Published in final edited form as:

Chem Rev. 2016 October 26; 116(20): 12711–12729. doi:10.1021/acs.chemrev.6b00191.

Role of Base Excision “Repair” Enzymes in Erasing Epigenetic Marks from DNA

Alexander C. Drohat* and Christopher T. Coey

Department of Biochemistry and Molecular Biology, University of Maryland School of Medicine, Baltimore, MD 21201, USA

Abstract

Base excision repair (BER) is one of several DNA repair pathways found in all three domains of life. BER counters the mutagenic and cytotoxic effects of damage that occurs continuously to the nitrogenous bases in DNA, and its critical role in maintaining genomic integrity is well established. However, BER also performs essential functions in processes other than DNA repair, where it acts on naturally modified bases in DNA. A prominent example is the central role of BER in mediating active DNA demethylation, a multi-step process that erases the epigenetic mark, 5-methylcytosine (5mC), or derivatives thereof, converting them back to cytosine. Here, we review recent advances in the understanding of how BER mediates this critical component of epigenetic regulation, in plants and animals.

Graphical abstract

1. INTRODUCTION

Base excision repair (BER) is one of the central DNA repair pathways present in all three domains of life.¹ BER counters the mutagenic and cytotoxic effects of damage that occurs to the nucleobases of DNA, and its critical role in maintaining genomic integrity is well established.^{2–4} In this review, we focus on the emerging and critical role of BER in epigenetic regulation, where it acts on nucleobases that have been purposefully modified rather than on lesions generated by damaging agents. The main emphasis is on the role of BER in active DNA demethylation, a multistep pathway that erases the epigenetic mark, 5-methylcytosine (5mC), or derivatives of 5mC, and converts them to back cytosine.^{5–8} The

*Corresponding Author: Phone: 410-706-8118. adrohat@som.umaryland.edu.

The authors declare no competing financial interests

enzymatic conversion of cytosine to 5-methylcytosine (5mC) generates an epigenetic signal in DNA that is critical for numerous cellular processes.^{9,10} We review recent advances in our understanding of how BER functions to erase this epigenetic mark, and highlight the similarities and substantial differences in this process in plants and animals. To set the stage, we begin with a brief overview of BER and highlight its important functions in DNA repair.

1.1. Overview of Base Excision Repair (BER)

The integrity of genetic information is under constant threat due to the susceptibility of DNA to a broad range of chemical modifications imparted by endogenous and exogenous agents.¹¹ The threat is countered by several DNA repair systems found in all three domains of life.^{1,12} One of these critical protective systems is base excision repair (BER), a pathway that handles damage occurring to the nitrogenous bases of DNA, in addition to other types of DNA damage.^{2–4,13}

The nucleobases of DNA are subject to three broad categories of chemical modification, including alkylation by endogenous and exogenous electrophiles, hydrolytic deamination of exocyclic amino groups, and oxidation by a variety of reactive oxygen species.^{11,14} Prominent examples include methylation of adenine to 3-methyladenine, deamination of cytosine to uracil, and oxidation of guanine to 8-oxoguanine. These examples represent just a few of the many dozens of base lesions that arise in DNA and are corrected by the BER pathway. If not repaired, these lesions can lead to mutations (upon replication), hinder critical DNA transactions such as replication or transcription, and trigger apoptosis. Unrepaired base damage is implicated in premature aging and diseases including cancer.¹⁵

While the detailed chemical reactions of BER and the enzymes that mediate them vary from prokaryotes to higher eukaryotes, the basic steps of the pathway are largely conserved (Figure 1). For a more general discussion of the BER pathway, we suggest other recent reviews.^{2–4} The basic steps of BER include (i) excision of the damaged or modified base by a DNA glycosylase, giving an abasic or apurinic/apyrimidinic (AP) site, (ii) cleavage of the phosphodiester bond at one or both sides of the abasic sugar, (iii) cleanup of the resulting termini, if needed, to generate the 3'-OH and 5'-phosphate needed for DNA synthesis and ligation, (iv) DNA synthesis to replace the excised nucleotide(s), and (v) ligation of the residual nick. Given the critical role of DNA glycosylases, which find damaged or modified bases and excise them to initiate the BER pathway, we discuss these enzymes in greater detail below.

In addition to damaged bases, BER also handles AP sites generated by spontaneous hydrolysis of the *N*-glycosyl bond, an event that is far more frequent for purines than for pyrimidines. BER also processes single strand breaks (SSBs), together with other enzymes that remove adducts at the termini of SSBs, as needed to generate the 3'-OH and 5'-phosphate needed for DNA synthesis and ligation in the final stage of BER.^{16,17}

1.2. Introduction to DNA Glycosylases

The BER pathway is initiated by DNA glycosylases, enzymes that employ a nucleotide-flipping mechanism to recognize damaged or modified bases in DNA and remove them by cleaving the *N*-glycosyl bond that links the base to the 2-deoxyribose sugar (Figure 2).^{18–20}

Base-flipping facilitates the search for lesions that in many cases do not dramatically distort the overall structure of DNA and provides access to the scissile C-N bond. While the base-flipping mechanism is used by all known mammalian glycosylases, some bacterial glycosylases remove bulky base lesions without flipping them out of the DNA duplex.²¹ Many different DNA glycosylases are needed to handle the broad range of base lesions that arise in a given organism; mammals have 11 distinct glycosylases.³ While some glycosylases act predominantly on a single type of lesion, such as uracil DNA glycosylase (UNG), others remove many different modified bases, in keeping with findings that the number of distinct lesions is far greater than the number of glycosylases in a given organism.

While most glycosylases remove bases that are foreign to DNA, a select few act on mismatches involving two canonical bases. Three glycosylases act on G/T mismatches, including TDG (thymine DNA glycosylase),^{22,23} MBD4 (methyl binding domain IV),^{24–26} and MIG (mismatch DNA glycosylase, Mig-Mth).^{27,28} Notably, these G/T mismatch glycosylases excise only thymine from G/T mismatches, as needed to protect against lesions arising by deamination of 5mC to thymine. By contrast, mismatches that arise from errors in DNA replication, including G/T mismatches, are handled by mismatch repair (MMR), a conserved pathway that acts just after the replication machinery.^{29,30} One of the mismatch glycosylases, TDG, also removes oxidized forms of 5mC, as discussed in detail below.

There are two general categories of DNA glycosylases, with respect to the reactions that they catalyze, monofunctional and bifunctional (Figure 3).^{31,32} Both types of glycosylases are featured in the epigenetic functions of BER discussed below. Monofunctional glycosylases catalyze hydrolysis of the *N*-glycosyl bond, yielding an AP site in DNA and the liberated nucleobase. By contrast, bifunctional glycosylases typically use an amine nucleophile, such as a Lys side chain, to cleave the *N*-glycosyl bond, generating a Schiff base (imine) intermediate.³³ This intermediate initiates the second activity, cleavage of the phosphodiester backbone on the 3' side of the lesion through a β -elimination reaction. Some bifunctional DNA glycosylases can also cleave a second phosphodiester bond, on the 5' side of the lesion, through δ -elimination. It is important to note that bifunctional glycosylases can also perform lyase activity on the AP sites generated by monofunctional glycosylases, although AP sites are typically handled by an AP endonuclease. In addition, some bifunctional glycosylases can catalyze hydrolytic excision of the *N*-glycosyl bond to yield an AP site.

The various reaction products generated by monofunctional and bifunctional glycosylases can themselves be mutagenic and cytotoxic, and it is therefore vital that they be recognized and processed by follow-on BER enzymes. The products generated by bifunctional glycosylases require processing by other BER enzymes to generate DNA termini (3'-OH, 5'-phosphate) that enable follow-on steps in BER, including DNA synthesis and DNA ligation. These steps are discussed in relevant sections below.

1.4. BER in Epigenetic Regulation

Perhaps the most prominent role for BER outside of DNA repair is in active DNA demethylation, a multi-step enzymatic pathway that converts 5-methylcytosine (5mC), or derivatives of 5mC, back to C (Figure 4).^{5,26} We provide an overview of active DNA

demethylation in this section, to set the stage for a detailed discussion of this important epigenetic process in plants and animals.

