Published in final edited form as:

DNA Repair (Amst). 2015 August ; 32: 33–42. doi:10.1016/j.dnarep.2015.04.011.

Role of base excision repair in maintaining the genetic and epigenetic integrity of CpG sites

Alfonso Bellacosa^{a,*} and Alexander C. Drohat^{b,*}

^aCancer Epigenetics Program, Fox Chase Cancer Center, 333 Cottman Avenue, Philadelphia PA 19111, United States

^bDepartment of Biochemistry and Molecular Biology, University of Maryland School of Medicine, 108 N. Greene St., Baltimore MD 21201, United States

Abstract

Cytosine methylation at CpG dinucleotides is a central component of epigenetic regulation in vertebrates, and the base excision repair (BER) pathway is important for maintaining both the genetic stability and the methylation status of CpG sites. This perspective focuses on two enzymes that are of particular importance for the genetic and epigenetic integrity of CpG sites, Methyl Binding Domain 4 (MBD4) and Thymine DNA Glycosylase (TDG). We discuss their capacity for countering C to T mutations at CpG sites, by initiating base excision repair of G·T mismatches generated by deamination of 5-methylcytosine (5mC). We also consider their role in active DNA demethylation, including pathways that are initiated by oxidation and/or deamination of 5mC.

Keywords

Base excision repair; DNA glycosylase; Deamination; Oxidation; Demethylation; 5-methylcytosine; CpG site; Methyl binding domain 4; Thymine DNA glycosylase; G·T mispair

1. Introduction

The susceptibility of cytosine to a variety of chemical modifications [1] creates an interesting conundrum during evolution: its intrinsic reactivity allows for enzymatic modifications that could be adapted to articulate novel epigenetic functions, but it can also alter the coded genetic information and lead to mutations. For this reason, a discourse on the genetic stability of cytosine must also take into consideration the related and intertwined problem of epigenetic integrity. Thus, enzymatic modifications at the 5' position of cytosine, generally in the context of CpG dinucleotides, create 5-methylcytosine (5mC), and its

*Corresponding authors: Tel.: +1-215-728-1041; fax: +1-215-214-1737, Alfonso.Bellacosa@fccc.edu (Alfonso Bellacosa). Tel.: +1-410-706-8118; fax: +1-410-706-8297, adrohat@som.umaryland.edu (Alexander C. Drohat).

Conflict of interest statement

The authors declare no conflict of interest

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

progressively oxidized variants, 5-hydroxymethylcytosine (5hmC), 5-formylcytosine(5fC) and 5-carboxylcytosine (5caC); these bases are all important for epigenetic regulation [2]. On the other hand, deamination of cytosine and some of its 5' derivatives has the potential to create miscoding transition mutations.

There are numerous examples of DNA repair processes closely linked to transcriptional regulation; e.g. nucleotide excision repair factors involved in basal transcription or chromatin remodeling. However, only in the case of CpG sites are genomic and epigenomic stability so jointly rooted in the unique chemistry of cytosine and therefore, unsurprisingly, handled by the same repair system: base excision repair (BER). Research conducted in the past five years is beginning to answer key questions about how this repair system keeps the detrimental consequences of genomic instability of CpG sites under control, while regulating complex epigenetic processes such as DNA demethylation.

2. Two sides of cytosine chemical modifications: regulating gene expression and susceptibility to mutagenic deamination

2.1 Cytosine methylation and oxidation: features and functions

The 5' position of cytosine is not involved in hydrogen bonding with guanine and can therefore accept many substitutions with minimal to no structural impact to the DNA helix. Cytosine methylation by DNA methyltransferases occurs in bacteria, as a part of restriction-modification systems that limit bacteriophage propagation[3], in plants, as required for development, transcriptional regulation and genome architecture[4], and at low levels in fruit flies, where it is associated with genome stability and silencing of retrotransposons [5]. It is with the emergence of vertebrates that DNA methylation becomes present at high levels throughout the genome, and the benefits afforded by a fifth DNA base, 5mC, are fully realized. In mammals, 5mC is present at levels corresponding to 1% of the other bases, found mostly at CpG dinucleotides. Approximately 80–90% of CpG sites are methylated in a temporally and spatially restricted fashion [6–8], whereas the so-called CpG islands found largely in the promoters of housekeeping genes and some tissue-specific genes typically have a high density of non-methylated CpG sites [9].

In mammals, 5mC is extensively used for transcriptional repression, a function derived from its vestigial role in the suppression of parasitic sequences (transposable elements, proviral DNA, and transgenes). DNA methylation is associated with silenced genes and reinforces repressive chromatin either by directly inhibiting the binding of transactivating transcription factors to promoters and regulatory sequences, or by attracting proteins containing domains that recognize methylated DNA, such as Methyl-CpG binding domain (MBD) proteins, among others, which in turn recruit a repressive complex[10, 11]. Related to transcriptional silencing are two other functions of DNA methylation, X-chromosome inactivation and imprinting [12–16].

Importantly, cytosine methylation is deranged in cancer. CpG islands and other typically nonmethylated regions can become hypermethylated, silencing many genes including tumor suppressors[17–19]. Similarly, loss of imprinting, i.e. loss of the allelic methylation

differences at imprinted genes, has also been described in cancer[20]. In addition, methylation at some CpG islands may occur during aging[21].

Cytosine methylation is mediated by DNA methyltransferases (DNMTs). The maintenance enzyme, DNMT1, converts hemimethylated CpG sites to fully methylated ones, as needed to maintain the epigenetic mark following DNA replication[22]. The *de novo* DNA methyltransferases, DNMT3a and DNMT3b, prefer unmethylated DNA substrates and mediate the original placing of methylation marks during early development[23], but also contribute to the maintenance of methylation in adult tissues [24].

Whether the methylation mark could only be removed passively, by dilution through successive cell divisions, or also actively, via direct enzymatic activity, was under debate until a few years ago, when the molecular basis of active DNA demethylation began to emerge. The detailed pathways and biochemistry of active DNA demethylation are still being elucidated, but the overall features are relatively well defined and involve multi-step oxidation of 5mC by newly discovered enzymes followed by novel activities of known DNA base excision repair glycosylases.

Oxidation of 5mC is catalyzed by Fe^{2+} - and α -ketoglutarate-dependent dioxygenases of the Ten Eleven Translocation (TET) family. The family includes the prototypical and eponymous TET1, originally identified as a Mixed Lineage Leukemia (MLL) fusion partner in acute myelogenous leukemia characterized by the t(10;11) translocation[25, 26], and the highly related proteins TET2 and TET3 [27]. These enzymes oxidize 5mC to 5hmC[28] and successively to 5fC and 5caC (Fig. 1) [29, 30].

Whereas oxidation of 5mC represents an initial step of active demethylation, a full demethylation is achieved when 5fC and 5caC are removed by thymine DNA glycosylase (TDG) and subsequent BER leading to incorporation of cytosine (*vide infra*)[29, 31] (Fig. 1). However, it is likely that 5hmC, 5fC and 5caC are not merely intermediates of active demethylation, but also have epigenetic functions on their own. It remains to be determined how TET activity is regulated, i.e. how the oxidation of 5mC stops to 5hmC or proceeds to 5fC and 5caC.

5hmC levels correspond to 5–10% and 40% of 5mC in embryonic stem cells and Purkinje neurons of the cerebellum, respectively [28, 32]. 5hmC is very enriched at enhancers of tissue-specific genes and its levels further increase with enhancer activity and differentiation. One possibility is that 5mC oxidation acts as a switch that promotes transcription factor binding; alternatively, it is transcription factor binding that tethers TET proteins onto enhancers promoting an active transcriptional milieu.

The steady-state levels of 5fC and 5caC are much lower than those of 5hmC, corresponding to approximately 0.03% and 0.01% of 5mC levels, respectively [29, 30, 33]. Interestingly, 5fC and 5caC levels increase upon TDG knockdown, particularly at promoters and enhancers, suggesting that they may be involved in transcriptional activation[34–39]. In other words, since 5mC and 5hmC levels are very low at promoters, it is possible that rapid generation of 5mC and its immediate oxidation are actually associated with promoter activity. More evidence is needed to confirm this model that defies the traditional view of

methylation as a silencing, repressive mark; however, rapid cycles of methylation and demethylation have been described at estrogen-regulated promoters during transcriptional activation[40, 41]and would be consistent with this model.

2.2 Mutagenesis by deamination of 5mC

The spontaneous decay of DNA bases includes hydrolytic deamination of purines and pyrimidines that contain an exocyclic amino group. While hypoxanthine and xanthine are generated at a relatively slow rate by deamination of adenine and guanine, respectively, deamination of pyrimidines is potentially highly mutagenic, as it occurs at a 50-fold higher rate of approximately 200300 events per cell per day [42–45]. Deamination of C and 5mC generates G·U and G·T mismatches, respectively, which, upon DNA replication, cause C to T transition mutations. Notably, for 5mC deamination, these mutations will arise predominantly in the context of CpG sites. Thus, the benefits of a methylated genome come at a price of increased susceptibility to mutations. Uracil was excluded as a component of DNA early in evolution and is efficiently excised by uracil glycosylase activities [44]. However, the fate of 5mC CpG sites is more complex. First, the deamination rate of 5mC is approximately three times that of C [46]. Second, whereas uracil can be recognized as being foreign to DNA, enzymatic removal of thymine from G·T mispairs has to be balanced against the erroneous and mutagenic excision of thymine from A·T base pairs (*vide infra*). Similarly, deamination of 5hmC to 5hmU, which base-pairs with adenine, also has the potential to cause transition mutations at CpG sites. Notably, 5hmU can be excised by TDG, MBD4, SMUG1 and NEIL1 [47]. During evolution, reiteration of this mutagenic process is believed to have led to under-representation of the CpG dinucleotide in the mammalian genome [48]. Evolutionary consequences are an excess of transition mutations in ortholog genes of human *vs.* other primates (interspecies evolution), and elevated single nucleotide polymorphisms at CpG sites among humans (intraspecies genetic variation) [49, 50]. Similarly, the accelerated genetic evolution in somatic cells, which characterizes tumorigenesis, also exploits the spontaneous decay of pyrimidines, and it is estimated that nearly one-third of all cancer mutations in coding regions originates by deamination of cytosine and 5mC at CpG sites [51, 52].

