in complex with 5-methylcytosine DNA. Nature. 2014; 506:391–5. [PubMed: 24390346]
58. Onizuka K, Yeo J, David SS, Beal PA. NEIL1 binding to DNA containing 2'-fluorothymidine glycol stereoisomers and the effect of editing. Chembiochem. 2012; 13:1338–48. [PubMed: 22639086]
59. Gawantka V, Delius H, Hirschfeld K, Blumenstock C, Niehrs C. Antagonizing the Spemann organizer: role of the homeobox gene Xvent-1. EMBO J. 1995; 14:6268–79. [PubMed: 8557046]
60. Villanueva S, Glavic A, Ruiz P, Mayor R. Posteriorization by FGF, Wnt, and retinoic acid is required for neural crest induction. Dev Biol. 2002; 241:289–301. [PubMed: 11784112]
61. Bradley L, Wainstock D, Sive H. Positive and negative signals modulate formation of the Xenopus cement gland. Development. 1996; 122:2739–50. [PubMed: 8787748]
62. Sive, HL.; Grainiger, RM.; Harland, RM. Early Development of Xenopus laevis: A Laboratory Manual. Cold Spring Harbor Laboratory Press; New York: 2000.
63. Saka Y, Smith JC. Spatial and temporal patterns of cell division during early Xenopus embryogenesis. Dev Biol. 2001; 229:307–18. [PubMed: 11150237]
64. Hensey C, Gautier J. Programmed cell death during Xenopus development: a spatio-temporal analysis. Dev Biol. 1998; 203:36–48. [PubMed: 9806771]
65. Kellner S, et al. Absolute and relative quantification of RNA modifications via biosynthetic isotopomers. Nucleic Acids Res. 2015; 42:e142. [PubMed: 25129236]

66. Ravanat JL, et al. Cellular background level of 8-oxo-7,8-dihydro-2’-deoxyguanosine: an isotope based method to evaluate artefactual oxidation of DNA during its extraction and subsequent work-up. Carcinogenesis. 2002; 23:1911–8. [PubMed: 12419840]