The methylation of cytosine bases in DNA generates an epigenetic mark that is typically associated with a repressed chromatin state.³⁴ Methylation occurs most often at cytosine bases within the context of a CG dinucleotide (or CpG site) in animals,¹⁰ but is increasingly recognized to arise in other DNA contexts (non-CpG sites).^{35,36} In plants, cytosine methylation is observed not only at CpG sites but in other sequence contexts, including CHH and CHG (H = A, C, or T).^{8,37,38}

It has been long known that cytosine is converted to 5mC by DNA methyltransferases in plants and animals (Figure 4), but the process for actively erasing this epigenetic mark has only recently emerged. We note that passive DNA demethylation refers to the depletion of 5mC (or derivatives thereof) via successive rounds of DNA replication in the absence of maintenance methylation.⁵ Early models for active DNA demethylation included direct cleavage of the methyl group from 5mC, and it was reported that an enzyme capable of catalyzing this challenging chemical reaction had been identified.³⁹ However, this initial report has not been substantiated, and no enzyme is presently known to possess this catalytic activity.

Research over the past decade or so shows that BER modulates the levels of cytosine methylation, in plants and animals (vertebrates), by mediating the conversion of 5mC, or derivatives of 5mC, back to C (Figure 4). The process begins with glycosylase-mediated excision of 5mC, or a derivative generated via deamination or oxidation of 5mC. Plants have DNA glycosylases that can directly excise 5mC from DNA.^{40–42} By contrast, the glycosylases found in vertebrates cannot remove 5mC but they can readily excise modified forms of 5mC.^{43,44} Thus, in vertebrates, other enzymes modify 5mC to yield derivatives that are excised by a DNA glycosylase, and then follow-on BER finishes the process of active demethylation.

2. BER-MEDIATED DNA DEMETHYLATION IN PLANTS

In flowering plants such as *Arabidopsis thaliana*, BER-mediated DNA demethylation enables allele-specific expression of genes in the endosperm, which are otherwise silenced by imprinting.^{38,41,45} Active DNA demethylation in both the male and female gametophytes is essential for normal seed development. DNA demethylation also appears to activate transcription of precursors for small interfering RNAs (siRNA) which guide *de novo* methylation and can reinforce silencing of transposons and other repetitive sequences.^{46,47} BER also performs genome-wide demethylation in adult plant cells (vegetative tissues).^{40,48–51} In this section, we review BER-mediated DNA demethylation in plants, with an emphasis on the DNA glycosylases that initiate this epigenetic process.

2.1. BER Erases 5-methylcytosine in Plants

In plants, erasure of the 5mC epigenetic mark is handled entirely by the BER pathway (Figure 5). DNA demethylation is initiated by one of four related DNA glycosylases, which can excise 5mC directly from DNA. These four enzymes, sometimes referred to as 5mC

glycosylases, are DEMETER (DME), Repressor of Silencing 1 (ROS1), and DME-like 2 and 3 (DML2, DML3).^{40–42,51–54} These homologous enzymes all belong to the HhH (helix-hairpin-helix) or Endo III superfamily of DNA glycosylases. Moreover, all four of these 5mC glycosylases are bifunctional enzymes, possessing both glycosylase (base excision) and AP lyase activity (Figures 3, 5).

The lyase activity of these enzymes includes a β -elimination reaction, giving a 3'-phospho- α,β -unsaturated aldehyde (3'-PUA), and δ -elimination, which converts 3'-PUA to 3'-phosphate. Other BER enzymes are required to process each of these lyase products, because neither is a substrate for a DNA polymerase, which requires a 3'-OH terminus for DNA synthesis. It is also possible that the DME enzymes may in some cases perform in a monofunctional mode, catalyzing only the base excision reaction to generate an abasic site that would be recognized by an AP endonuclease, as suggested by some studies of OGG1,^{55–60} a bifunctional glycosylase belonging to the same superfamily as the DMEs.

APE1L, one of three AP endonucleases found in *Arabidopsis*, catalyzes hydrolysis of 3'-PUA (created by the lyase reaction) to give the 3'-OH needed for DNA synthesis. Another AP endonuclease, ARP, can also hydrolyze 3'-PUA, though not as efficiently as APE1L.⁶¹ Processing of the 3'-phosphate that is generated by combined β - and δ -elimination is handled by a phosphatase known as zinc finger DNA 3' phosphoesterase (ZDP), an enzyme that generates a 3'-OH terminus and is required for ROS1-initiated DNA demethylation.⁶²

The δ -elimination activity of ROS1 is inefficient relative to the β -elimination activity, raising questions about the biological relevance of δ -elimination, given that the resulting 3'-PUA intermediate is potentially toxic.^{62,63} Accordingly, recent studies indicate that the predominant product of DME-ROS1 enzymes is 3'-PUA generated by β -elimination,⁶¹ which is then handled by APE1L and follow-on BER enzymes (Figure 6).

Another BER protein that mediates DNA demethylation in plants is XRCC1 (x-ray cross-complementing group protein 1), which interacts with ROS1 and ZDP and is reported to stimulate the activity of each, thereby facilitating 5mC excision, 3'-end processing, and DNA ligation.⁶⁴

2.2. Structure of 5mC Glycosylases

Each of the four DME-ROS1 5mC glycosylases in plants contain three conserved domains, including the “A”, the “glycosylase”, and the “B” domain (Figure 6A).^{65–67} The glycosylase domain is homologous to the helix-hairpin-helix (HhH) catalytic motif, which is the hallmark of enzymes in the HhH (or Endo III) superfamily. There are no reported crystal or NMR structures of a DME-ROS1 glycosylase, but models have been generated based on similarity to a related enzyme, Endo III (also known as Nth) (Figure 6B).^{66,67}

Remarkably, results from sequence alignments, homology modeling, and site-directed mutagenesis indicate that the catalytic core of DME-ROS enzymes includes elements from two separate domains, “A” and “glycosylase”, which are joined by a very large interdomain region (IDR1) that is probably disordered (Figure 6).^{66,67} The “glycosylase” domain of DME-ROS1 enzymes lacks the first three α -helices that are conserved in the HhH

superfamily and are essential for activity. However, these missing α -helices are provided by domain “A” of DME enzymes. Notably, the huge size of the interdomain region (IDR1) that joins these domains, which can include hundreds of amino acid residues, is unprecedented among members of the Endo III superfamily, which typically exhibit a small loop of about a dozen residues in this region.⁶⁶

The “glycosylase” domain of the DME-ROS1 enzymes contains the signature “helix-hairpin-helix” (HhH) motif that is present in all HhH enzymes (Figure 6B). It also contains an iron-sulfur or [4Fe-4S] cluster, as do many HhH superfamily members (Endo III, MutY, and Mig). Disruption of the iron-sulfur cluster causes a decrease in glycosylase activity for DME-ROS1 enzymes, but its exact role remains unclear.⁶⁵ Notably, the iron-sulfur cluster found in the related HhH enzymes Endo III and MutY functions in DNA-mediated charge-transport, which may provide a signalling mechanism to facilitate target site location or to communicate with other DNA-bound proteins that also have a redox-active [4Fe-4S] cluster.^{68–72}

Domain B is conserved in the four DME-ROS1 enzymes, but it does not appear to contain regions that are homologous to other HhH enzymes. While Domain B is required for glycosylase activity in DMEs, its function remains unknown.^{65–67} Notably, the variable regions joining the three conserved domains in DMEs, IDR1 and IDR2, are not essential for catalytic activity, and their function remains unclear. DME-ROS1 enzymes also have an N-terminal Lys rich region that enhances binding to nonspecific DNA and is required for efficient action on G:5mC pairs but not on G/T mismatches, as indicated by studies on ROS1.⁷³

2.3. Catalytic Challenge for Excision of 5mC

Before we consider mechanism(s) for enzymatic excision of 5mC from DNA, it is useful to review what is known about corresponding enzymatic and non-enzymatic reactions. The mechanism of *N*-glycosyl hydrolysis for deoxynucleotides, free and in DNA, has been studied by many approaches, including structure-activity relationships (SAR) and transition-state analysis via kinetic isotope effects (KIEs).^{32,74} The potential mechanisms can be grouped into two broad categories, including (i) a stepwise (two-step) mechanism, where departure of the nucleobase leaving group gives a discrete but short-lived oxacarbenium ion intermediate prior to nucleophile addition, or (ii), a concerted mechanism with a single transition state that features some degree of bond order to the nucleophile and leaving group (Figure 7).^{74,75}

Importantly, a stepwise mechanism has been observed in all studies to date of deoxynucleotide *N*-glycosyl hydrolysis using transition-state analysis (KIEs), including both enzymatic and non-enzymatic reactions. The studies comprise four reactions by unrelated enzymes, including two DNA glycosylases, UNG and MutY, which excise uracil and adenine from DNA, respectively.^{76,77} They also include excision of adenine from DNA by ricin (its biological target is an Ade in ribosomal RNA),⁷⁸ and *N*-glycosyl hydrolysis for thymidine deoxynucleotide by thymidine phosphorylase (performed in the absence of phosphate such that water is the nucleophile).⁷⁹ The KIE studies also include non-enzymatic *N*-glycosyl hydrolysis for the deoxynucleotide dAMP.⁸⁰ In addition, hybrid quantum-mechanical/molecular mechanical (QM/MM) studies also concluded the UNG reaction

follows a stepwise mechanism,⁸¹ confirming the conclusion from transition-state analysis.⁷⁶ Together, these rigorous experimental and theoretical findings suggest that, in addition to UNG and MutY, a stepwise mechanism may be employed for *N*-glycosyl hydrolysis by many, perhaps all DNA glycosylases.