3. Role of TDG and MBD4 in maintaining the genetic integrity of CpG sites

3.1 Overview – countering the threat posed by 5mC deamination

While the repair of G·T mismatches resulting from DNA replication errors is directed at the new strand via mismatch repair (MMR), repair of 5mC deamination must involve excision of T (not G) from the G·T mispair, and subsequent BER to restore the G·C pair. Two unrelated mammalian enzymes possess this capability: thymine DNA glycosylase (TDG) [53, 54] and methyl binding domain 4 (MBD4) [55, 56]. Consistent with a role in countering mutations caused by 5mC deamination, MBD4 and TDG have specificity for acting on mispairs arising in a CpG context, though they employ different mechanisms to do so[57]. While most DNA glycosylases remove lesions that are clearly foreign to DNA (e.g., uracil, 8-oxoguanine), MBD4 and TDG remove a *canonical* base from a mismatched pair. As such, these enzymes strike a balance between potentially conflicting needs for efficient removal of T from mutagenic G·T lesions and minimal excision of T from the vast

background of A·T pairs, which can be mutagenic and cytotoxic [58, 59]. Such a compromise could lead to suboptimal G·T glycosylase activity, as indicated for TDG [60], which in turn could potentially contribute to the high mutational frequency at CpG sites in cancer and genetic disease [61, 62]. Structural and biochemical studies have answered some questions about how these enzymes selectively excise T from G·T mispairs, but the detailed mechanism remains largely unknown.

3.2 MBD4-initiated repair of deaminated 5mC

Human MBD4 (580 residues) contains a methyl-CpG binding domain (MBD^{MBD}; 82–147) and a C-terminal glycosylase domain (426–580) separated by a region of unknown function that appears to be largely disordered (Fig. 2A, B, C) [55, 56, 63, 64]. The glycosylase domain (MBD4^{GD}) retains the catalytic activity of full-length MBD4 [55, 56, 63]. The MBD binds preferably to fully methylated CpG sites and their deamination derivatives (5mCG/5mCG, 5mCG/TG) [55, 64, 65], thereby targeting the protein to regions enriched in these sites (Fig. 2B, D). Confirming a biological role for MBD4 in countering mutations caused by 5mC deamination, deficiency of MBD4 (*Mbd4*^{-/-}) in mice causes elevated CT mutations at CpG sites, and accelerates intestinal tumor formation for mice with a predisposing mutation in the adenomatous polyposis coli gene (*Apc*^{Min/+}) [62, 66]. MBD4 is mutated in many cancers that exhibit microsatellite instability (MSI), which can lead to expression of a truncated protein (MBD4^{tru}) that contains the MBD but not the glycosylase domain [67–69]. Notably, ectopic expression of MBD4^{tru} causes a dominant negative impairment of repair in cancer cells and elevated mutations at CpG sites and other locations [67, 68]. Recent studies have advanced our understanding of how MBD4 protects against mutations at CpG sites.

3.2.1 Glycosylase domain of MBD4—Structures of the DNA-free glycosylase domain (MBD4^{GD}) confirmed that it belongs to the Endo III or helix-hairpin-helix (HhH) superfamily of repair enzymes [47, 63, 70]. MBD4^{GD} is most similar (structurally) to two HhH enzymes that recognize mismatched pairs, the 8oxoG·A mismatch glycosylase MutY [71, 72] and the thermophilic G·T mismatch glycosylase MIG [73]. However, MBD4^{GD} (~155 residues) is smaller than MutY (352 residues) and MIG (221 residues) and it lacks the iron-sulfur cluster [4Fe4S] domain found in these other enzymes.

The first structure of DNA-bound MBD4^{GD} provided a snapshot of the enzyme-product (E·P) complex, with a tetrahydrofuran (THF) a basic-site analog flipped into the active site [74]. Structures were also reported for lesion recognition (enzyme-substrate) complexes, with a catalytically inactive form of MBD4^{GD} bound to a G·T or a G·5hmU mispair (Fig. 2C, E) [75, 76]. These structures revealed that the mismatched guanine remains in the DNA helix and is contacted at N1H and N2H₂ by backbone oxygens of the enzyme (Fig. 2C, E). These contacts are not compatible with adenine and may account for some degree of the stringent specificity of MBD4 for G·T mispairs relative to A·T pairs [56, 77]. However, the role of these selective contacts has not been directly examined, and mismatch specificity may also depend on the nature of G·T mispairs, which adopt a wobble structure and are less stable (thermodynamically) and more dynamic than A·T pairs [78–80].

Remarkably, MBD4^{GD} does not contact the G-C pair immediately 3' of the flipped nucleotide, indicating that the glycosylase domain lacks specificity for acting on lesions in a CpG context. This accounts for findings that MBD4^{GD} activity does not depend on the methylation status of a CpG site that harbors a target lesion [55, 56, 77] and that activity for a G-T mispair is not dramatically altered when the flanking G-C pair (CpG context) is replaced by one of the other three base pairs [74]. By contrast, the G-T activity of TDG depends strongly on the presence of a guanine on the 3' side of the target thymine [57, 81, 82]. As discussed below, specificity for CpG sites resides in the MBD of MBD4.

The structures of lesion-recognition (ES) complexes reveal that the enzyme contacts O2 and O4 of the flipped base using a Thr hydroxyl and a backbone amide NH, respectively (Fig. 2E) [75, 76]. MBD4^{GD} also uses a Gln side chain to contact N3H and O4 of the flipped T, 5hmU, or other thymine analogue. Notably, these Gln contacts would not be compatible with cytosine or its C5-substituted analogues (5mC, 5hmC, 5fC, 5caC). Although rotation of the Gln side chain could potentially allow favorable contacts with cytosines, its orientation is stabilized by a hydrogen bond to the backbone (Fig. 2E). These observations may account, in part, for the inability of MBD4 to excise cytosine analogues. Another potential factor is the much greater stability of guanine base pairs with cytosine analogues (G·5xC) as compared to guanine mispairs with thymine and its analogues (G·T, G·5xU). Lack of MBD4 activity for oxidized forms of 5mC is not likely explained by steric effects, because the enzyme tolerates bulky substituents at the C5 position (Fig. 2E) [74, 77].

3.2.2 Targeting by the methyl-CpG-methyl-binding domain of MBD4—MBD4 localizes to methylated foci when it is ectopically expressed in mammalian cells, a function that is attributable to its methyl-CpG binding domain (MBD4^{MBD}) [55, 64, 65]. MBD4 binds more tightly to fully methylated CpG sites (5mCG/5mCG) than to hemi- or non-methylated sites, similar to other MBD proteins (MBD1, MBD2, MeCP2). Unlike these others, MBD4 also binds methylated sites that have undergone deamination, to give a G-T mispair (5mCG/TG) [55, 65], or oxidation by a Tet enzyme to give 5hmC (5mCG/5hmCG) [65]. MBD4^{MBD} binds the major groove and employs a conserved Arg and other side chains, together with a network of water molecules to recognize these sites (Fig. 2D) [65].

Observations that MBDs bind the major groove of DNA and glycosylases bind the minor groove prompted the idea that targeting of MBD4 might involve simultaneous binding of the MBD and the glycosylase domain to a single damaged site [73]. However, as noted previously [63], and shown in Figures 2B and 2C, the sharp (57°) bend in DNA imposed by the glycosylase domain [74–76] seems likely to disrupt binding of the MBD, which binds straight DNA [65]. Taken together, findings to date indicate that the two domains bind unique sites of DNA, and that the MBD targets full-length MBD4 to regions of methylated CpG sites so that the glycosylase domain can more efficiently find deaminated 5mC bases [55, 56, 63, 83, 84]. The 280-residue linker joining the domains, which is likely disordered, could allow up to 50 bp separation between binding sites for the two domains [83]. NMR studies show that the MBD exchanges rapidly between 5mCG sites on DNA, which could enhance the efficiency of lesion search by full-length MBD4 [83].

3.3 TDG-initiated BER

3.3.1 Specificity of TDG for deaminated 5mC—Human TDG is comprised of a catalytic domain (190 residues) flanked by two 110-residue regions that are largely disordered yet important for activity, protein interactions, and regulation by post-translational modifications such as acetylation, phosphorylation, and SUMO conjugation (Fig. 3A) [85–90]. TDG was discovered as an enzyme that excises T from G·T mispairs *in vitro* [91, 92], as needed to repair 5mC deamination, and subsequent studies indicate that deaminated 5mC is one of its key biological targets [57, 82]. For example, studies showing that depletion of MBD4 in mice gives an increase in CT transitions at CpG sites also suggested that another factor, likely TDG, also contributes to the repair of deaminated 5mC [62, 66]. TDG is in fact the main G·T repair activity in cell-free extracts [93]. More recently, loss of TDG was found to cause an increase in CT mutations at CpG sites in tumor cells from a patient exhibiting a deficiency in MMR [94].