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

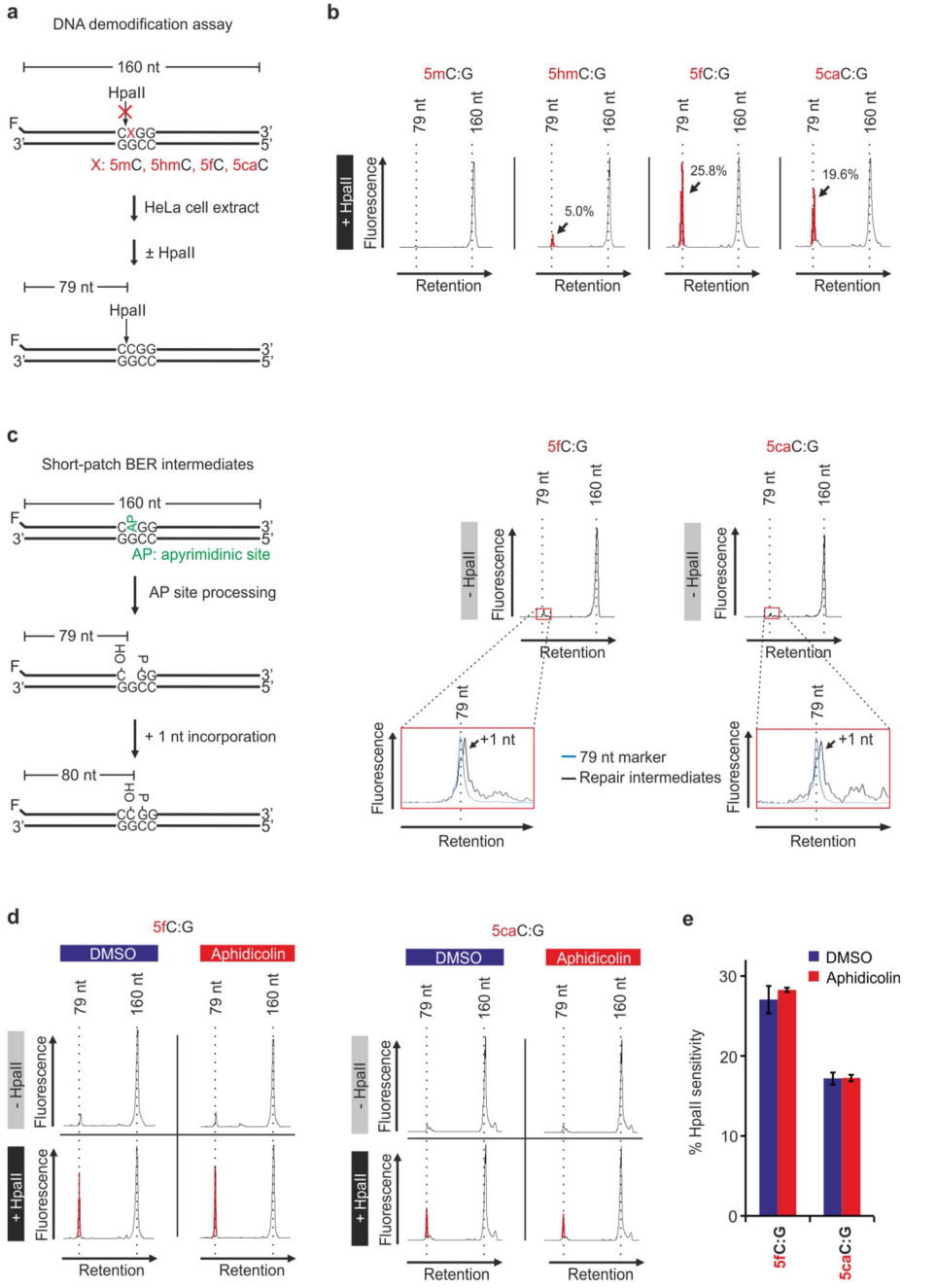

Figure 1. 5fC and 5caC are demodified by short-patch base excision repair in HeLa cell extracts.

(**a**) Scheme of DNA demodification assay. A 160 bp synthetic, fluorescently-labeled oligonucleotide containing the indicated cytosine derivative within a HpaII recognition sequence was incubated with HeLa cell extracts. Replacement of cytosine derivatives by unmodified cytosine was monitored by gain of HpaII sensitivity. Reaction without HpaII treatment served to monitor repair intermediates, which otherwise would be masked by the HpaII cleavage product.

(**b**) Denaturing gel electrophoresis of reaction products from DNA demodification assay (+ HpaII) using 50 μg HeLa extract in a 50 μl reaction on 20 nM 5mC, 5hmC, 5fC and 5caC containing oligonucleotides. Repair efficiencies are represented as integral ratio between HpaII signal peak (red) to the total fluorescent signal per electropherogram. Data are representative of three independent experiments.
(**c**) Demodification intermediate assay (– HpaII) on 5fC and 5caC containing oligonucleotides. Left, scheme and expected lengths of short-patch BER intermediates. Right top, denaturing gel electrophoresis of reaction products. Repair intermediates are boxed in red. Right bottom, magnification of intermediates. A 79mer marker oligonucleotide (blue) was overlaid with the reaction products (black). Data are representative of three independent experiments.
(**d**) Demodification assay (± HpaII treatment as indicated) as in **b** but in absence (DMSO, carrier) or presence of Aphidicolin. Repair products are highlighted in red. Data are representative of three independent experiments.
(**e**) Quantification of HpaII sensitivity (repair products) in absence or presence of Aphidicolin. Error bars, s.d. (n = 3 independent demodification assays).