The rate of deoxynucleotide *N*-glycosyl hydrolysis can be dramatically increased by catalysis of leaving group departure. The leaving group takes on increased electron density upon cleavage of the C-N bond, such that a neutral base (substrate) departs as a monoanion and a cationic base departs as a neutral species (Figure 8A). The departing anionic base can be highly unstable and is typically a good nucleophile, which may favor reformation of the ruptured C-N bond. Accordingly, *N*-glycosyl hydrolysis can be acid catalyzed, through protonation of a ring nitrogen, giving a leaving group that has an overall neutral charge. We will consider catalysis of leaving group departure for two pyrimidines of interest here, deoxycytidine (dCyd) and deoxyuridine (dUrd), and for their 5-substituted derivatives.

Non-enzymatic *N*-glycoside hydrolysis for dCyd is acid catalyzed via protonation at N3, which has a pK_a of 4.3 (Figure 8B).^{82–84} The same mechanism is observed for 5-methyl-dCyd (dMeCyd), which has a higher N3 pK_a of about 4.45, owing to the electron-donating properties of its methyl substituent.⁸² The pH profile (log rate versus pH) for these reactions indicates that the neutral form of these deoxynucleosides is not hydrolyzed. This likely reflects the fact that the N1 anion of C and 5mC are highly unstable in solution, as indicated by an N1 pK_a of 12.2 for C (and higher for 5mC).⁸⁵ As such, the *N*-glycoside hydrolysis reactions are extremely slow at pH 7.4, given the low population of the reactive, N3-protonated species. Indeed, non-catalyzed hydrolysis of dCyd and dMeCyd is much slower at neutral pH compared to the other deoxynucleotides in DNA (dThd, dAdo, dGuo).^{74,82} Notably, the rate *N*-glycoside hydrolysis is even slower for dCyd in DNA, relative to free dCyd,⁷⁴ and a similar trend is likely for 5mdCyd.

These findings suggest that enzymatic excision of C and 5mC from DNA could also be acid catalyzed via N3 protonation, but little is currently known about the mechanism of such reactions. A variant of uracil DNA glycosylase (UNG) removes C from canonical G:C pairs,^{86,87} and the mechanism appears to involve N3 protonation. The Glu side chain introduced by the mutation is proposed to serve as the general acid that protonates C at N3, though it might also stabilize an N3-protonated species arising from a solvent-derived proton.

In contrast to dCyd, non-enzymatic *N*-glycoside hydrolysis for dUrd is not acid catalyzed at neutral pH, and the reaction features the expulsion of a uracil anion (Figure 9A).^{74,81,82} This is favored in part by resonance stabilization of the Ura anion, via charge delocalization from N1 to O2, and O4 (Figure 9B). Electron-withdrawing substituents at C5 (e.g., F, Br) stabilize the departing anion and increase the reaction rate.^{32,82,88} Protonation of dUrd is highly unfavorable, with a pK_a of about -3.4 at O4 (Figure 9C).⁸⁰ As such, acid catalysis of dUrd hydrolysis is ineffective at neutral pH, but is observed as the pH approaches the pK_a .

Following suit, the enzymatic reaction mediated by UNG does not employ acid catalysis of uracil expulsion. Rather, UNG stabilizes the departing uracil anion using electrostatic

catalysis via interactions with the two carbonyl oxygens, including a strong hydrogen bond to O2.^{81,89–92} Similarly, evidence suggests that other glycosylases which excise Ura, or 5-substituted Ura analogues, do so by stabilizing a departing anionic base through electrostatic catalysis,^{32,88,93–95} rather than acid catalysis via protonation of the departing base.

2.4. Mechanism of 5mC DNA Glycosylases

Given this background, we next consider potential mechanisms for 5mC excision by the DME-ROS1 enzymes. Our focus is on the glycosylase (base excision) reaction rather than the subsequent lyase reactions, as the latter are likely common to other bifunctional enzymes (Figure 3). Some studies suggest that a conserved Lys may be the nucleophile for the glycosylase reaction of DME-ROS1 and other bifunctional HhH enzymes (Figure 10A), but it is also possible that the nucleophile is a water molecule (Figure 10B), as is the case for monofunctional DNA glycosylases. Regarding the conserved Lys, studies of the related bifunctional enzymes Endo III and OGG1 demonstrate that the Lys forms the covalent enzyme-substrate interaction (Schiff base) that is required for the lyase reaction.^{96,97} Sequence homology indicates that the same is likely true for the conserved catalytic Lys of DME-ROS1 enzymes (Lys¹²⁸⁶ of DME, Lys⁹⁵³ of ROS1, *Arabidopsis*).^{40,65–67} While the conserved Lys has been shown to be important for the glycosylase reaction of DME-ROS1, Endo III and OGG1, as indicated by large mutational effects,^{45,60,67,96,98} this does not require that it serves as the nucleophile for base excision. Indeed, some studies indicate that the Lys mediates substrate binding or perhaps another role, and that a water molecule serves as the nucleophile, yielding an AP site as the product for base excision (Figure 10B).^{60,99}

Notably, DME-ROS1 enzymes have a conserved Asp-Asn pair (e.g., Asp¹³⁰⁴, Asn¹³⁰⁶, *Arabidopsis* DME) corresponding to the Asp-Asn pair that binds the putative nucleophilic water molecule in the reaction catalyzed by MutY (Asp¹⁴⁴, Asn¹⁴⁶, *Bst* MutY), a monofunctional HhH glycosylase that employs hydrolytic base excision,^{65,66,100} raising the possibility that DME-ROS1 enzymes use a similar mechanism. The Asp of this pair is a signature residue of HhH glycosylases and one of its functions is to form a helix cap.²⁰ Other proposed roles for the conserved Asp include stabilizing the putative oxacarbenium ion intermediate or binding and/or activating the nucleophilic water molecule (or Lys side chain).^{20,74}

Some studies suggest that OGG1 functions in a monofunctional mode *in vivo*, releasing AP DNA that is processed by downstream BER enzymes.^{55–60} However, this does not necessarily require that a water molecule is the nucleophile for the glycosylase reaction, because the covalent enzyme-DNA product generated from a glycosylase reaction with a Lys nucleophile could be hydrolyzed to yield an AP DNA product (Figure 10A). Thus, additional studies are needed to establish whether the conserved Lys or a water molecule is the nucleophile in the glycosylase reaction for bifunctional HhH glycosylases, including the DME-ROS1 enzymes. Nevertheless, we consider potential mechanisms for the glycosylase reaction below, including their consistency with experimental findings reported to date.

As discussed above (Section 2.3), a stepwise reaction mechanism has been observed for all deoxynucleotide *N*-glycosyl hydrolysis reactions examined to date using transition-state analysis (TSA) and QM/MM theoretical methods.⁷⁴ Importantly, one TSA study focused on

MutY,⁷⁷ a member of the HhH superfamily that is homologous to DME-ROS1 and other bifunctional glycosylases Endo III and OGG1. Indeed, crystal structures reveal a very similar active-site configuration for MutY, Endo III, and OGG1.^{100–105} As such, we limit the discussion to some of the potential stepwise mechanisms that might be employed by DME-ROS1 enzymes, involving a Lys side chain or a water molecule as the nucleophile. While concerted mechanisms are not discussed, for the reasons articulated above, a concerted mechanism cannot presently be excluded.