Unlike the glycosylase domain of MBD4, TDG has stringent specificity for removing T bases that have a 3' guanine (5'-TG-3') [57, 82], due likely to contacts it makes with the 3'-G [95, 96]. This context specificity, unique for a glycosylase, provides full activity for G·T mispairs that arise by 5mC deamination. By contrast, activity will be weak for most G·T mispairs arising from replication errors, which is important because replication errors require action on the misincorporated nucleotide (G or T) via MMR processing of the nascent strand. Suppression of TDG activity during DNA replication is also accomplished by its degradation in S phase, as discussed below [97].

TDG activity is much greater for removing T from G·T relative to A·T pairs, as required for minimizing aberrant and mutagenic action on A·T pairs. Crystal structures suggest this specificity may derive in part from selective interactions with N1H and N2H2 of the mismatched guanine (Figs. 2E, 3B) [60]. Remarkably, MBD4 and TDG both use backbone oxygen atoms to contact the same two nitrogen atoms of the mismatched guanine. However, the weaker and more dynamic base-pairing properties of G·T mispairs relative to A·T pairs could also contribute to mismatch specificity, as noted above for MBD4. Additional studies are needed to reveal how mismatch glycosylases act selectively on G·T mispairs.

3.3.2 Regulation of TDG by post-translational modifications—The catalytic turnover of TDG is severely impeded by tight binding to its AP-DNA product [98–102], a problem that could potentially be circumvented by post-translational modification [87]. TDG is modified by small ubiquitin-like modifier (SUMO) proteins at a single Lys residue (K330, human) [87], which weakens its binding to AP-DNA and depletes its G·T glycosylase activity [87, 103, 104]. It was proposed that selective modification of *product-bound* TDG is required for efficient product release and subsequent substrate processing [87, 105]. This model gained much attention [106, 107], given that TDG is the only enzyme for which catalytic turnover is proposed to be regulated by sumoylation, and one of a few enzymes [108] whose activity is altered by SUMO. However, it was unknown whether SUMO conjugation is fast enough to hasten product release or specific for *product-bound* TDG. Recent studies show that modification by E2~SUMO is not selective for DNA-bound

TDG, but it is relatively efficient such that it could potentially stimulate product release [109].

Sumoylation of TDG could potentially serve purposes other than (or in addition to) regulating product release. TDG has a SUMO interacting motif (SIM), located adjacent to the *conjugation* site, that binds noncovalently to SUMO proteins [104, 105]. Many proteins that interact with TDG are themselves sumoylated and/or contain a SIM, including p53 and p300 [110–113]. Sumoylation of TDG could potentially stabilize its binding to proteins that have a SIM. Alternatively, sumoylation of TDG could hinder its binding to other sumoylated proteins. This is because the SUMO conjugation site of TDG (K330) flanks its own SIM, thus an intramolecular SUMO domain can occupy the SIM of TDG, as indicated by crystal structures (Fig. 3)[103, 104]. This could explain findings that sumoylation of TDG suppresses its binding to free SUMO proteins [114] and hinders its association with (and acetylation by) p300 [113].

TDG is degraded in S phase by the ubiquitin-proteasome pathway [97], in a mechanism that requires PCNA and the E3 ubiquitin ligase CRL4^{Cdt2} [115, 116]. Notably, TDG is also destroyed through the same process in response to DNA damage [115]. If not degraded, TDG slows S phase progression and cell proliferation[97, 117]. The reason for S-phase degradation of TDG is not yet clear. One possibility is that TDG may generate and bind tightly to AP sites, which could block the replication fork and/or suppress efficient downstream repair, causing DNA strand breaks [97, 115, 117, 118]. TDG might also initiate inappropriate (mutagenic) repair of G·T mismatches that arise by misincorporation of G during replication, which should be handled by MMR[97]. Alternatively, TDG might erroneously remove 5fC or 5caC during DNA replication, when heterochromatin is less condensed. In addition, it is possible that degradation of TDG is needed to regulate levels of DNA methylation during early stages of development, when most cells are undergoing DNA replication and division [115, 116].

3.3.3 TDG repair is curtailed by residues that minimize action on undamaged DNA—While numerous findings indicate that G·T mispairs are a key biological target for TDG, its activity is weaker for these lesions relative to nearly all other substrates *in vitro*[57, 98, 119–122]. Many studies indicate that access to the TDG active site is limited for thymine and other bases with a bulky C-5 substituent, due likely to steric hindrance. A structure of TDG bound to a G·U mispair suggested a steric clash between the methyl groups of a flipped thymine base and Ala145 (Fig. 3D) [60]. Removing the methyl group of Ala145 (via A145G) yielded a 13-fold *increase* in G·T activity, with no impact on G·U, as expected[60]. Consistent with these experimental findings, molecular dynamics (MD) simulations suggest that wild-type TDG stabilizes thymine in a partially flipped conformation that is not catalytically competent (Fig. 3E), whereas A145G-TDG exhibits complete thymine flipping [60]. Unexpectedly, a second active-site residue also curtails repair by TDG; G·T activity is 13-fold *higher* for H151A-TDG relative to native enzyme[60]. Why are two residues that curtail repair activity conserved in TDG enzymes (vertebrates)? One possibility is that they limit aberrant action on A·T pairs, which can be mutagenic and cytotoxic. Indeed, A·T activity is 38- and 34-fold higher for A145G- and H151A-TDG, respectively, compared to wild-type enzyme. To our knowledge, TDG is the only glycosylase for which repair activity

is sharply curtailed by conserved active-site residues. It seems likely that this reflects a balance between the conflicting needs for efficient repair and minimal erroneous action on undamaged DNA. The suboptimal G·T repair activity of TDG could potentially account, in part, for the prevalence of CT mutations at CpG sites, which are among the most frequent mutations in cancer and genetic disease [61, 123–126].

4. Role of base excision repair in active DNA demethylation

4.1 Overview

Many studies over the past two decades have suggested a potential role for MBD4 and TDG in active DNA demethylation, and the details have become increasingly clear in the last few years. Early reports suggested that both enzymes could directly remove 5mC from DNA [127, 128], but subsequent studies demonstrate that their 5mC glycosylase activity is exceedingly weak and not biologically relevant [93, 129, 130]. Notably, 5mC glycosylase activity is found in plants [131]. More recent studies indicate that TDG and MBD4 act on derivatives of 5mC resulting from active deamination and/or oxidation (Fig. 1) [2]. The demethylation pathway involving TDG has been established biochemically and biologically, whereas the same cannot be said conclusively for MBD4.

4.2 A role for MBD4 in active DNA demethylation?

Rai *et al* reported in 2008 that MBD4 mediates active DNA demethylation in zebrafish embryos, by initiating base excision repair of G·T mispairs that are generated by activation-induced deaminase (AID), and that Gadd45 facilitates a functional interaction between MBD4 and AID [132]. However, a recent report concludes that key findings of Rai *et al* are not reproducible and that there is no evidence for AID-MBD4 mediated DNA demethylation in zebrafish embryos [133]. Another study reported a role for MBD4 in hormone-induced DNA demethylation in mice [134], but this paper was subsequently retracted [135]. More recently, Sabag *et al* proposed that a TET1-AID-MBD4-Gadd45 pathway mediates DNA demethylation during reprogramming of somatic cells to generate induced pluripotent stem (iPS) cells [136]. Thus, biological studies have not conclusively demonstrated a role for MBD4 in active DNA demethylation.

In addition, biochemical studies raise questions about the role of MBD4 in DNA demethylation. Because MBD4 cannot excise 5mC, 5hmC, 5fC or 5caC [55, 56, 74, 93], an MBD4-mediated pathway would necessitate that a deaminase of the AID-Apolipoprotein B RNA-editing catalytic component (APOBEC) family convert 5mC to T, or 5hmC to 5hmU, to generate a substrate for MBD4 (G·T or G·5hmU). However, the plausibility of a pathway involving 5hmC deamination is questioned by findings that deamination of 5hmC by AID or APOBEC enzymes is not detectable *in vitro* or in cells [137, 138]. While deamination of 5hmC occurs *in vitro*, it is ~10-fold slower than deamination of C. Additionally, the preference of AID-APOBEC enzymes for single stranded DNA and for particular DNA sequences, together with findings that deaminase deficiency causes no significant developmental defects (with the caveat of possible redundancy), suggest a limited role for these enzymes in 5mC deamination [1, 2]. Even if deamination of 5mC or 5hmC is efficient *in vivo*, findings that MBD4-deficient mice are viable and exhibit no obvious developmental

or DNA methylation defects also argues against a substantial role for a deaminase-MBD4-BER pathway for DNA demethylation. Thus, compared to the TET-TDG pathway, discussed below, the (TET)-deaminase-MBD4 pathway for DNA demethylation is not well established by *in vitro* or *in vivo* studies reported to date.