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

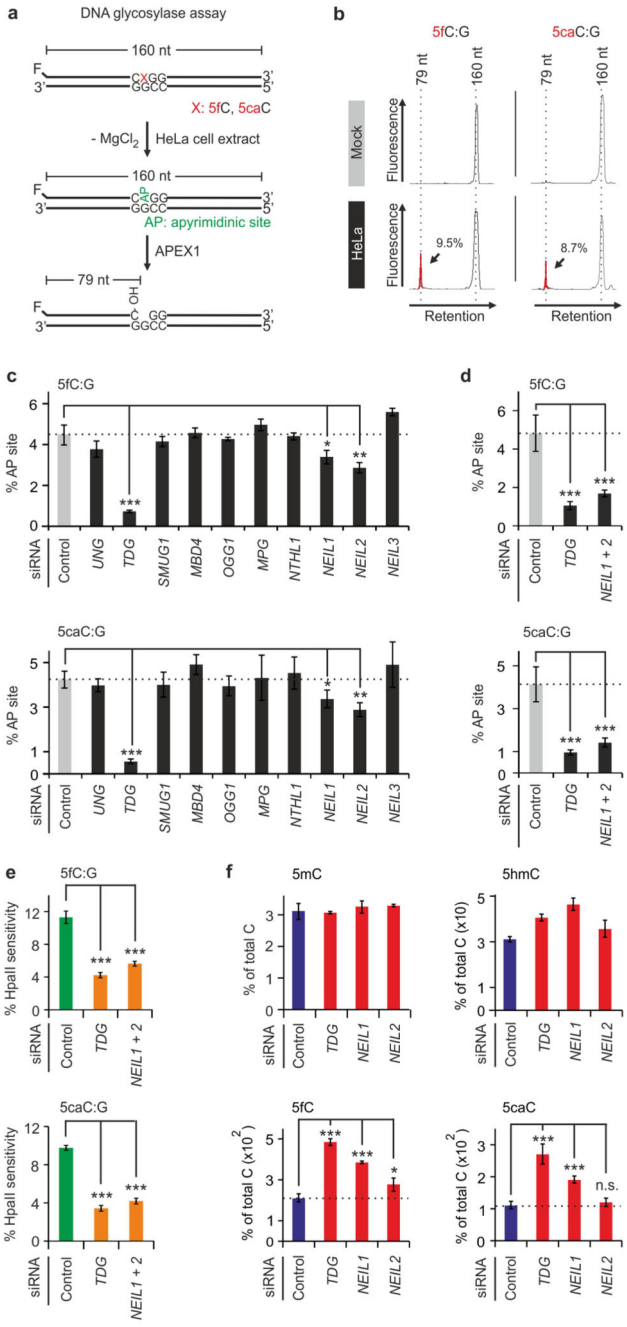

Figure 2. NEIL1 and NEIL2 are required for removal of 5fC and 5caC in HeLa cells.

(**a**) Scheme of DNA glycosylase assay. A 160 bp synthetic, fluorescently-labeled oligonucleotide containing the indicated cytosine derivative was incubated for N-glycosidic bond cleavage by HeLa extract DNA glycosylases in absence of $MgCl_2$. Base excision generates an abasic site, which can be converted into a single-strand break (nucleotide position 79) by recombinant APEX1, and monitored by denaturing PAGE.

(**b**) Electropherograms of DNA glycosylase assays. Product peaks of glycosylase activities (79mers; arrows) are highlighted red with efficiencies shown as % of 79 nt peak integral

relative to the total fluorescent signal per electropherogram. Data are representative of three independent experiments.

(**c**) siRNA screen for DNA glycosylases required for 5fC and 5caC removal. Data represent percentage of abasic sites (% AP site) generated during the assay as quantified from strand cleavage products (79mer) post APEX1 treatment.

(**d**) DNA glycosylase assay using HeLa extracts depleted of *TDG* and *NEIL1* + *NEIL2* by siRNAs.

(**e**) DNA demodification assay as in Figure 1 b but with HeLa cell extracts depleted as in **d**.

(**f**) LC-MS/MS quantification of genomic cytosine modifications from HeLa cells siRNA depleted of the indicated genes. Cells were transfected with TET1 catalytic domain to elevate levels of rare demethylation intermediates and to facilitate their mass-spectrometric detection.

Error bars, s.d. (n = 3 assay repetitions (**c-e**) or cell culture transfections (**f**)). n.s., not significant. $*P < 0.05$, $**P < 0.01$, $***P < 0.005$ by two-tailed unpaired Student's t-test.