One stepwise mechanism that has been proposed for OGG1 features product-assisted catalysis, where the expelled 8oxoG anion deprotonates the ammonium ($-\text{NH}_3^+$) of the putative Lys nucleophile, activating it for addition to the oxacarbenium ion intermediate.^{74,106} We note that this mechanism is not based on transition-state analysis for OGG1, but inspired by such studies showing that UNG employs a stepwise mechanism.^{76,77,81} Here, we explore similar stepwise mechanisms for DME-ROS1 enzymes involving a Lys nucleophile (Figure 10A), and consider their compatibility with previous biochemical findings. We also consider a stepwise mechanism involving a water molecule as the nucleophile (Figure 10B).

Similar to the proposal for OGG1, DME-ROS1 enzymes could potentially catalyze a stepwise reaction with expulsion of an anionic base (Figure 10A, “anionic”). While the N1 anion of 5mC is not stable in solution, given an N1 $\text{p}K_a$ of over 12.2,⁸⁵ it might be stabilized in the active site, through charge delocalization to N1, N3, and O2, and enzyme contacts to these nuclei (electrostatic catalysis), including perhaps the cationic side chain of the catalytic Lys. As proposed for OGG1, the expelled anionic base might deprotonate the Lys nucleophile, activating it for addition to the oxacarbenium ion intermediate. Given the previous findings for non-enzymatic *N*-glycoside hydrolysis of dCyd and dMeCyd (section 3.3), the DME-ROS1 enzymes could employ acid catalysis of 5mC expulsion via N3 protonation, such that 5mC departs as a neutral species (Figure 10A, “neutral” LG). While acid catalysis would activate the base for departure, deprotonation of the Lys nucleophile would probably be less effective for the neutral versus anionic 5mC leaving group. However, a Lys nucleophile might be activated by another group or have a reduced $\text{p}K_a$ such that it is neutral in the active site. Moreover, the nucleophile might be a water molecule, which does not necessarily need to be activated for addition to an oxacarbenium ion intermediate.

These stepwise mechanisms seem consistent with findings from previous biochemical studies of DME-ROS1 enzymes in which the methyl of 5mC (or of thymine) was replaced with an electron-withdrawing halogen (5BrC, 5BrU).⁶³ ROS1 activity is much lower for 5BrC relative to 5mC, and for 5BrU versus T, and it was proposed that the enzyme specifically recognizes the C5-methyl of 5mC (and T) via steric and/or electrostatic effects.⁶³ This model is supported by findings that ROS1 efficiently removes 5mC and thymine (from G/T mispairs) but does not excise the methyl-deficient counterparts, Cyt or Ura.⁶³ However, the Br substituent is relatively isosteric with methyl, and other factors could account for the previous findings.

The enhanced ROS1 activity for 5mC versus 5BrC seems consistent with a stepwise mechanism involving acid-catalyzed base expulsion (Figure 10A “neutral”, Figure 10B).

The electron-withdrawing Br substituent will reduce the N3 pK_a for 5BrC versus 5mC ($pK_a \sim 1.6$),⁸⁵ rendering acid catalysis less effective at neutral pH, as observed for corresponding non-enzymatic reactions.⁸² The 5Br substituent could also have an adverse effect on the subsequent Lys nucleophile activation step, by reducing the N1 pK_a such that N1-mediated deprotonation of a Lys nucleophile would likely be less effective for 5BrC versus 5mC. However, nucleophile activation by a departing base leaving group would likely not be a factor for a mechanism in which a water molecule is the nucleophile.

A mechanism involving rate-limiting departure of an anionic base would be expected to give a faster rate for 5BrC versus 5mC, in contrast with the biochemical findings. However, if the rate-limiting step in such a mechanism is activation of a Lys nucleophile by the expelled anionic base, then a slower rate might be expected for excision of 5BrC versus 5mC. The Br substituent will reduce the N1 pK_a , rendering Lys nucleophile activation less effective for the anion of 5BrC versus 5mC.

The proposed mechanisms also seem consistent with findings that DME excises 5-hydroxymethylcytosine (5hmC) with efficiency similar to 5mC,⁶⁷ given the similar substituent electronic effects (σ_m is -0.07 for CH_3 and 0.0 for CH_2OH).¹⁰⁷ Findings that DMEs do not excise 5-formylcytosine⁶⁷ are also consistent with our proposed mechanisms, for reasons similar to those cited above regarding 5BrC. The lack of 5fC activity might be explained by the strong electronic effect of the formyl substituent (σ_m is 0.35), and enhanced resonance stabilization for 5fC conferred by the formyl group. These effects greatly reduce the N1 (and N3) pK_a for 5fC versus 5mC,⁸⁴ reducing its ability to deprotonate the Lys nucleophile. For similar reasons, the mechanism seems consistent with findings that ROS1 activity is slower for 5BrU (and 5FU) relative to T (i.e., 5mU).

Thus, the potential mechanisms in Figure 10 seem consistent with previous biochemical studies of DME-ROS1 enzymes and previous mechanistic studies of glycosylase-mediated and non-enzymatic *N*-glycosyl hydrolysis reactions. Of course, additional studies will be needed to test these and other potential catalytic mechanisms for DME-ROS1 enzymes.

3. BER MEDIATES ACTIVE DNA DEMETHYLATION VERTEBRATES

Cytosine methylation is also a key epigenetic signal in animals (vertebrates) and it is critical for processes including X-chromosome inactivation, silencing of imprinted genes and repetitive DNA elements such as retrotransposons, and it plays a role in transcriptional regulation.^{108–111} Aberrant patterns of methylation have been observed in tumors and are implicated in carcinogenesis.^{112,113} While cytosine methylation occurs predominantly at CG dinucleotides (CpG sites) in higher eukaryotes, methylation at non-CpG sites is becoming increasingly recognized, particularly in neurons.³⁶ It is well established that modification of C to 5mC is mediated by one of three *S*-adenosylmethionine (SAM)-dependent DNA methyltransferase (DNMT) enzymes, including DNMT1, DNMT3A, or DNMT3B.¹¹⁴ However, the process by which this epigenetic mark can be erased has only recently become clear, and is an active area of current research.

3.1. Potential BER-Mediated Pathways for Active DNA Demethylation

Early studies suggested that MBD4 (methyl binding domain IV) and TDG can directly excise 5mC from DNA,^{115,116} but subsequent work has shown that the 5mC glycosylase activity of these enzymes is very weak and not biologically significant.^{88,117,118} While the vertebrate DNA glycosylases cannot directly remove 5mC, they can excise derivatives of 5mC generated by deamination and/or oxidation.^{5,119} Accordingly, the potential pathways for active DNA demethylation feature enzymatic modification of 5mC, and subsequent BER processing of the 5mC derivative (Figure 11).^{5,34}

To date, the only biochemically and biologically validated demethylation pathway is that initiated by one of three different TET (ten-eleven translocation) enzymes, TET1, 2, and 3 (Figure 11).^{43,120–124} These α -ketoglutarate-Fe(II)-dependent dioxygenases, catalyze the oxidation of 5mC to give three different products, 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC), and 5-carboxylcytosine (5caC). The BER component of this pathway is initiated by thymine DNA glycosylase (TDG), which removes 5fC and 5caC (but not 5hmC) from DNA.^{43,44} Follow-on BER enzymes then convert the TDG-generated AP site to a G:C base pair, completing the demethylation process. We consider below several proposed DNA demethylation pathways that involve BER, and then focus in more detail on the TET-TDG-BER pathway.

Some studies have suggested that DNA demethylation could involve active deamination of 5mC to T by a deaminase enzyme, excision of T from the resulting G/T mispair by TDG or MBD4, and subsequent BER processing to give unmodified cytosine.^{118,125–131} In these models the deamination activity involves a member of the activation-induced deaminase (AID)/APOBEC family. Similarly, some studies suggest demethylation via active deamination of 5hmC to give 5-hydroxymethyluracil (5hmU), excision of 5hmU by TDG, MBD4, or SMUG1, and follow-on BER to give cytosine.^{117,131} 5hmU is efficiently removed by all three of these DNA glycosylases.^{88,132–135} Notably, 5fU and 5caU, the products resulting from deamination 5fC and 5caC, would be likely be readily excised by one or more mammalian glycosylases, including TDG, SMUG1, NTH1, and perhaps MBD4.^{88,94,132,133,136–142} However, deamination 5fC and 5caC has not yet been implicated in DNA demethylation.