4.3 TDG in active DNA demethylation and transcriptional regulation

TDG plays complex roles in DNA demethylation and transcriptional regulation, through both catalytic and non-catalytic (structural) mechanisms. Recent genetic and biochemical studies have converged to provide strong evidence for a prominent role of TDG in DNA demethylation and epigenetic regulation. First, mice with targeted inactivation of TDG in the germ line revealed embryonic lethality and developmental abnormalities associated with hypermethylation of promoter CpG islands [93, 130]. In addition, demethylation of tissue specific, developmentally regulated enhancers was impaired in the absence of TDG [93]. More importantly, knock-in of a catalytically inactive TDG variant also manifested embryonic lethality in mice, and reproduced the enhancer demethylation defect, thus providing compelling genetic evidence for an active and enzymatically-driven process for DNA demethylation [93].

In parallel, biochemical studies have begun to clarify the mechanistic role of TDG in DNA demethylation. TDG has robust glycosylase activity on 5fC and 5caC opposite G, and the levels of these epigenetic bases increase upon TDG knock-down, thus implying a necessary role for TDG acting downstream of TET enzymes in a pathway for DNA demethylation (Fig. 1) [29, 31, 96, 139, 140]. Whereas the TET-TDG-BER appears to be the main demethylation pathway, TDG may play a role in alternative pathways as well. Although the involvement of AID/APOBEC deaminases in DNA demethylation is controversial, TDG could potentially act on G-T mismatches originated by enzymatic deamination of 5mC. In fact, a physical interaction was detected between AID and TDG [93]. Although it is unlikely that AID-APOBEC deaminases mediate deamination of 5hmC to 5hmU due to the increased hindrance of 5hmC in comparison to 5mC and C [137], TDG can excise hmU [93, 129, 141].

In addition to its catalytic role, TDG may affect DNA demethylation and transcriptional activation in other ways, such as a structural adaptor/scaffold or a co-activator. In fact, it is known that active transcription contributes to maintaining CpG islands in their unmethylated state [6, 142], and it is possible that TDG is involved in promoting an open chromatin environment via protein-protein interactions. TDG interacts with DNMT3a and DNMT3b; as a result of these interactions, the two DNMTs stimulate TDG glycosylase activity whereas TDG inhibits DNMT3a [143, 144]. TDG also interacts with and is acetylated by the histone acetyl-transferases p300 and CREB-binding protein (CBP) [85], and is necessary for recruitment of p300 and MLL to active promoters [93, 130]. Due to this transcriptional co-activator role, TDG modulates the activity of several nuclear hormone receptors, including estrogen receptor alpha (ER α), retinoic acid and retinoic X receptors, and the transcriptional activators p300, CBP and p160 [85, 93, 145–147]. Thus, the absence of TDG is associated with the loss of activating histone marks (H3K9ac, H3K14ac, H3K4me2) and an increase of repressive histone marks (H3K27me3 and H3K9me3) [93, 130]. In the future, it will be

important to define more precisely the coordination of the catalytic and non-catalytic functions of TDG in transcriptional control.

5. Perspectives and Future Challenges

The last two decades have seen an increased comprehension of how BER maintains the genetic and epigenetic stability of CpG sites. Indeed, studies over the past five years have resulted in the characterization of active DNA demethylation, the identification of novel substrates of MBD4 and TDG, and the determination of DNA-bound structures of these enzymes. Nevertheless, many important challenges remain for the years ahead.

The respective roles of MBD4 and TDG in genomic and epigenomic integrity need to be established: for the shared substrates (G·T, G·U, G·hmU), do the two enzymes exhibit redundant activities or do they protect distinct areas of the genome at different stages of life (development, adult) and/or indifferent tissues?

At the moment, the available data appear to indicate a role for TDG, but not MBD4, in development and active DNA demethylation. Is it possible, however, that MBD4 has limited but defined roles in developmental control and cytosine demethylation? The same can be said for the mechanisms of DNA demethylation: most of the evidence indicates that the TET-TDG pathway is prevalent. However, it may be premature to rule out any role for deaminases in the process.

It is important to fully characterize the epigenetic roles of 5fC and 5caC, in addition to serving intermediates for cytosine demethylation, as is the case for 5hmC. This and other knowledge is likely to come from the improvement and widespread application of techniques for single-base resolution mapping of these bases, among other methods [34, 148, 149].

It is certainly an exciting time to study the connections between the genetic and epigenetic integrity of CpG sites. The next decade will allow an exploration of these open questions, coupled with a better understanding of BER functions in cancer and genetic diseases and hopefully with potential therapeutic applications of this knowledge.

Acknowledgments

We would like to thank past and present members of our respective groups for feedback and open discussions over the years; and R. Sonlin and R. Brooks for secretarial assistance. The authors apologize to colleagues for having failed to cite their work due to space constraints. Work in the authors' laboratories is supported by National Cancer Institute grant CA078412 (to AB), National Institute of General Medical Sciences grant GM072711 (to ACD), National Cancer Institute grant CA06927 to Fox Chase Cancer Center, a grant with the Pennsylvania Department of Health, and an appropriation from the Commonwealth of Pennsylvania.

References

1. Nabel CS, Manning SA, Kohli RM. The Curious Chemical Biology of Cytosine: Deamination, Methylation, and Oxidation as Modulators of Genomic Potential. *ACS Chem Biol*. 2012; 7:20–30. [PubMed: 22004246]
2. Kohli RM, Zhang Y. TET enzymes, TDG and the dynamics of DNA demethylation. *Nature*. 2013; 502:472–479. [PubMed: 24153300]

3. Kelleher JE, Raleigh EA. A novel activity in *Escherichia coli* K-12 that directs restriction of DNA modified at CG dinucleotides. *Journal of bacteriology*. 1991; 173:5220–5223. [PubMed: 1830580]
4. Zemach A, Grafi G. Methyl-CpG-binding domain proteins in plants: interpreters of DNA methylation. *Trends in plant science*. 2007; 12:80–85. [PubMed: 17208509]
5. Krauss V, Reuter G. DNA methylation in *Drosophila*--a critical evaluation. *Progress in molecular biology and translational science*. 2011; 101:177–191. [PubMed: 21507351]
6. Bird A. DNA methylation patterns and epigenetic memory. *Genes Dev*. 2002; 16:6–21. [PubMed: 11782440]
7. Colot V, Rossignol JL. Eukaryotic DNA methylation as an evolutionary device. *BioEssays : news and reviews in molecular, cellular and developmental biology*. 1999; 21:402–411.
8. Ehrlich M, Gama-Sosa MA, Huang LH, Midgett RM, Kuo KC, McCune RA, Gehrke C. Amount and distribution of 5-methylcytosine in human DNA from different types of tissues of cells. *Nucleic Acids Res*. 1982; 10:2709–2721. [PubMed: 7079182]
9. Bird AP. CpG-rich islands and the function of DNA methylation. *Nature*. 1986; 321:209–213. [PubMed: 2423876]
10. Menafrà R, Stunnenberg HG. MBD2 and MBD3: elusive functions and mechanisms. *Frontiers in genetics*. 2014; 5:428. [PubMed: 25538734]
11. Sansom OJ, Maddison K, Clarke AR. Mechanisms of disease: methyl-binding domain proteins as potential therapeutic targets in cancer. *Nat Clin Pract Oncol*. 2007; 4:305–315. [PubMed: 17464338]
12. Bird A. The essentials of DNA methylation. *Cell*. 1992; 70:5–8. [PubMed: 1377983]
13. Kass SU, Landsberger N, Wolffe AP. DNA methylation directs a time-dependent repression of transcription initiation. *Curr Biol*. 1997; 7:157–165. [PubMed: 9395433]
14. Martienssen RA, Richards EJ. DNA methylation in eukaryotes. *Curr Opin Genet Dev*. 1995; 5:234–242. [PubMed: 7613094]
15. Nan X, Campoy FJ, Bird A. MeCP2 is a transcriptional repressor with abundant binding sites in genomic chromatin. *Cell*. 1997; 88:471–481. [PubMed: 9038338]
16. Siegfried Z, Cedar H. DNA methylation: a molecular lock. *Curr Biol*. 1997; 7:R305–R307. [PubMed: 9115385]
17. Baylin SB, Jones PA. A decade of exploring the cancer epigenome - biological and translational implications. *Nat Rev Cancer*. 2011; 11:726–734. [PubMed: 21941284]
18. Jones PA, Laird PW. Cancer epigenetics comes of age. *Nat Genet*. 1999; 21:163–167. [PubMed: 9988266]
19. Issa JP. CpG island methylator phenotype in cancer. *Nat Rev Cancer*. 2004; 4:988–993. [PubMed: 15573120]
20. Jelinic P, Shaw P. Loss of imprinting and cancer. *The Journal of pathology*. 2007; 211:261–268. [PubMed: 17177177]
21. Issa JP. Aging and epigenetic drift: a vicious cycle. *The Journal of clinical investigation*. 2014; 124:24–29. [PubMed: 24382386]
22. Jeltsch A. On the enzymatic properties of Dnmt1: specificity, processivity, mechanism of linear diffusion and allosteric regulation of the enzyme. *Epigenetics*. 2006; 1:63–66. [PubMed: 17965604]
23. Okano M, Bell DW, Haber DA, Li E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell*. 1999; 99:247–257. [PubMed: 10555141]
24. Jeltsch A, Jurkowska RZ. New concepts in DNA methylation. *Trends in Biochemical Sciences*. 2014; 39:310–318. [PubMed: 24947342]
25. Ono R, Taki T, Taketani T, Taniwaki M, Kobayashi H, Hayashi Y. LCX, leukemia-associated protein with a CXXC domain, is fused to MLL in acute myeloid leukemia with trilineage dysplasia having t(10;11)(q22;q23). *Cancer Res*. 2002; 62:4075–4080. [PubMed: 12124344]
26. Lorschach RB, Moore J, Mathew S, Raimondi SC, Mukatira ST, Downing JR. TET1, a member of a novel protein family, is fused to MLL in acute myeloid leukemia containing the t(10;11)(q22;q23). *Leukemia*. 2003; 17:637–641. [PubMed: 12646957]