Europe PMC Funders Author Manuscripts

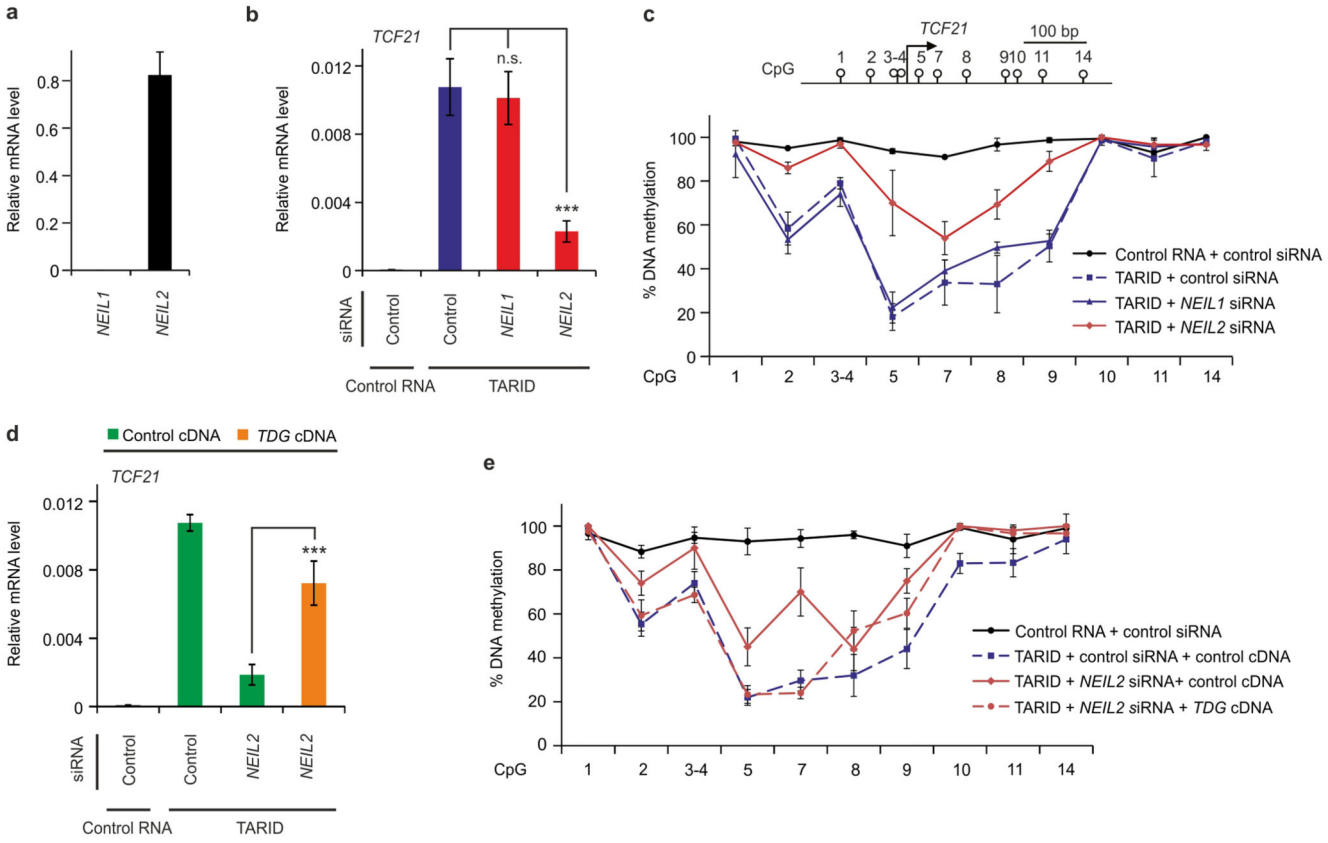

Figure 3. NEIL2 is required for demethylation of *TCF21* in HNO387 cells.

(**a**) qPCR expression analysis of *NEIL1* and *NEIL2* in HNO387 cells relative to *HPRT1* (housekeeping gene).

(**b**) qPCR expression analysis of *TCF21* relative to *HPRT1* in HNO387 cells upon transfection with control RNA (Control RNA), TARID and the indicated siRNAs.

(**c**) Top, schematic representation of *TCF21* promoter with transcription start site and CpGs analyzed. Bottom, methylation analysis of the *TCF21* promotor by MassARRAY32 using HNO387 cells transfected as in **b**.

(**d**) qPCR expression analysis of *TCF21* as in (b). HNO387 cells were transfected with control RNA, TARID RNA, control siRNA and *NEIL2* siRNA co-transfected with control or *TDG* expression plasmids as indicated.

(**e**) Methylation analysis of the *TCF21* promotor as in (**c**) using HNO387 cells transfected as in **d**.

Error bars, s.d. (n = 3 cell culture transfections). n.s., not significant. *** $P < 0.005$ by two-tailed unpaired Student's t-test.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