Unlike the TET-TDG pathway, direct biochemical evidence for deamination-mediated demethylation is not particularly strong. The AID/APOBEC enzymes prefer single strand DNA and typically exhibit preferences for certain DNA sequences (hotspots) or have other features that suggest they are poorly suited for a central role in active DNA demethylation.^{34,143} While these enzymes exhibit robust activity for cytosine in DNA, their canonical substrate, their activity is typically much lower for 5mC and is not detected for 5hmC.^{34,144–148} This can be explained by findings that deaminase activity is reduced with increasing steric bulk at the C5 position for a series of 5-substituted cytosine substrates in DNA.^{144,145} Humans have eight APOBEC enzymes, APOBEC1 and seven APOBEC3 variants.¹⁴⁹ Notably, APOBEC3A can deaminate 5mC to T, and this activity is comparable to that for deamination of C to U.^{146,147} However, APOBEC3A expression patterns suggest a role in foreign DNA restriction.¹⁴⁷ Thus, while additional studies are needed, the findings

reported to date suggest that active deamination of 5mC or 5hmC is not likely to be a predominant mechanism for initiating DNA demethylation in vertebrates.

3.2. TET-TDG-BER Pathway for DNA Demethylation

In contrast to the findings for a putative deaminase-mediated pathway, purified TET enzymes readily convert 5mC to 5hmC, 5fC, and 5caC for DNA substrates *in vitro*, and TET-mediated formation of these 5mC derivatives can be detected in the genomic DNA of mammalian cells.^{43,120–122,144,150–153} Notably, the abundance of 5fC and 5caC in genomic DNA is enhanced by depletion of TDG, consistent with its efficient excision of these bases *in vitro*.^{43,44} TET enzymes can catalyze the three different oxidation reactions in a stepwise manner,^{43,122,154} or they can generate 5fC and 5caC iteratively, by acting in a single encounter with 5mC-containing DNA (without releasing the 5hmC or 5fC intermediate).¹⁵⁵ Based on the mechanism of other Fe(II)- and α -ketoglutarate-dependent dioxygenases,^{119,156} structures of TET enzymes suggest that they catalyze oxidative decarboxylation of α -ketoglutarate, using molecular oxygen, giving a reactive, enzyme-bound, high-valent Fe(IV)-oxo intermediate that converts 5mC to 5hmC.^{157–160} Thus, the TET-initiated steps of active DNA demethylation have been demonstrated in multiple biochemical and cellular studies.

It is also possible that an enzyme possessing decarboxylase activity directly converts 5caC to cytosine (Figure 12), which would avoid BER intermediates that can potentially be mutagenic and toxic.^{161–163} Although 5caC decarboxylase activity was reported in stem cells, the responsible enzyme(s) was not identified.¹⁶² One study finds that DNA methyltransferases, including the human Dnmt3A/B Dnmt3L complex, converts 5caC to C *in vitro*, but the activity was reduced in the presence of physiological levels of *S*-adenosylmethionine (SAM), a cofactor for DNA methyltransferases, so the biological relevance remains to be established.¹⁶⁴ Studies of non-enzymatic decarboxylation of 5caC (in DNA) revealed an efficient thiol-mediated, acid catalyzed reaction.¹⁶⁵ The proposed mechanism involves attack at the C6 position of 5caC by a Cys side chain, giving an activated intermediate with a saturated C5-C6 bond that leads to decarboxylation, thiol expulsion, and re-aromatization (Figure 12).^{162,164,165} Notably, the catalytic Cys is essential for 5caC decarboxylation by the methyltransferases.¹⁶⁴ A similar mechanism has been proposed for dehydroxymethylation (conversion of 5hmC to C) by cytosine methyltransferases.^{165–167} Non-enzymatic deformylation (5fC to C) is much less efficient than decarboxylation, and non-enzymatic dehydroxymethylation is slower still.¹⁶⁵

Biochemical and biological studies demonstrate that TET-generated 5fC and 5caC can be converted to cytosine through TDG-initiated BER. Indeed, TDG, which was discovered as a G/T mismatch glycosylase,^{22,23} has robust activity for excising fC and caC from DNA.^{43,44} Importantly, two studies have demonstrated that TDG, and its glycosylase activity in particular, is essential for development in mice.^{117,118} This phenotype likely reflects a critical role in active DNA demethylation, rather than a need for repair of spontaneously generated G/T mispairs. It is important to note that no other mammalian DNA glycosylases have been found to be essential for embryonic development. In addition, no other mammalian DNA glycosylases have been shown to have significant biochemical activity for

excision of 5fC or 5caC *in vitro*.^{26,44,132} It was reported that UNG2 may function in TET-initiated DNA demethylation, but it is unclear what role it may play, given that purified UNG2 does not remove 5fC or 5caC from DNA *in vitro*.^{132,168}

A reconstituted TET1-TDG-BER pathway, comprised of purified TET1 (catalytic domain; TET1_{CD}), TDG, APE1, DNA Pol β , and XRCC1-LIG3, can convert a G:5mC base pair to a G:C pair in duplex DNA.¹²⁴ Findings that TET enzymes interact with TDG raises the possibility that their activity might be coordinated to some extent *in vivo*.^{124,169} This coordination could be as simple as TET recruitment of TDG to sites of demethylation, or might involve a TET-TDG oxidation-excision complex. A number of studies have implicated Gadd45 (Growth Arrest and DNA Damage 45) in active DNA demethylation,^{117,131,170,171} though a role for Gadd45 has been controversial.¹⁷² Recent studies find that Gadd45a interacts with and enhances the activity of TET enzymes and TDG.^{173,174}

One report raised the possibility that the bifunctional NEIL (nei endonuclease VIII-like) glycosylases might excise 5fC or 5caC from DNA, since they seemed to partially compensate for the loss of TDG in TET1-initiated expression of a reporter gene silenced by methylation.¹⁶⁹ But this study did not demonstrate such activity for NEIL enzymes *in vitro*, and other studies show that purified NEIL1 and NEIL2 lack significant 5fC or 5caC glycosylase activity.^{132,175} However, NEIL enzymes could contribute in another way to a TET-TDG-BER demethylation pathway. NEIL1 and NEIL2 appear to process the abasic sites generated by TDG excision of 5fC and 5caC, in HeLa whole cell extracts, and they enhance the catalytic turnover of TDG for processing 5fC and 5caC substrates *in vitro*.¹⁷⁵ Many studies show that APE1 greatly enhances the catalytic turnover of TDG for 5caC and other substrates *in vitro*.^{176–179} However, depletion of APE1 in HeLa cells was not found to cause an elevation of genomic 5fC and 5caC in HeLa cells, whereas 5fC and 5caC levels were elevated upon depletion of NEIL1 and NEIL2.¹⁷⁵ Notably, NEIL enzymes bind to TDG,¹⁷⁵ and to the TET enzymes.¹⁶⁹ The TET enzymes also interact with BER proteins LIG3 and XRCC1.¹⁶⁹

Thus, the post-TDG steps of DNA demethylation (Figure 12) could involve at least two potential pathways, with APE1 or one of the NEIL enzymes acting on the TDG-generated AP site. Observations that both APE1 and NEIL enzymes stimulate the catalytic turnover of TDG *in vitro*, for several different substrates,^{176–179} is consistent with other studies indicating that efficient catalytic turnover of TDG does not require SUMO modification of product-bound TDG,^{176,180} in contrast with previous proposals.^{181–183}

3.3. Enzymatic Excision of 5fC and 5caC from DNA

Given that the BER component of active DNA demethylation begins with TDG excision of 5fC or 5caC from DNA, we turn our attention to these two enzymatic reactions, which happen to exhibit rather different mechanisms. Our focus will be limited to the TDG-mediated reactions, because the subsequent BER reactions are likely to involve mechanisms that are largely general to the pathway. We begin by considering how the chemical properties of 5fC and 5caC impact their enzymatic excision from DNA, and mechanisms of relevant enzymatic and non-enzymatic reactions.

As discussed above (Section 2.3), nonenzymatic *N*-glycoside hydrolysis of dCyd is acid catalyzed via N3 protonation, such that the departing base is neutral rather than anionic (section 2.3). Acid catalysis was also observed for hydrolysis of 5-Br-substituted dCyd (dBrCyd). The electron-withdrawing Br substituent renders N3 and N1 more acidic, with two counteractive effects on the reaction rate. Acid catalysis is less effective, because the N3-protonated species is less abundant, while leaving-group quality, which increases with N1 acidity, is better for 5BrC versus C. Given these counteracting effects, *N*-glycosyl hydrolysis is found to occur at similar rates for dCyd and dBrCyd. We will consider the substituent effects on N1 and N3 as we examine how the formyl and carboxylate groups impact base excision. First, we consider the role of substituent effects on the non-enzymatic and TDG-catalyzed hydrolysis of dUrd and its derivatives.