27. Pastor WA, Aravind L, Rao A. TETonic shift: biological roles of TET proteins in DNA demethylation and transcription. *Nature reviews Molecular cell biology*. 2013; 14:341–356. [PubMed: 23698584]
28. Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, Agarwal S, Iyer LM, Liu DR, Aravind L, Rao A. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science*. 2009; 324:930–935. [PubMed: 19372391]
29. He YF, Li BZ, Li Z, Liu P, Wang Y, Tang Q, Ding J, Jia Y, Chen Z, Li L, Sun Y, Li X, Dai Q, Song CX, Zhang K, He C, Xu GL. Tet-Mediated Formation of 5-Carboxylcytosine and Its Excision by TDG in Mammalian DNA. *Science*. 2011; 333:1303–1307. [PubMed: 21817016]
30. Ito S, Shen L, Dai Q, Wu SC, Collins LB, Swenberg JA, He C, Zhang Y. Tet Proteins Can Convert 5-Methylcytosine to 5-Formylcytosine and 5-Carboxylcytosine. *Science*. 2011; 333:1300–1303. [PubMed: 21778364]
31. 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; 286:35334–35338. [PubMed: 21862836]
32. Kriaucionis S, Heintz N. The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. *Science*. 2009; 324:929–930. [PubMed: 19372393]
33. Pfaffeneder T, Hackner B, Truss M, Munzel M, Muller M, Deiml CA, Hagemeyer C, Carell T. The Discovery of 5-Formylcytosine in Embryonic Stem Cell DNA. *Angew Chem Int Ed Engl*. 2011; 50:7008–7012. [PubMed: 21721093]
34. Shen L, Wu H, Diep D, Yamaguchi S, D'Alessio AC, Fung HL, Zhang K, Zhang Y. Genome-wide analysis reveals TET- and TDG-dependent 5-methylcytosine oxidation dynamics. *Cell*. 2013; 153:692–706. [PubMed: 23602152]
35. Wu H, Wu X, Shen L, Zhang Y. Single-base resolution analysis of active DNA demethylation using methylase-assisted bisulfite sequencing. *Nat Biotechnol*. 2014; 32:1231–1240. [PubMed: 25362244]
36. Neri F, Incarnato D, Krepelova A, Rapelli S, Anselmi F, Parlato C, Medana C, Dal Bello F, Oliviero S. Single-Base Resolution Analysis of 5-Formyl and 5-Carboxyl Cytosine Reveals Promoter DNA Methylation Dynamics. *Cell reports*. 2015
37. You C, Ji D, Dai X, Wang Y. Effects of Tet-mediated oxidation products of 5-methylcytosine on DNA transcription in vitro and in mammalian cells. *Scientific reports*. 2014; 4:7052. [PubMed: 25394478]
38. Hahn MA, Szabo PE, Pfeifer GP. 5-Hydroxymethylcytosine: a stable or transient DNA modification? *Genomics*. 2014; 104:314–323. [PubMed: 25181633]
39. Lu X, Han D, Boxuan Simen Z, Song CX, Zhang LS, Dore LC, He C. Base-resolution maps of 5-formylcytosine and 5-carboxylcytosine reveal genome-wide DNA demethylation dynamics. *Cell Res*. 2015
40. Kangaspeska S, Stride B, Metivier R, Polycarpou-Schwarz M, Ibberson D, Carmouche RP, Benes V, Gannon F, Reid G. Transient cyclical methylation of promoter DNA. *Nature*. 2008; 452:112–115. [PubMed: 18322535]
41. Metivier R, Gallais R, Tiffocche C, Le Peron C, Jurkowska RZ, Carmouche RP, Ibberson D, Barath P, Demay F, Reid G, Benes V, Jeltsch A, Gannon F, Salbert G. Cyclical DNA methylation of a transcriptionally active promoter. *Nature*. 2008; 452:45–50. [PubMed: 18322525]
42. Hickson, ID. General introduction to DNA base excision repair. In: Hickson, ID., editor. *Base Excision Repair of DNA Damage*, Landes Bioscience. Springer-Verlag; Austin, TX: 1997. p. 1-5.
43. Lindahl T, Wood RD. Quality control by DNA repair. *Science*. 1999; 286:1897–1905. [PubMed: 10583946]
44. Pearl LH, Savva R. The problem with pyrimidines. *Nat Struct Biol*. 1996; 3:485–487. [PubMed: 8646528]
45. Shen JC, Rideout WM 3rd, Jones PA. The rate of hydrolytic deamination of 5-methylcytosine in double-stranded DNA. *Nucleic Acids Res*. 1994; 22:972–976. [PubMed: 8152929]
46. Wang RY, Kuo KC, Gehrke CW, Huang LH, Ehrlich M. Heat- and alkali-induced deamination of 5-methylcytosine and cytosine residues in DNA. *Biochimica et biophysica acta*. 1982; 697:371–377. [PubMed: 7104364]

47. Brooks SC, Adhikary S, Robinson EH, Eichman BF. Recent advances in the structural mechanisms of DNA glycosylases. *Biochim Biophys Acta*. 2013; 1834:247–271. [PubMed: 23076011]
48. Schorderet DF, Gartler SM. Analysis of CpG suppression in methylated and nonmethylated species. *Proc Natl Acad Sci USA*. 1992; 89:957–961. [PubMed: 1736311]
49. Hacia JG, Fan JB, Ryder O, Jin L, Edgemon K, Ghandour G, Mayer RA, Sun B, Hsie L, Robbins CM, Brody LC, Wang D, Lander ES, Lipshutz R, Fodor SP, Collins FS. Determination of ancestral alleles for human single-nucleotide polymorphisms using high-density oligonucleotide arrays. *Nat Genet*. 1999; 22:164–167. [PubMed: 10369258]
50. Halushka MK, Fan JB, Bentley K, Hsie L, Shen N, Weder A, Cooper R, Lipshutz R, Chakravarti A. Patterns of single-nucleotide polymorphisms in candidate genes for blood-pressure homeostasis. *Nat Genet*. 1999; 22:239–247. [PubMed: 10391210]
51. Jones PA, Rideout WM 3rd, Shen JC, Spruck CH, Tsai YC. Methylation, mutation and cancer. *BioEssays : news and reviews in molecular, cellular and developmental biology*. 1992; 14:33–36.
52. Schmutte C, Jones PA. Involvement of DNA methylation in human carcinogenesis. *Biol Chem*. 1998; 379:377–388. [PubMed: 9628328]
53. Wiebauer K, Jiricny J. In vitro correction of G.T mispairs to G.C pairs in nuclear extracts from human cells. *Nature*. 1989; 339:234–236. [PubMed: 2716851]
54. Neddermann P, Jiricny J. The purification of a mismatch-specific thymine-DNA glycosylase from HeLa cells. *J Biol Chem*. 1993; 268:21218–21224. [PubMed: 8407958]
55. Hendrich B, Hardeland U, Ng HH, Jiricny J, Bird A. The thymine glycosylase MBD4 can bind to the product of deamination at methylated CpG sites. *Nature*. 1999; 401:301–304. [PubMed: 10499592]
56. Petronzelli F, Riccio A, Markham GD, Seeholzer SH, Stoerker J, Genuardi M, Yeung AT, Matsumoto Y, Bellacosa A. Biphasic kinetics of the human DNA repair protein MED1 (MBD4), a mismatch-specific DNA N-glycosylase. *J Biol Chem*. 2000; 275:32422–32429. [PubMed: 10930409]
57. Morgan MT, Bennett MT, Drohat AC. Excision of 5-halogenated uracils by human thymine DNA glycosylase: Robust activity for DNA contexts other than CpG. *J Biol Chem*. 2007; 282:27578–27586. [PubMed: 17602166]
58. Boiteux S, Guillet M. Abasic sites in DNA: repair and biological consequences in *Saccharomyces cerevisiae*. *DNA Repair*. 2004; 3:1–12. [PubMed: 14697754]
59. Kavli B, Slupphaug G, Mol CD, Arvai AS, Peterson SB, Tainer JA, Krokan HE. Excision of cytosine and thymine from DNA by mutants of human uracil-DNA glycosylase. *EMBO J*. 1996; 15:3442–3447. [PubMed: 8670846]
60. Maiti A, Noon MS, Mackerell AD Jr, Pozharski E, Drohat AC. Lesion processing by a repair enzyme is severely curtailed by residues needed to prevent aberrant activity on undamaged DNA. *Proc Natl Acad Sci U S A*. 2012; 109:8091–8096. [PubMed: 22573813]
61. Waters TR, Swann PF. Thymine-DNA glycosylase and G to A transition mutations at CpG sites. *Mutat Res*. 2000; 462:137–147. [PubMed: 10767625]
62. Millar CB, Guy J, Sansom OJ, Selfridge J, MacDougall E, Hendrich B, Keightley PD, Bishop SM, Clarke AR, Bird A. Enhanced CpG mutability and tumorigenesis in MBD4-deficient mice. *Science*. 2002; 297:403–405. [PubMed: 12130785]
63. Wu P, Qiu C, Sohail A, Zhang X, Bhagwat AS, Cheng X. Mismatch repair in methylated DNA. Structure and activity of the mismatch-specific thymine glycosylase domain of methyl-CpG-binding protein MBD4. *J Biol Chem*. 2003; 278:5285–5291. [PubMed: 12456671]
64. Bellacosa A, Cicchillitti L, Schepis F, Riccio A, Yeung AT, Matsumoto Y, Golem EA, Genuardi M, Neri G. MED1, a novel human methyl-CpG-binding endonuclease, interacts with DNA mismatch repair protein MLH1. *Proc Natl Acad Sci U S A*. 1999; 96:3969–3974. [PubMed: 10097147]
65. Otani J, Arita K, Kato T, Kinoshita M, Kimura H, Suetake I, Tajima S, Ariyoshi M, Shirakawa M. Structural basis of the versatile DNA recognition ability of the methyl-CpG binding domain of methyl-CpG binding domain protein 4. *J Biol Chem*. 2013; 288:6351–6362. [PubMed: 23316048]
66. Wong E, Yang K, Kuraguchi M, Werling U, Avdievich E, Fan K, Fazzari M, Jin B, Brown AM, Lipkin M, Edelmann W. Mbd4 inactivation increases Cright-arrowT transition mutations and