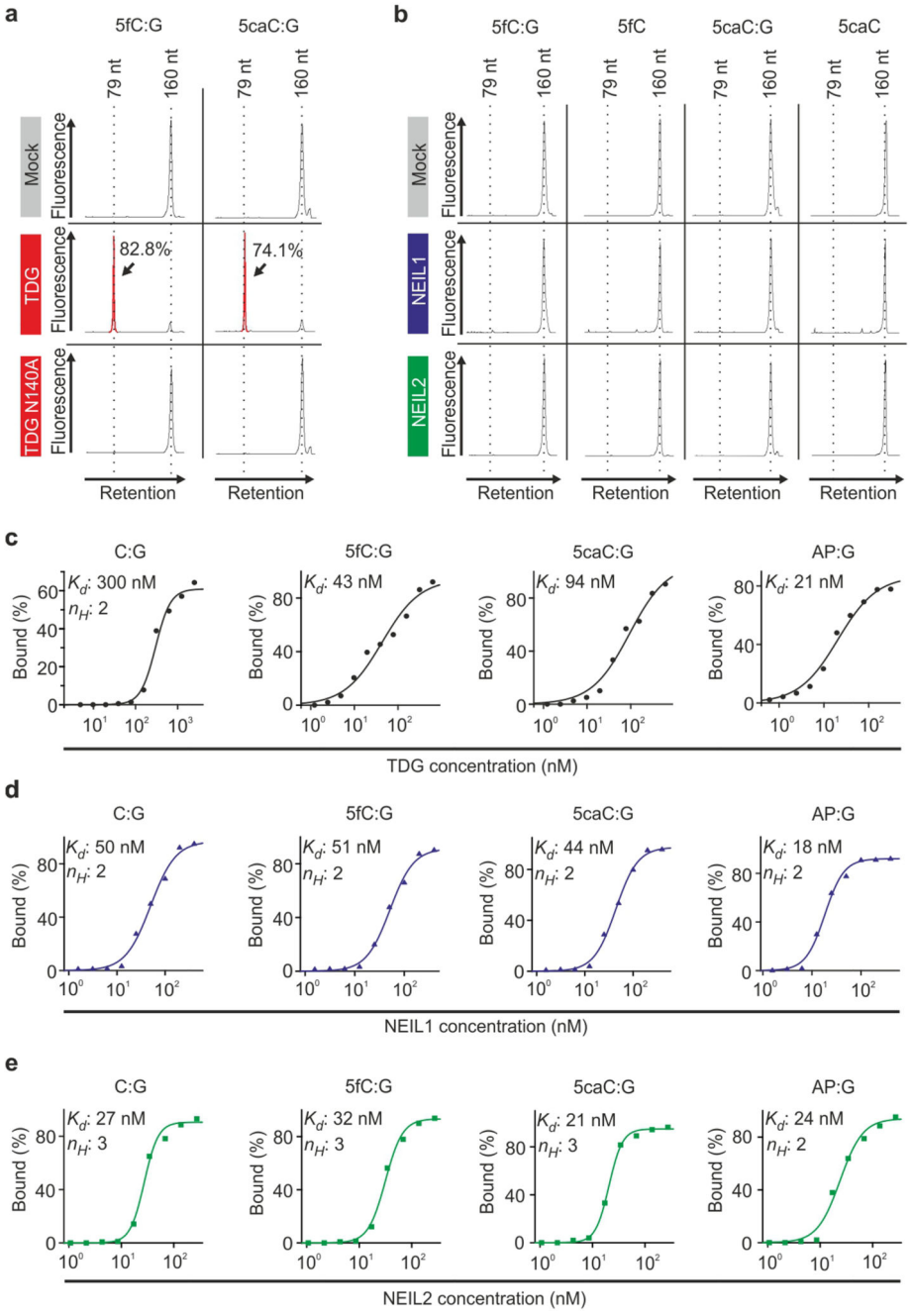

Figure 4. NEIL1 and NEIL2 do not process and bind 5fC and 5caC *in vitro*.

(**a**,**b**) Electropherograms of reaction products from DNA glycosylase assays with 20 nM ds (5fC:G; 5caC:G) and ss (5fC; 5caC) oligonucleotide substrates. Recombinant DNA glycosylase was in 10-fold excess over oligonucleotide substrate (single turnover conditions). (**a**) Human TDG (reaction products highlighted in red with efficiencies shown in %) and catalytically inactive mutant TDG^{N140A}. (**b**) Human NEIL1 and NEIL2. Data are representative of three independent experiments.

(**c-e**) DNA binding assays with TDG, NEIL1 and NEIL2. Fitted curves, calculated dissociation constants (K_d) and Hill coefficients (n_H) derived from electrophoretic mobility shift assays towards indicated ds oligonucleotides are shown. Binding assays are representative of two independent experiments with similar outcome.