Compared to findings for *N*-glycosyl hydrolysis of dCyd, the effect of C5 substituents is less complicated for non-enzymatic dUrd *N*-glycosyl hydrolysis, since it is not acid catalyzed (section 2.3). The reaction rate increases with leaving-group quality (N1 acidity) of the departing base, giving a Brønsted-type linear free energy relationship (LFER) (Figure 13).^{32,82,88} An electron-withdrawing C-5 substituent confers faster hydrolysis by stabilizing the departing anionic base.

Similar findings were obtained for TDG-catalyzed reactions, with key implications for its excision of 5fC and 5caC.^{44,88} A Brønsted-type LFER was observed for TDG excision of C5-substituted uracils and cytosines from DNA (Figure 14).⁸⁸ For uracil substrates, the TDG reaction is not acid catalyzed and the base likely departs as an anion, stabilized by active-site interactions (electrostatic catalysis),^{88,184–186} similar to findings for UNG.^{89–91,187} The results indicated that TDG excision of 5-fluoroC (5fC) and 5BrC is also not acid catalyzed, and, more generally, that TDG can catalyze departure of a cytosine anion if is stabilized by an electron-withdrawing substituent.⁸⁸

Observation that a halogen substituent at C5 activates cytosine for excision by TDG indicated that the newly recognized epigenetic base, 5fC, with its electron-withdrawing formyl group, could be a good TDG substrate. Indeed, TDG rapidly excises 5fC from DNA.⁴⁴ In fact, TDG activity is 17-fold greater for 5fC compared to 5-fluoroC, even though the substituent electronic effects are the same (σ_m is 0.35 for CHO and 0.34 for F). This is likely explained by more extensive resonance stabilization for the 5fC anion, including charge delocalization to the formyl oxygen (Figure 15A). This stabilizes the N1 anion, enhancing N1 acidity (leaving group quality) for 5fC relative to 5FC.⁸⁴ In addition, TDG may stabilize negative charge that develops on the formyl and the O2 oxygen of the departing 5fC anion, as suggested by structures of the enzyme-substrate (ES) complex for other TDG substrates (U, 5caC).^{185,188} However, confirmation of this model awaits a crystal structure of TDG bound to 5fC-containing DNA.

Findings that 5fC is likely an inherently good substrate for glycosylase excision begs the question of why this base is not excised by other DNA glycosylases such as MBD4. Structural and biochemical studies of MBD4 indicate its lack of 5fC and 5caC activity is likely not due to steric hindrance, because MBD4 can accommodate and excise uracil analogs with bulky groups at C5.^{140,189} Rather, it seems likely that stable flipping of

cytosine analogs may be hindered by electrostatic interactions that favor a carbonyl rather than an exocyclic amino at the C4 position.^{132,133}

Unlike 5fC, it turns out that 5caC is not an inherently good substrate for excision by TDG. It is now well established that at physiological pH, 5caC exists as a monoanion, owing to its carboxylate (-COO^-) substituent (Figure 15B). The 5caC anion is converted to a neutral species at low pH, with a $\text{p}K_a$ of ~ 4.3 (Figure 15B).^{84,190,191} This is highly relevant, because carboxylate is an electron donating substituent ($\sigma_m -0.10$ for -COO^-),¹⁹² and the dianionic form of 5caC that would result from C-N cleavage is expected to be a very poor leaving group. Accordingly, theoretical calculations reveal weak N1 acidity for the 5caC anion (Figure 15B), also indicating that it is likely a poor leaving group.⁸⁴ Indeed, the 5caC anion is less acidic (N1) than C and 5hmC (Figure 15B), which cannot be removed by TDG.⁸⁴ Together, these observations suggest TDG cannot excise the 5caC anion, the form of 5caC that predominates under physiological conditions. However, theoretical calculations reveal that the three neutral forms of 5caC have far greater N1 acidity than the 5caC anion and should be much more amenable to enzymatic excision.⁸⁴

Consistent with the N1 acidity calculations, TDG excision of 5caC is acid catalyzed, where activity increases with decreasing pH over the range of pH 5.5 to 9.25 (Figure 16).^{84,134} By contrast, excision of 5fC is not acid catalyzed (Figure 16), which is expected given that 5fC is protonated (at N3) with a $\text{p}K_a$ of 2.6,^{83,191} far below the experimentally accessible pH range for TDG (inactive for $\text{pH} < 5.5$).^{84,184} Similarly, TDG does not employ acid catalysis for excision T from G-T (or U from G-U) mispairs.¹⁸⁴ Thus, 5caC is the only known substrate for which TDG employs acid catalysis.

The pH profile for 5caC excision by TDG, collected under conditions that report on events in the enzyme-substrate (E-S) complex, reveals an essential protonated group that ionizes with an apparent $\text{p}K_a$ of 5.8.⁸⁴ The results indicate that this essential protonated group is the 5caC base, flipped into the TDG active site, rather than a TDG side chain acting as a general acid. Thus, TDG flips the 5caC anion out of the DNA duplex and protonates it in the active site, converting it to a neutral species to enable *N*-glycosyl bond cleavage. This appears to involve a proton derived from solvent in the active site. Consistent with this idea, high-resolution crystal structures of TDG reveal water molecules in the active site and a solvent-filled channel from the enzyme surface to the active site, for DNA-bound TDG.¹⁹³

It is not presently known which of the three neutral species of 5caC is formed in the TDG active site prior to C-N cleavage, the zwitterion, the imino, the amino, or perhaps more than one of these forms (Figure 15B). Crystal structures indicate that the Asn191 side chain of TDG contacts the N3 and the exocyclic NH_2 of 5caC (Figure 17).¹⁸⁸ Importantly, the N191A-TDG mutant binds normally to a G:5caC DNA substrate but has no detectable 5caC excision activity. By contrast, N191-TDG has full activity for 5fC, which is excised without the need for acid catalysis. These findings suggest that Asn191 may be needed to stabilize one of the two neutral 5caC species that feature an N3-H moiety (zwitterion or imino, Figure 15B). However, additional studies are needed to determine which neutral form of 5caC is formed in the active site prior to C-N bond cleavage.

Two studies examining the ionization of the free d5caCyd deoxynucleoside are relevant.^{84,191} Both studies find two ionization events with decreasing pH; the first converts the anion to a neutral species and the second gives a monocation. However, the studies differ with regard to which neutral species of d5caCyd is generated by protonation of the anion. Studies employing UV absorbance suggested that d5caCyd anion is protonated at N3 (pK_a 4.3) to give the zwitterion (Figure 16), which is subsequently protonated at the carboxylate (pK_a 2.5) to give a the cation.⁸⁴ A higher pK_a for N3 (rather than $-\text{COO}^-$) is consistent with the idea that the TDG active site likely stabilizes a neutral species with N3-H (zwitterion or imino). A study employing FTIR spectroscopy finds similar pK_a values but with opposite ionization assignments, concluding that the anion-to-neutral transition involves protonation of the exocyclic carboxylate and the neutral-to-cation transition involves N3 protonation.¹⁹¹ This raises the possibility that TDG may protonate the 5caC anion at the carboxylate rather than at N3.¹⁴⁸ Regarding the free d5caCyd nucleoside, additional studies are needed to determine which neutral species of d5caCyd, perhaps more than one, predominates upon protonation of the monoanion. Of course, it is possible that the neutral form(s) of 5caC that is generated in the TDG active site differs from that generated at low pH (pK_a 4.3) for the free d5caCyd deoxynucleoside in solution.

3.4. Properties of 5fC and 5caC in DNA that Could Favor Enzymatic Excision

It was proposed that 5fC and 5caC adopt the rare imino tautomeric state and thereby form a wobble structure when forming base-pairs with Gua in DNA, similar to G·T and G·U mispairs, and that a wobble structure is unifying feature of substrate recognition by TDG (Figure 18).^{134,194} However, theoretical calculations reveal much lower stability for the imino versus the canonical amino forms of d5fCyd and d5caCyd nucleosides, suggesting the imino forms are rare, as observed for dCyd.⁸⁴ Moreover, experimental studies show that the amino form predominates for d5fCyd and d5caCyd nucleosides, and NMR and crystallographic studies of duplex DNA show that G:5fC and G:5caC base pairs form the canonical Watson-Crick structure rather than the wobble structure that would result from the imino tautomer of 5fC or 5caC.^{191,195–197} While the studies do not exclude the possibility that G:5fC and G:5caC base pairs adopt a transient wobble structure, the canonical Watson-Crick structure predominates in DNA.