- promotes gastrointestinal tumor formation. *Proc Natl Acad Sci U S A*. 2002; 99:14937–14942. [PubMed: 12417741]
67. Bader S, Walker M, Hendrich B, Bird A, Bird C, Hooper M, Wyllie A. Somatic frameshift mutations in the MBD4 gene of sporadic colon cancers with mismatch repair deficiency. *Oncogene*. 1999; 18:8044–8047. [PubMed: 10637515]
68. Bader SA, Walker M, Harrison DJ. A human cancer-associated truncation of MBD4 causes dominant negative impairment of DNA repair in colon cancer cells. *British journal of cancer*. 2007; 96:660–666. [PubMed: 17285135]
69. Riccio A, Aaltonen LA, Godwin AK, Loukola A, Percesepe A, Salovaara R, Masciullo V, Genuardi M, Paravatou-Petsotas M, Bassi DE, Ruggeri BA, Klein-Szanto AJ, Testa JR, Neri G, Bellacosa A. The DNA repair gene MBD4 (MED1) is mutated in human carcinomas with microsatellite instability. *Nat Genet*. 1999; 23:266–268. [PubMed: 10545939]
70. Zhang W, Liu Z, Crombet L, Amaya MF, Liu Y, Zhang X, Kuang W, Ma P, Niu L, Qi C. Crystal structure of the mismatch-specific thymine glycosylase domain of human methyl-CpG-binding protein MBD4. *Biochemical and biophysical research communications*. 2011; 412:425–428. [PubMed: 21820404]
71. Guan Y, Manuel RC, Arvai AS, Parikh SS, Mol CD, Miller JH, Lloyd S, Tainer JA. MutY catalytic core, mutant and bound adenine structures define specificity for DNA repair enzyme superfamily. *Nat Struct Biol*. 1998; 5:1058–1064. [PubMed: 9846876]
72. Lee S, Verdine GL. Atomic substitution reveals the structural basis for substrate adenine recognition and removal by adenine DNA glycosylase. *Proc Natl Acad Sci U S A*. 2009; 106:18497–18502. [PubMed: 19841264]
73. Mol CD, Arvai AS, Begley TJ, Cunningham RP, Tainer JA. Structure and activity of a thermostable thymine-DNA glycosylase: evidence for base twisting to remove mismatched normal DNA bases. *J Mol Biol*. 2002; 315:373–384. [PubMed: 11786018]
74. Manvilla BA, Maiti A, Begley MC, Toth EA, Drohat AC. Crystal Structure of Human Methyl-Binding Domain IV Glycosylase Bound to Abasic DNA. *J Mol Biol*. 2012; 420:164–175. [PubMed: 22560993]
75. Morera S, Grin I, Vigouroux A, Couve S, Henriot V, Saparbaev M, Ishchenko AA. Biochemical and structural characterization of the glycosylase domain of MBD4 bound to thymine and 5-hydroxymethyluracil-containing DNA. *Nucleic Acids Res*. 2012; 40:9917–9926. [PubMed: 22848106]
76. Hashimoto H, Zhang X, Cheng X. Excision of thymine and 5-hydroxymethyluracil by the MBD4 DNA glycosylase domain: structural basis and implications for active DNA demethylation. *Nucleic Acids Res*. 2012
77. Turner DP, Cortellino S, Schupp JE, Caretti E, Loh T, Kinsella TJ, Bellacosa A. The DNA N-glycosylase MED1 exhibits preference for halogenated pyrimidines and is involved in the cytotoxicity of 5-iododeoxyuridine. *Cancer Res*. 2006; 66:7686–7693. [PubMed: 16885370]
78. Moe JG, Russu IM. Kinetics and energetics of base-pair opening in 5'-d (CGCGAATTCGCG)-3' and a substituted dodecamer containing G.T mismatches. *Biochemistry*. 1992; 31:8421–8428. [PubMed: 1327102]
79. Krosky DJ, Schwarz FP, Stivers JT. Linear free energy correlations for enzymatic base flipping: how do damaged base pairs facilitate specific recognition? *Biochemistry*. 2004; 43:4188–4195. [PubMed: 15065862]
80. Darwanto A, Theruvathu JA, Sowers JL, Rogstad DK, Pascal T, Goddard W 3rd, Sowers LC. Mechanisms of base selection by human single-stranded selective monofunctional uracil-DNA glycosylase. *J Biol Chem*. 2009; 284:15835–15846. [PubMed: 19324873]
81. Sibghat U, Gallinari P, Xu YZ, Goodman MF, Bloom LB, Jiricny J, Day RS 3rd. Base analog and neighboring base effects on substrate specificity of recombinant human G:T mismatch-specific thymine DNA-glycosylase. *Biochemistry*. 1996; 35:12926–12932. [PubMed: 8841138]
82. Abu M, Waters TR. The main role of human thymine-DNA glycosylase is removal of thymine produced by deamination of 5-methylcytosine and not removal of ethenocytosine. *J Biol Chem*. 2003; 278:8739–8744. [PubMed: 12493755]

83. Walavalkar NM, Cramer JM, Buchwald WA, Scarsdale JN, Williams DC Jr. Solution structure and intramolecular exchange of methyl-cytosine binding domain protein 4 (MBD4) on DNA suggests a mechanism to scan for mCpG/TpG mismatches. *Nucleic Acids Res.* 2014; 42:11218–11232. [PubMed: 25183517]
84. Bellacosa A. Role of MED1 (MBD4) Gene in DNA repair and human cancer. *J Cell Physiol.* 2001; 187:137–144. [PubMed: 11267993]
85. Tini M, Benecke A, Um SJ, Torchia J, Evans RM, Chambon P. Association of CBP/p300 acetylase and thymine DNA glycosylase links DNA repair and transcription. *Mol Cell.* 2002; 9:265–277. [PubMed: 11864601]
86. Mohan RD, Litchfield DW, Torchia J, Tini M. Opposing regulatory roles of phosphorylation and acetylation in DNA mispair processing by thymine DNA glycosylase. *Nucleic Acids Res.* 2009
87. Hardeland U, Steinacher R, Jiricny J, Schar P. Modification of the human thymine-DNA glycosylase by ubiquitin-like proteins facilitates enzymatic turnover. *EMBO J.* 2002; 21:1456–1464. [PubMed: 11889051]
88. Guan X, Madabushi A, Chang DY, Fitzgerald M, Shi G, Drohat AC, Lu AL. The Human Checkpoint Sensor Rad9-Rad1-Hus1 Interacts with and Stimulates DNA Repair Enzyme TDG Glycosylase. *Nucleic Acids Res.* 2007; 35:6207–6218. [PubMed: 17855402]
89. Madabushi A, Hwang BJ, Jin J, Lu AL. Histone deacetylase SIRT1 modulates and deacetylates DNA base excision repair enzyme thymine DNA glycosylase. *Biochem J.* 2013; 456:89–98. [PubMed: 23952905]
90. Smet-Nocca C, Wieruszkeski JM, Chaar V, Leroy A, Benecke A. The Thymine-DNA Glycosylase Regulatory Domain: Residual Structure and DNA Binding. *Biochemistry.* 2008
91. Wiebauer K, Jiricny J. 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.* 1990; 87:5842–5845. [PubMed: 2116008]
92. Neddermann P, Gallinari P, Lettieri T, Schmid D, Truong O, Hsuan JJ, Wiebauer K, Jiricny J. Cloning and expression of human G/T mismatch-specific thymine-DNA glycosylase. *J Biol Chem.* 1996; 271:12767–12774. [PubMed: 8662714]
93. Cortellino S, Xu J, Sannai M, Moore R, Caretti E, Cigliano A, Le Coz M, Devarajan K, Wessels A, Soprano D, Abramowitz LK, Bartolomei MS, Rambow F, Bassi MR, Bruno T, Fanciulli M, Renner C, Klein-Szanto AJ, Matsumoto Y, Kobi D, Davidson I, Alberti C, Larue L, Bellacosa A. Thymine DNA glycosylase is essential for active DNA demethylation by linked deamination-base excision repair. *Cell.* 2011; 146:67–79. [PubMed: 21722948]
94. Vasovcak P, Krepelova A, Menigatti M, Puchmajerova A, Skapa P, Augustinakova A, Amann G, Wernstedt A, Jiricny J, Marra G, Wimmer K. Unique mutational profile associated with a loss of TDG expression in the rectal cancer of a patient with a constitutional PMS2 deficiency. *DNA Repair (Amst).* 2012; 11:616–623. [PubMed: 22608206]
95. Maiti A, Morgan MT, Pozharski E, Drohat AC. Crystal Structure of Human Thymine DNA Glycosylase Bound to DNA Elucidates Sequence-Specific Mismatch Recognition. *Proc Natl Acad Sci USA.* 2008; 105:8890–8895. [PubMed: 18587051]
96. Zhang L, Lu X, Lu J, Liang H, Dai Q, Xu GL, Luo C, Jiang H, He C. Thymine DNA glycosylase specifically recognizes 5-carboxylcytosine-modified DNA. *Nat Chem Biol.* 2012; 8:328–330. [PubMed: 22327402]
97. Hardeland U, Kunz C, Focke F, Szadkowski M, Schar P. Cell cycle regulation as a mechanism for functional separation of the apparently redundant uracil DNA glycosylases TDG and UNG2. *Nucleic Acids Res.* 2007; 35:3859–3867. [PubMed: 17526518]
98. Waters TR, Swann PF. Kinetics of the action of thymine DNA glycosylase. *J Biol Chem.* 1998; 273:20007–20014. [PubMed: 9685338]
99. 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]
100. Steinacher R, Schar P. Functionality of human thymine DNA glycosylase requires SUMO-regulated changes in protein conformation. *Current biology : CB.* 2005; 15:616–623. [PubMed: 15823533]