Europe PMC Funders Author Manuscripts

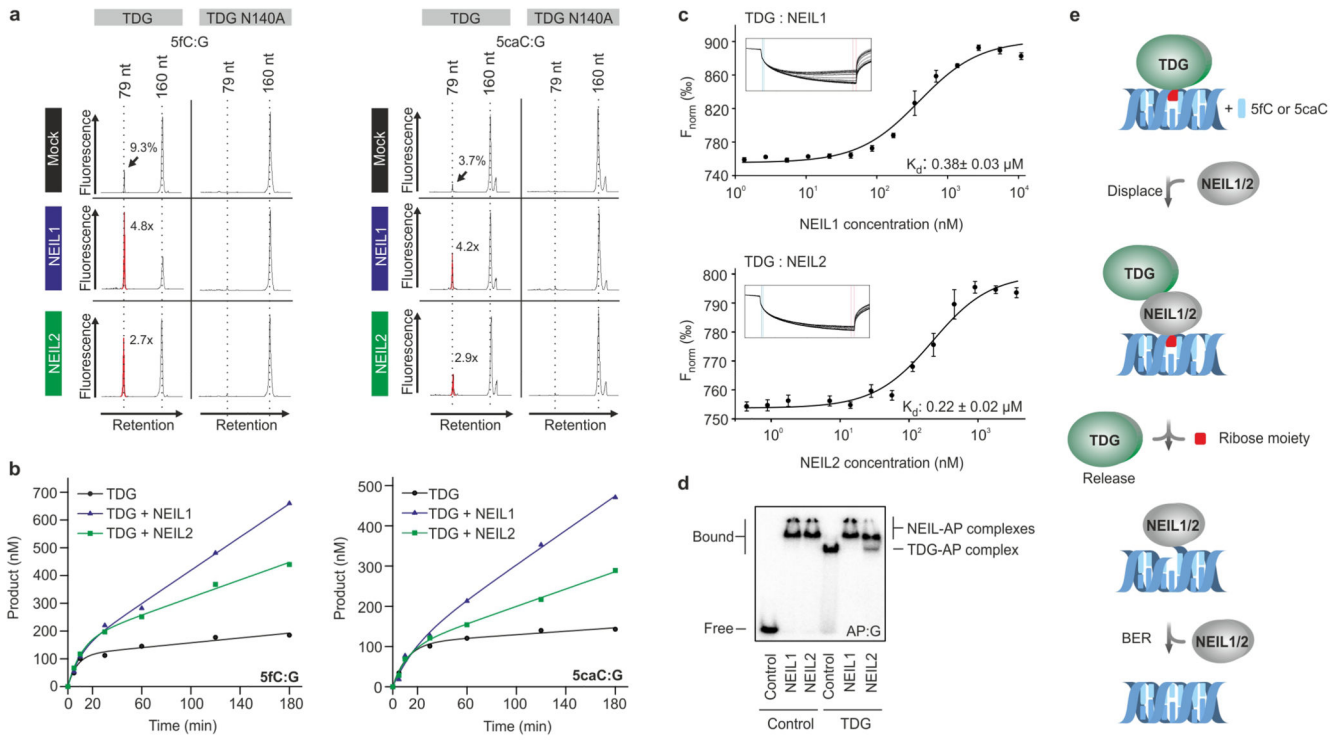

Figure 5. NEIL1 and 2 promote TDG-mediated 5fC and 5caC excision.

(**a**) DNA glycosylase assay with 100 nM oligonucleotide substrates under multiple turnover conditions with TDG and TDG^{N140A} in absence or presence of NEIL1 or NEIL2. Stimulation of base excision in presence of NEIL1 or NEIL2 is highlighted by red peaks with fold changes (left subpanels). Data are representative of three independent experiments.

(**b**) TDG base release kinetics under multiple turnover conditions towards 20bp 5fC or 5caC containing oligonucleotides (rate constants are depicted in Table 1). Kinetics are representative of two independent experiments with similar outcome.

(**c**) Binding assays of TDG to NEIL1 and NEIL2. Fitted curves for each binding experiment, normalized fluorescence timetraces and calculated K_d-values are shown. Error bars, s.d. (n = 3 binding assays). K_d errors are calculated technical errors derived from curve fittings. F_{norm} (‰), normalized fluorescence per mill.

(**d**) AP site binding competition assay. Electrophoretic mobility shift assays with TDG (500 nM), NEIL1 (250 nM) and NEIL2 (750 nM) towards an AP site containing oligonucleotide (20 pM). To probe for TDG displacement the TDG-AP complex was preformed prior to NEIL addition. Control, BSA. Experiment was reproduced two times.

(**e**) Model for NEIL1 and NEIL2 in TET–TDG-mediated DNA demethylation. TDG is product inhibited and stalls after 5fC and 5caC excision at the resulting AP site (red). NEIL1 or NEIL2 (NEIL1/2) displace TDG from the AP site by transiently contacting TDG and competing for AP site binding. NEIL1/2 lyase activity generates a strand break and downstream base excision repair (BER) factors mend the lesion.