It was suggested that the pH dependence of TDG 5caC excision can be explained by a pH-dependent weakening of G:5caC base pairs in DNA, which has been observed for free DNA.¹⁹¹ This idea stems from findings that a 10 bp DNA duplex containing six G:5caC pairs exhibits pH-dependent melting; DNA duplex stability (T_m) decreases with pH, due to protonation of the 5caC anion. However, the stability (T_m) of this 5caC-rich duplex was constant (independent of pH) for pH above 5.5, which is the range for which pH-dependence of 5caC excision by TDG is observed (pH 5.5 to 9.0).⁸⁴ In fact, TDG is unstable (and inactive) in solution for pH below 5.5.¹⁸⁴ Thus, the pH-dependent weakening of G:5caC base pairs, observed only for pH <5.5, cannot account for the pH dependence of 5caC excision activity, for pH >5.5. Rather, the pH dependence of 5caC excision is most reasonably explained by acid catalysis of 5caC departure, as described above.⁸⁴

It has also been proposed that the specificity of TDG for excising 5fC, but not C, 5mC, or 5hmC can be explained by weaker guanine base-pairing for 5fC, due to the electron-withdrawing formyl group.¹⁹¹ Observation that the formyl substituent increases N3 acidity for 5fC relative to C (and 5mC, 5hmC) does indicate reduced electron density at N3, suggesting weaker hydrogen bonding to the imino N-H of Gua. However, multiple studies indicate that the formyl group has a fairly small effect on base-pairing stability, and the presence of G:5fC base pairs in DNA has been found to cause both increases or decreases in duplex stability (T_m), depending on the study. In one example, the T_m for a 10 bp DNA duplex containing six G:5fC pairs was reduced by 3 °C (as determined by UV absorbance),¹⁹¹ a small change given that the DNA was comprised of 60% G:5fC pairs. Studies using the Dickerson-Drew dodecamer (DDD), with two internal cytosine bases replaced by 5mC, 5hmC, or 5fC, found no large differences in T_m (C, 48 °C; 5mC, 46 °C; 5hmC, 48 °C; 5fC, 46 °C).¹⁹⁷ Three other studies report the same or modestly higher T_m for DNA in which one or two G:C pairs was replaced with G:5fC.^{190,196,198}

Thus, there are no clear and consistent findings that G:5fC pairs have substantially reduced stability compared to G:C, G:5mC, and G:5hmC pairs. As such, it seems unlikely that small differences in base-pairing stability account for huge differences in TDG excision activity for G:5fC relative to G:C, G:5mC, and G:5hmC. Indeed, TDG activity is at least 50,000-fold greater for G:5fC relative to G:C or G:5hmC, where the DNA substrate contains only a single G:5fC site.⁴⁴ By contrast, there is a good correlation between TDG activity and N1 acidity (leaving-group quality) for the cytosine bases (Figure 15B), which provides a reasonable explanation for the observed specificity.⁸⁴ It seems likely that specificity is also due in part to substrate binding, which is much tighter for G:5fC (and G:5caC) relative to G:C, G:5mC, and G:5hmC.¹⁸⁸ Additional studies are needed to uncover the basis of these differences in binding affinity.

Findings that a single G:5fC pair can increase the flexibility of DNA raises the possibility that this may contribute to specificity for G:5fC pairs, although G:5caC pairs were not found to significantly effect DNA flexibility.¹⁹⁹ Other structural and biophysical studies using a DNA dodecamer that contained six consecutive G:5fC pairs found changes in DNA groove geometry and the base pairs associated with 5fC, and underwinding of the DNA.¹⁴⁴ Such changes could potentially contribute to 5fC recognition by TDG, though the effects of 5fC on DNA structure seem likely to be far smaller for a single G:5fC pair (compared to six consecutive pairs). Additional studies are needed to investigate whether and how the effects of 5fC on DNA structure contribute to recognition of 5fC in DNA by TDG.

4. CONCLUDING REMARKS

As summarized above, there has been much progress in our understanding of how BER mediates active DNA demethylation in both plants and animals, but many important questions still need to be addressed. In plants, the mechanism employed by DME-ROS1 enzymes to excise 5mC remains mysterious and will be an important area for future research. Elucidating this mechanism will require rigorous biochemical approaches and could be greatly advanced by crystal structures, ideally of an enzyme-substrate complex.

Other questions include how the 5mC excision activity of DME-ROS1 enzymes might be controlled to regulate the initiation of active DNA demethylation.

Important questions also remain for understanding how BER mediates active DNA demethylation in vertebrates. The mechanism by which TDG excises 5fC and 5caC from DNA remains unresolved. While 5fC seems to be an inherently good substrate, additional structural and biochemical studies are needed to understand how it is rapidly excised by TDG, but not other glycosylases such as MBD4. Studies to date indicate acid catalyzed 5caC excision, but does this involve protonation at N3 or the carboxylate group? Many other questions also remain to be addressed. For example, how does TDG-initiated BER avoid generating toxic double strand breaks (DSBs) when processing TET-generated 5fC or 5caC bases that are located close together on opposite strands? How is TDG activity on 5fC and 5caC, hence active DNA demethylation, regulated? Can TDG efficiently excise 5fC and 5caC from non-CpG sites in DNA? It is well known that TDG activity is greatly reduced for G/T mispairs in a non-CpG context, but it is unclear if the same is true for excision of 5fC and 5caC. We anticipate that these and many other important questions will be addressed in future studies of BER-mediated active DNA demethylation.

Acknowledgments

Studies in our laboratory have been supported by National Institutes of Health grant GM72711 (to ACD).

ABBREVIATIONS

5mC	5-methylcytosine
5hmC	5-hydroxymethylcytosine
5fC	5-formylcytosine
5caC	5-carboxylcytosine
5hmU	5-hydroxymethyluracil
5FU	5-fluorouracil
5fU	5-formyluracil
5caU	5-carboxyluracil
AID	activation-induced deaminase
BER	base excision repair
DSB	double strand break
LINE-1	Long interspersed element-1
MBD4	methyl binding domain IV
MMR	mismatch repair

NEIL	nei endonuclease VIII-like
OGG1	8-oxoguanine DNA glycosylase 1
Pol β	DNA polymerase beta
SMUG1	single-strand selective monofunctional uracil DNA glycosylase
SSB	single strand break
TDG	thymine DNA glycosylase
TET	ten-eleven translocation
UNG	uracil DNA glycosylase

References

1. Lindahl T, Wood RD. Quality Control by DNA Repair. *Science*. 1999; 286:1897–1905. [PubMed: 10583946]
2. Wallace SS. Base Excision Repair: A Critical Player in Many Games. *DNA Repair (Amst)*. 2014; 19:14–26. [PubMed: 24780558]
3. Krokan HE, Bjoras M. Base Excision Repair. *Cold Spring Harb Perspect Biol*. 2013; 5:a012583. [PubMed: 23545420]
4. Kim YJ, Wilson DM 3rd. Overview of Base Excision Repair Biochemistry. *Curr Mol Pharmacol*. 2012; 5:3–13. [PubMed: 22122461]
5. Kohli RM, Zhang Y. Tet Enzymes, Tdg and the Dynamics of DNA Demethylation. *Nature*. 2013; 502:472–9. [PubMed: 24153300]
6. Wyatt MD. Advances in Understanding the Coupling of DNA Base Modifying Enzymes to Processes Involving Base Excision Repair. *Adv Cancer Res*. 2013; 119:63–106. [PubMed: 23870509]
7. Fu Y, He C. Nucleic Acid Modifications with Epigenetic Significance. *Curr Opin Chem Biol*. 2012; 16:516–24. [PubMed: 23092881]
8. Zhang H, Zhu JK. Active DNA Demethylation in Plants and Animals. *Cold Spring Harb Symp Quant Biol*. 2012; 77:161–73. [PubMed: 23197304]
9. Bird A. DNA Methylation Patterns and Epigenetic Memory. *Genes Dev*. 2002; 16:6–21. [PubMed: 11782440]
10. Deaton AM, Bird A. CpG Islands and the Regulation of Transcription. *Genes Dev*. 2011; 25:1010–22. [PubMed: 21576262]
11. Lindahl T. Instability and Decay of the Primary Structure of DNA. *Nature*. 1993; 362:709–15. [PubMed: 8469282]
12. Scharer OD. Chemistry and Biology of DNA Repair. *Angew Chem Int Ed Engl*. 2003; 42:2946–2974. [PubMed: 12851945]
13. Lindahl T. An N-Glycosidase from *Escherichia Coli* That Releases Free Uracil from DNA Containing Deaminated Cytosine Residues. *Proc Natl Acad Sci U S A*. 1974; 71:3649–53. [PubMed: 4610583]
14. David SS, O'Shea VL, Kundu S. Base-Excision Repair of Oxidative DNA Damage. *Nature*. 2007; 447:941–50. [PubMed: 17581577]
15. Wilson DM 3rd, Bohr VA. The Mechanics of Base Excision Repair, and Its Relationship to Aging and Disease. *DNA Repair (Amst)*. 2007; 6:544–59. [PubMed: 17112792]
16. Hegde ML, Hazra TK, Mitra S. Early Steps in the DNA Base Excision/Single-Strand Interruption Repair Pathway in Mammalian Cells. *Cell Res*. 2008; 18:27–47. [PubMed: 18166975]
17. Dianov GL, Hubscher U. Mammalian Base Excision Repair: The Forgotten Archangel. *Nucleic Acids Res*. 2013; 41:3483–90. [PubMed: 23408852]