101. 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. *The Journal of biological chemistry*. 2008; 283:32680–32690. [PubMed: 18805789]
102. Schärer OD, Nash HM, Jiricny J, Laval J, Verdine GL. Specific binding of a designed pyrrolidine abasic site analog to multiple DNA glycosylases. *J Biol Chem*. 1998; 273:8592–8597. [PubMed: 9535832]
103. Baba D, Maita N, Jee JG, Uchimura Y, Saitoh H, Sugasawa K, Hanaoka F, Tochio H, Hiroaki H, Shirakawa M. Crystal structure of thymine DNA glycosylase conjugated to SUMO-1. *Nature*. 2005; 435:979–982. [PubMed: 15959518]
104. Baba D, Maita N, Jee JG, Uchimura Y, Saitoh H, Sugasawa K, Hanaoka F, Tochio H, Hiroaki H, Shirakawa M. Crystal structure of SUMO-3-modified thymine-DNA glycosylase. *J Mol Biol*. 2006; 359:137–147. [PubMed: 16626738]
105. Baba D, Maita N, Jee JG, Uchimura Y, Saitoh H, Sugasawa K, Hanaoka F, Tochio H, Hiroaki H, Shirakawa M. Crystal structure of thymine DNA glycosylase conjugated to SUMO-1. *Nature*. 2005; 435:979–982. [PubMed: 15959518]
106. Johnson ES. Protein modification by SUMO. *Annual review of biochemistry*. 2004; 73:355–382.
107. Geiss-Friedlander R, Melchior F. Concepts in sumoylation: a decade on. 2007; 8:947–956.
108. Pilla E, Moller U, Sauer G, Mattioli F, Melchior F, Geiss-Friedlander R. A novel SUMO1-specific interacting motif in dipeptidyl peptidase 9 (DPP9) that is important for enzymatic regulation. *The Journal of biological chemistry*. 2012; 287:44320–44329. [PubMed: 23152501]
109. Coey CT, Fitzgerald ME, Maiti A, Reiter KH, Guzzo CM, Matunis MJ, Drohat AC. E2-mediated small ubiquitin-like modifier (SUMO) modification of thymine DNA glycosylase is efficient but not selective for the enzyme-product complex. *J Biol Chem*. 2014; 289:15810–15819. [PubMed: 24753249]
110. Kim EJ, Um SJ. Thymine-DNA glycosylase interacts with and functions as a coactivator of p53 family proteins. *Biochemical and biophysical research communications*. 2008; 377:838–842. [PubMed: 18951877]
111. Hock A, Vousden KH. Regulation of the p53 pathway by ubiquitin and related proteins. *The international journal of biochemistry & cell biology*. 2010; 42:1618–1621. [PubMed: 20601087]
112. Stindt MH, Carter S, Vigneron AM, Ryan KM, Vousden KH. MDM2 promotes SUMO-2/3 modification of p53 to modulate transcriptional activity. *Cell cycle (Georgetown, Tex)*. 2011; 10:3176–3188.
113. Mohan RD, Rao A, Gagliardi J, Tini M. SUMO-1-dependent allosteric regulation of thymine DNA glycosylase alters subnuclear localization and CBP/p300 recruitment. *Molecular and cellular biology*. 2007; 27:229–243. [PubMed: 17060459]
114. Takahashi H, Hatakeyama S, Saitoh H, Nakayama KI. Noncovalent SUMO-1 binding activity of thymine DNA glycosylase (TDG) is required for its SUMO-1 modification and colocalization with the promyelocytic leukemia protein. *The Journal of biological chemistry*. 2005; 280:5611–5621. [PubMed: 15569683]
115. Slenn TJ, Morris B, Havens CG, Freeman RM Jr, Takahashi TS, Walter JC. Thymine DNA Glycosylase is a CRL4Cdt2 Substrate. *J Biol Chem*. 2014
116. Shibata E, Dar A, Dutta A. CRL4Cdt2 E3 Ubiquitin Ligase and PCNA Cooperate to Degrade Thymine DNA Glycosylase in S-phase. *J Biol Chem*. 2014
117. Sjolund A, Nemec AA, Paquet N, Prakash A, Sung P, Doublet S, Sweasy JB. A germline polymorphism of thymine DNA glycosylase induces genomic instability and cellular transformation. *PLoS genetics*. 2014; 10:e1004753. [PubMed: 25375110]
118. Kunz C, Focke F, Saito Y, Schuermann D, Lettieri T, Selfridge J, Schar P. Base excision by thymine DNA glycosylase mediates DNA-directed cytotoxicity of 5-fluorouracil. *PLoS Biol*. 2009; 7:e91. [PubMed: 19402749]
119. 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; 283:32680–32690. [PubMed: 18805789]

120. Maiti A, Morgan MT, Drohat AC. Role of two strictly conserved residues in nucleotide flipping and N-glycosylic bond cleavage by human thymine DNA glycosylase. *J Biol Chem.* 2009; 284:36680–36688. [PubMed: 19880517]
121. Morgan MT, Maiti A, Fitzgerald ME, Drohat AC. Stoichiometry and affinity for thymine DNA glycosylase binding to specific and nonspecific DNA. *Nucleic Acids Res.* 2011; 39:2319–2329. [PubMed: 21097883]
122. Maiti A, Drohat AC. Dependence of substrate binding and catalysis on pH, ionic strength, and temperature for thymine DNA glycosylase: Insights into recognition and processing of G.T mispairs. *DNA Repair.* 2011; 10:545–553. [PubMed: 21474392]
123. Rideout WM 3rd, Coetzee GA, Olumi AF, Jones PA. 5-Methylcytosine as an endogenous mutagen in the human LDL receptor and p53 genes. *Science.* 1990; 249:1288–1290. [PubMed: 1697983]
124. Cooper DN, Yousoufian H. The CpG dinucleotide and human genetic disease. *Hum Genet.* 1988; 78:151–155. [PubMed: 3338800]
125. Sjoblom T, Jones S, Wood LD, Parsons DW, Lin J, Barber TD, Mandelker D, Leary RJ, Ptak J, Silliman N, Szabo S, Buckhaults P, Farrell C, Meeh P, Markowitz SD, Willis J, Dawson D, Willson JK, Gazdar AF, Hartigan J, Wu L, Liu C, Parmigiani G, Park BH, Bachman KE, Papadopoulos N, Vogelstein B, Kinzler KW, Velculescu VE. The consensus coding sequences of human breast and colorectal cancers. *Science.* 2006; 314:268–274. [PubMed: 16959974]
126. Pfeifer GP, Besaratinia A. Mutational spectra of human cancer. *Hum Genet.* 2009; 125:493–506. [PubMed: 19308457]
127. Zhu B, Zheng Y, Hess D, Angliker H, Schwarz S, Siegmann M, Thiry S, Jost JP. 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; 97:5135–5139. [PubMed: 10779566]
128. Zhu B, Zheng Y, Angliker H, Schwarz S, Thiry S, Siegmann M, Jost JP. 5-Methylcytosine DNA glycosylase activity is also present in the human MBD4 (G/T mismatch glycosylase) and in a related avian sequence. *Nucleic Acids Res.* 2000; 28:4157–4165. [PubMed: 11058112]
129. Bennett MT, Rodgers MT, Hebert AS, Ruslander LE, Eisele L, Drohat AC. Specificity of Human Thymine DNA Glycosylase Depends on N-Glycosidic Bond Stability. *J Am Chem Soc.* 2006; 128:12510–12519. [PubMed: 16984202]
130. Cortazar D, Kunz C, Selfridge J, Lettieri T, Saito Y, Macdougall E, Wirz A, Schuermann D, Jacobs AL, Siegrist F, Steinacher R, Jiricny J, Bird A, Schar P. Embryonic lethal phenotype reveals a function of TDG in maintaining epigenetic stability. *Nature.* 2011; 470:419–423. [PubMed: 21278727]
131. Feng S, Jacobsen SE, Reik W. Epigenetic reprogramming in plant and animal development. *Science.* 2010; 330:622–627. [PubMed: 21030646]
132. Rai K, Huggins JJ, James SR, Karpf AR, Jones DA, Cairns BR. DNA demethylation in zebrafish involves the coupling of a deaminase, a glycosylase, and gadd45. *Cell.* 2008; 135:1201–1212. [PubMed: 19109892]
133. Shimoda N, Hirose K, Kaneto R, Izawa T, Yokoi H, Hashimoto N, Kikuchi Y. No Evidence for AID/MBD4-Coupled DNA Demethylation in Zebrafish Embryos. *PLoS One.* 2014; 9:e114816. [PubMed: 25536520]
134. Kim MS, Kondo T, Takada I, Youn MY, Yamamoto Y, Takahashi S, Matsumoto T, Fujiyama S, Shiode Y, Yamaoka I, Kitagawa H, Takeyama K, Shibuya H, Ohtake F, Kato S. DNA demethylation in hormone-induced transcriptional derepression. *Nature.* 2009; 461:1007–1012. [PubMed: 19829383]
135. Kim MS, Kondo T, Takada I, Youn MY, Yamamoto Y, Takahashi S, Matsumoto T, Fujiyama S, Shiode Y, Yamaoka I, Kitagawa H, Takeyama K, Shibuya H, Ohtake F, Kato S. Retraction: DNA demethylation in hormone-induced transcriptional derepression. *Nature.* 2012; 486:280. [PubMed: 22699624]
136. Sabag O, Zamir A, Keshet I, Hecht M, Ludwig G, Tabib A, Moss J, Cedar H. Establishment of methylation patterns in ES cells. *Nature structural & molecular biology.* 2014; 21:110–112.