Europe PMC Funders Author Manuscripts

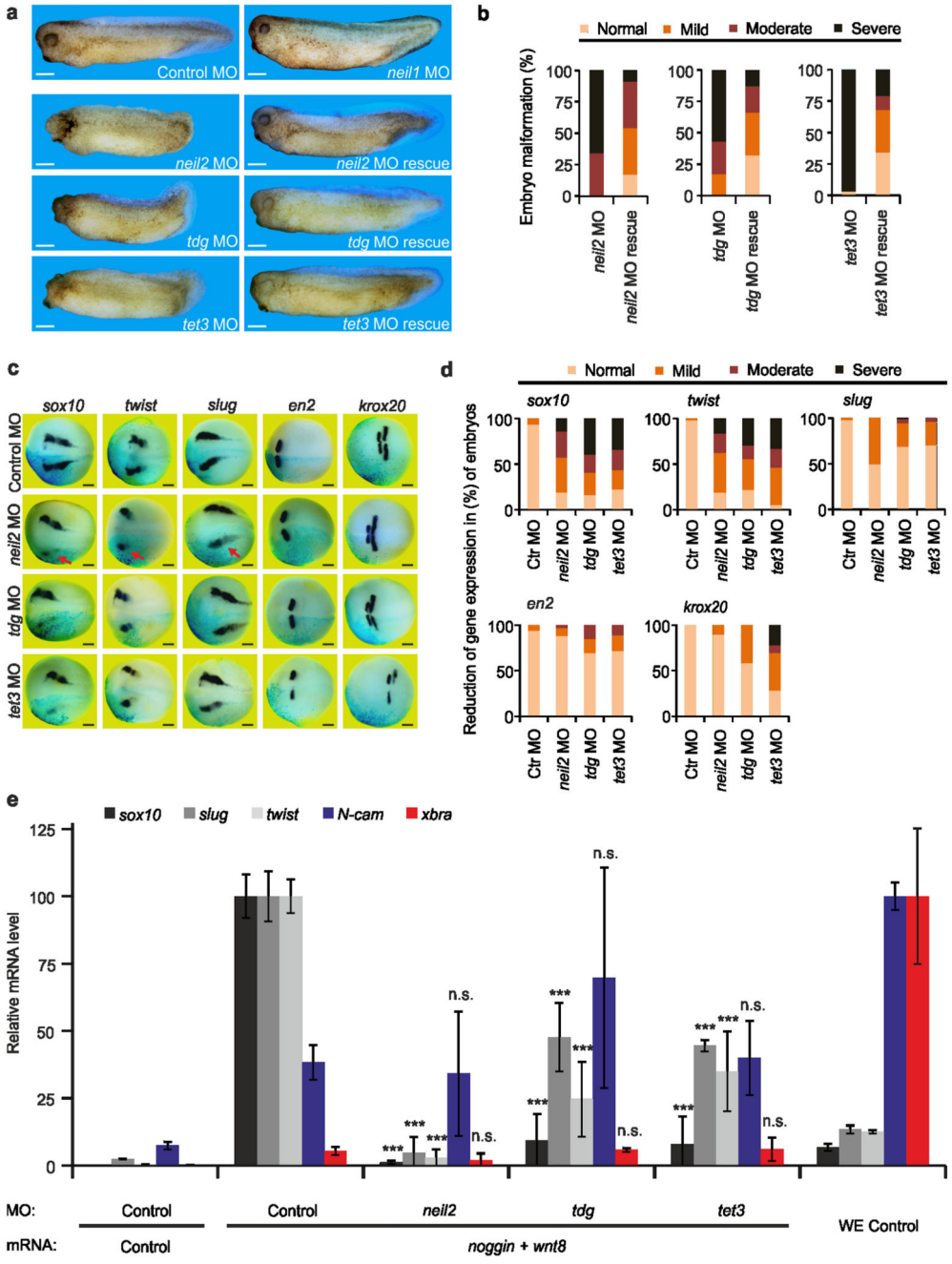

Figure 6. Neil2 is required for neural crest development in *Xenopus laevis*.

(**a**) Phenotypes of stage 34 embryos resulting from *neil1*, *neil2*, *tdg* or *tet3* morpholino (MO) injections (left). Corresponding human (*NEIL2* and *TDG*) and *X. tropicalis* (*tet3*) mRNAs were injected for rescue experiments (right). Scale bars, 200 μm

(**b**) Quantification of embryo malformations shown in **a** (n > 30 embryos per group, for details see source data).

(**c**) Expression of neural crest and brain marker genes in *neil2*, *tdg* or *tet3* morphants shown by *in situ* hybridization at stage 15. Embryos were unilaterally co-injected with *lacZ* mRNA

as lineage tracer (light blue speckles). Note reduced *sox10*, *twist* and *slug* expression in neural crest after *neil2* MO injection (red arrows). Scale bars, 200 μm

(**d**) Quantification of marker gene expression reduction shown in **c** (n > 30 embryos per group, for details see source data).

(**e**) qPCR expression analysis of the indicated marker genes in animal cap explants at stage 16. Injection with *noggin* and *wnt8* mRNA for neural crest induction and antisense morpholinos was as indicated. Expression of all markers was normalized to *histone h4* and represented as relative expression in percent of the respective control MO sample (*sox10*, *slug*, *twist*) or whole embryo (WE) control (*N-cam*, *xbra*). Error bars, s.d. (n = 3 explant or embryo batches consisting of 12, 10 and 14 individual explants and 5, 6 and 5 embryos, respectively). n.s., not significant. ***$P < 0.005$ by two-tailed unpaired Student's t-test.

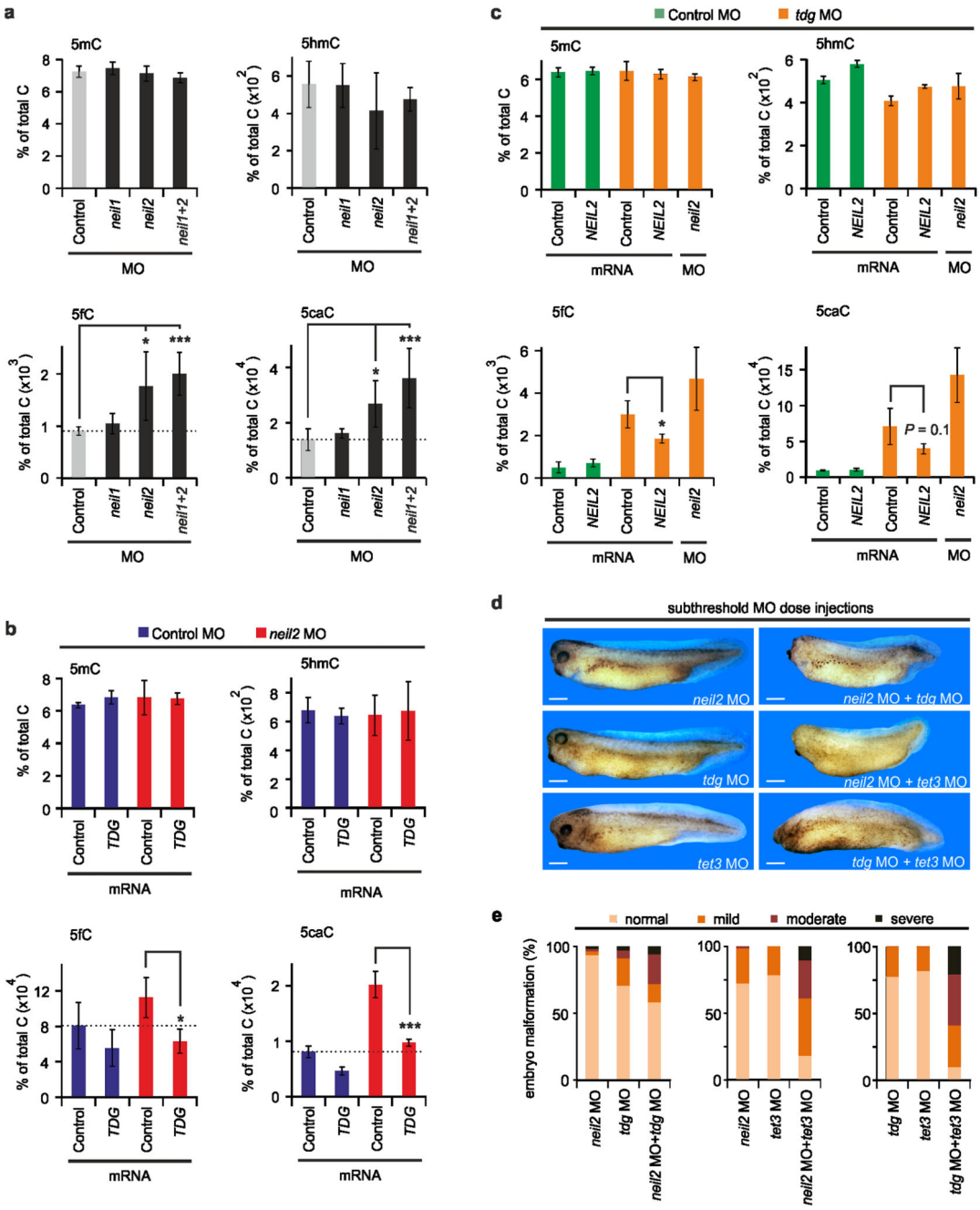

Figure 7. Neil2 is required for 5fC and 5caC removal in *Xenopus* embryos and cooperates with Tet3 and Tdg.

(**a**) LC-MS/MS analysis of genomic 5mC, 5hmC, 5fC and 5caC levels in *neil* morphant animal cap explants at stage 32. Error bars, s.d. (n = 4 explant batches each consisting of 20 animal cap explants). Grey, controls; black, morphants.

(**b**) LC-MS/MS analysis as in (a) with Control MO (blue) and *neil2* MO (red) injected animal caps without or with co-injection of human *TDG* mRNA as indicated. Error bars, s.d. (n = 3 explant batches each consisting of 20 animal cap explants).

Europe PMC Funders Author Manuscripts

(**c**) LC-MS/MS analysis as in (a), but in Control MO (green) and *tdg* MO (orange) injected animal caps (without or with co-injection of human *NEIL2* mRNA and *neil2* MO as indicated. Error bars, s.d. (n = 3 explant batches each consisting of 20 animal cap explants).
(**d**) Functional cooperation of *neil2*, *tdg and tet3* in *X. laevis* embryonic development. Subthreshold doses of *neil2, tdg or tet3* morpholinos (10 ng/embryo) were singly injected with Control MO (10 ng/embryo; top) or in combination (bottom) as indicated, to keep total MO dose in all samples constant. Embryos were fixed at stage 34. Scale bars, 200 μm
(**e**) Quantification of embryo malformations shown in **d** (n > 30 embryos per group, for details see source data).
$*P < 0.05$, $***P < 0.005$ by two-tailed unpaired Student's t-test.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Table 1

Rate constants of TDG in absence or presence of NEIL1 or NEIL2 as calculated from multiple turnover kinetics.

Rate constant errors are calculated technical errors derived from the curve fittings (see Figure 5b).

	5fC:G		5caC:G	
	pre-steady state base release	steady state base release	pre-steady state base release	steady state base release
	min^{-1}	min^{-1}	min^{-1}	min^{-1}
TDG	0.14 ± 0.04	0.0038 ± 0.0004	0.09 ± 0.02	0.0020 ± 0.0006
TDG + NEIL1	0.10 ± 0.03	0.0252 ± 0.0011	0.05 ± 0.03	0.0241 ± 0.0022
TDG + NEIL2	0.09 ± 0.02	0.0097 ± 0.0006	0.09 ± 0.02	0.0118 ± 0.0007