137. Nabel CS, Jia H, Ye Y, Shen L, Goldschmidt HL, Stivers JT, Zhang Y, Kohli RM. AID/APOBEC deaminases disfavor modified cytosines implicated in DNA demethylation. *Nat Chem Biol.* 2012; 8:751–758. [PubMed: 22772155]
138. Rangam G, Schmitz KM, Cobb AJ, Petersen-Mahrt SK. AID enzymatic activity is inversely proportional to the size of cytosine C5 orbital cloud. *PLoS One.* 2012; 7:e43279. [PubMed: 22916236]
139. Hashimoto H, Zhang X, Cheng X. Activity and crystal structure of human thymine DNA glycosylase mutant N140A with 5-carboxylcytosine DNA at low pH. *DNA Repair (Amst).* 2013
140. Maiti A, Michelson AZ, Armwood CJ, Lee JK, Drohat AC. Divergent mechanisms for enzymatic excision of 5-formylcytosine and 5-carboxylcytosine from DNA. *J Am Chem Soc.* 2013; 135:15813–15822. [PubMed: 24063363]
141. Hardeland U, Bentele M, Jiricny J, Schar P. The versatile thymine DNA-glycosylase: a comparative characterization of the human, *Drosophila* and fission yeast orthologs. *Nucleic Acids Res.* 2003; 31:2261–2271. [PubMed: 12711670]
142. Illingworth RS, Bird AP. CpG islands--‘a rough guide’. *FEBS Lett.* 2009; 583:1713–1720. [PubMed: 19376112]
143. Li YQ, Zhou PZ, Zheng XD, Walsh CP, Xu GL. Association of Dnmt3a and thymine DNA glycosylase links DNA methylation with base-excision repair. *Nucleic Acids Res.* 2007; 35:390–400. [PubMed: 17175537]
144. Boland MJ, Christman JK. Characterization of Dnmt3b:thymine-DNA glycosylase interaction and stimulation of thymine glycosylase-mediated repair by DNA methyltransferase(s) and RNA. *J Mol Biol.* 2008; 379:492–504. [PubMed: 18452947]
145. Chen D, Lucey MJ, Phoenix F, Lopez-Garcia J, Hart SM, Losson R, Buluwela L, Coombes RC, Chambon P, Schar P, Ali S. T:G Mismatch-specific Thymine-DNA Glycosylase Potentiates Transcription of Estrogen-regulated Genes through Direct Interaction with Estrogen Receptor {alpha}. *J Biol Chem.* 2003; 278:38586–38592. [PubMed: 12874288]
146. Missero C, Pirro MT, Simeone S, Pischetola M, Di Lauro R. The DNA glycosylase T:G mismatch-specific thymine DNA glycosylase represses thyroid transcription factor-1-activated transcription. *J Biol Chem.* 2001; 276:33569–33575. [PubMed: 11438542]
147. Um S, Harbers M, Benecke A, Pierrat B, Losson R, Chambon P. Retinoic acid receptors interact physically and functionally with the T:G mismatch-specific thymine-DNA glycosylase. *J Biol Chem.* 1998; 273:20728–20736. [PubMed: 9694815]
148. Song CX, Szulwach KE, Dai Q, Fu Y, Mao SQ, Lin L, Street C, Li Y, Poidevin M, Wu H, Gao J, Liu P, Li L, Xu GL, Jin P, He C. Genome-wide profiling of 5-formylcytosine reveals its roles in epigenetic priming. *Cell.* 2013; 153:678–691. [PubMed: 23602153]
149. Spruijt CG, Gnerlich F, Smits AH, Pfaffeneder T, Jansen PW, Bauer C, Munzel M, Wagner M, Muller M, Khan F, Eberl HC, Mensinga A, Brinkman AB, Lephikov K, Muller U, Walter J, Boelens R, van Ingen H, Leonhardt H, Carell T, Vermeulen M. Dynamic Readers for 5-(Hydroxy)Methylcytosine and Its Oxidized Derivatives. *Cell.* 2013; 152:1146–1159. [PubMed: 23434322]

Fig. 1. Pathway for active DNA demethylation mediated by TET enzymes followed by TDG-initiated BER. TET enzymes (noted in red) catalyze stepwise oxidation of 5mC to 5hmC, 5fC, and 5caC. TDG excises 5fC and 5caC and subsequent BER steps yield cytosine.

Fig. 2.

Structures of MBD4. (A) Schematic primary structure of MBD4. The MBD and the glycosylase domain are structured, while other regions appear to be disordered. (B) Crystal structure of the MBD bound to a G·T mismatch in a CpG context, i.e., 5mCG/TG (PDBID: 3VXV). The DNA is shown as gray lines with surface representation, the G·T mismatch is shown in cyan and the flanking C·G pair (CpG context) is in yellow (with N and O atoms blue and red). The same coloring scheme is used in panels C, D, and E. Note that the structures in panels B through E are oriented in the same way with respect to the position of the mismatched guanine (cyan, with an asterisk). (C) Structure of the MBD4 glycosylase domain bound to DNA, with thymine of a G·T mismatch flipped into the active site (PDBID: 4E9G). (D) Close up view of the MBD (as in panel B) shows that it recognizes an 5mCG/GT dinucleotide using conserved Arg and other residues together with water molecules. (E) Close-up view of contacts the glycosylase forms with the flipped thymine, via Tyr and Gln side chains and backbone contacts. Also shown are contacts with mismatched G, and the Arg side chain that fills the void created by thymine flipping.

Fig. 3.

Structures and molecular dynamics (MD) studies of human TDG. (A) Schematic primary structure of TDG showing the catalytic (glycosylase) domain and the two flanking regions that are disordered yet functionally important. Shown are sites of post-translational modifications, the SUMO-interacting motif (SIM), and PIP degron that mediates interaction with PCNA and subsequent ubiquitination and degradation. (B) Structure of TDG (catalytic domain) bound to DNA with U of a G·U mismatch flipped into active site (2.97 Å resolution, PDBID: 3UFJ). The DNA is shown as gray lines with surface representation; the G·U mismatch is shown in cyan and the flanking C·G pair (i.e., CpG context) is yellow. Contacts with the flipped U and mismatched G are shown. (C) Structure of SUMO-1 modified TDG (residues 117–331, PDBID: 1WYW), with DNA modeled in (by aligning DNA-bound TDG; PDBID 2RBA). The SUMO-1 domain is covalently tethered to K330 of TDG and it forms many non-covalent interactions with the SIM of TDG (left surface as shown). SUMO-binding stabilizes an otherwise disordered helix of TDG (magenta), which likely weakens TDG binding to AP-DNA. (D) To illustrate a potential steric clash involving the methyl groups of a flipped thymine and Ala145, uracil from the structure in panel B was replaced with thymine (interactions shown are from the structure in panel B). (E) MD simulations of TDG bound to a G·T mismatch suggest that the enzyme can transiently flip thymine into a position that lies just shy of the fully-flipped conformation which is likely needed for base excision. The last two panels are adapted from Ref [60].