18. Slupphaug G, Mol CD, Kavli B, Arvai AS, Krokan HE, Tainer JA. A Nucleotide-Flipping Mechanism from the Structure of Human Uracil-DNA Glycosylase Bound to DNA. *Nature*. 1996; 384:87–92. [PubMed: 8900285]
19. Friedman JI, Stivers JT. Detection of Damaged DNA Bases by DNA Glycosylase Enzymes. *Biochemistry*. 2010; 49:4957–4967. [PubMed: 20469926]
20. Brooks SC, Adhikary S, Robinson EH, Eichman BF. Recent Advances in the Structural Mechanisms of DNA Glycosylases. *Biochim Biophys Acta*. 2013; 1834:247–71. [PubMed: 23076011]
21. Mullins EA, Shi R, Parsons ZD, Yuen PK, David SS, Igarashi Y, Eichman BF. The DNA Glycosylase AlkD Uses a Non-Base-Flipping Mechanism to Excise Bulky Lesions. *Nature*. 2015; 527:254–8. [PubMed: 26524531]
22. 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]
23. Neddermann P, Jiricny J. The Purification of a Mismatch-Specific Thymine-DNA Glycosylase from Hela Cells. *J Biol Chem*. 1993; 268:21218–24. [PubMed: 8407958]
24. 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]
25. Bellacosa A, Cicchillitti L, Schepis F, Riccio A, Yeung AT, Matsumoto Y, Golemis 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; 96:3969–74. [PubMed: 10097147]
26. Bellacosa A, Drohat AC. Role of Base Excision Repair in Maintaining the Genetic and Epigenetic Integrity of CpG Sites. *DNA Repair (Amst)*. 2015; 32:33–42. [PubMed: 26021671]
27. Horst JP, Fritz HJ. Counteracting the Mutagenic Effect of Hydrolytic Deamination of DNA 5-Methylcytosine Residues at High Temperature: DNA Mismatch N-Glycosylase Mig. Mth of the Thermophilic Archaeon *Methanobacterium Thermoautotrophicum* Thf. *EMBO J*. 1996; 15:5459–5469. [PubMed: 8895589]
28. Begley TJ, Haas BJ, Morales JC, Kool ET, Cunningham RP. Kinetics and Binding of the Thymine-DNA Mismatch Glycosylase, Mig-Mth, with Mismatch-Containing DNA Substrates. *DNA Repair (Amst)*. 2003; 2:107–20. [PubMed: 12509271]
29. Iyer RR, Pluciennik A, Burdett V, Modrich PL. DNA Mismatch Repair: Functions and Mechanisms. *Chem Rev*. 2006; 106:302–23. [PubMed: 16464007]
30. Jiricny J. Postreplicative Mismatch Repair. *Cold Spring Harb Perspect Biol*. 2013; 5:a012633. [PubMed: 23545421]
31. Stivers JT, Jiang YL. A Mechanistic Perspective on the Chemistry of DNA Repair Glycosylases. *Chem Rev*. 2003; 103:2729–59. [PubMed: 12848584]
32. Drohat AC, Maiti A. Mechanisms for Enzymatic Cleavage of the N-Glycosidic Bond in DNA. *Org Biomol Chem*. 2014; 12:8367–8378. [PubMed: 25181003]
33. Prakash A, Doublet S, Wallace SS. The Fpg/Nei Family of DNA Glycosylases: Substrates, Structures, and Search for Damage. *Prog Mol Biol Transl Sci*. 2012; 110:71–91. [PubMed: 22749143]
34. 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]
35. Guo JU, Su Y, Shin JH, Shin J, Li H, Xie B, Zhong C, Hu S, Le T, Fan G, et al. Distribution, Recognition and Regulation of Non-CpG Methylation in the Adult Mammalian Brain. *Nat Neurosci*. 2014; 17:215–22. [PubMed: 24362762]
36. Kinde B, Gabel HW, Gilbert CS, Griffith EC, Greenberg ME. Reading the Unique DNA Methylation Landscape of the Brain: Non-CpG Methylation, Hydroxymethylation, and Mecp2. *Proc Natl Acad Sci U S A*. 2015; 112:6800–6. [PubMed: 25739960]

37. Cokus SJ, Feng S, Zhang X, Chen Z, Merriman B, Haudenschild CD, Pradhan S, Nelson SF, Pellegrini M, Jacobsen SE. Shotgun Bisulphite Sequencing of the Arabidopsis Genome Reveals DNA Methylation Patterning. *Nature*. 2008; 452:215–9. [PubMed: 18278030]
38. Bauer MJ, Fischer RL. Genome Demethylation and Imprinting in the Endosperm. *Curr Opin Plant Biol*. 2011; 14:162–7. [PubMed: 21435940]
39. Bhattacharya SK, Ramchandani S, Cervoni N, Szyf M. A Mammalian Protein with Specific Demethylase Activity for Mcpg DNA. *Nature*. 1999; 397:579–83. [PubMed: 10050851]
40. Gong Z, Morales-Ruiz T, Ariza RR, Roldan-Arjona T, David L, Zhu JK. Ros1, a Repressor of Transcriptional Gene Silencing in Arabidopsis, Encodes a DNA Glycosylase/Lyase. *Cell*. 2002; 111:803–14. [PubMed: 12526807]
41. Choi Y, Gehring M, Johnson L, Hannon M, Harada JJ, Goldberg RB, Jacobsen SE, Fischer RL. Demeter, a DNA Glycosylase Domain Protein, Is Required for Endosperm Gene Imprinting and Seed Viability in Arabidopsis. *Cell*. 2002; 110:33–42. [PubMed: 12150995]
42. Morales-Ruiz T, Ortega-Galisteo AP, Ponferrada-Marin MI, Martinez-Macias MI, Ariza RR, Roldan-Arjona T. Demeter and Repressor of Silencing 1 Encode 5-Methylcytosine DNA Glycosylases. *Proc Natl Acad Sci U S A*. 2006; 103:6853–8. [PubMed: 16624880]
43. He YF, Li BZ, Li Z, Liu P, Wang Y, Tang Q, Ding J, Jia Y, Chen Z, Li L, et al. Tet-Mediated Formation of 5-Carboxylcytosine and Its Excision by Tdg in Mammalian DNA. *Science*. 2011; 333:1303–1307. [PubMed: 21817016]
44. 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]
45. Gehring M, Huh JH, Hsieh TF, Penterman J, Choi Y, Harada JJ, Goldberg RB, Fischer RL. Demeter DNA Glycosylase Establishes Medea Polycomb Gene Self-Imprinting by Allele-Specific Demethylation. *Cell*. 2006; 124:495–506. [PubMed: 16469697]
46. Calarco JP, Borges F, Donoghue MT, Van Ex F, Jullien PE, Lopes T, Gardner R, Berger F, Feijo JA, Becker JD, et al. Reprogramming of DNA Methylation in Pollen Guides Epigenetic Inheritance Via Small Rna. *Cell*. 2012; 151:194–205. [PubMed: 23000270]
47. Ibarra CA, Feng X, Schoft VK, Hsieh TF, Uzawa R, Rodrigues JA, Zemach A, Chumak N, Machlicova A, Nishimura T, et al. Active DNA Demethylation in Plant Companion Cells Reinforces Transposon Methylation in Gametes. *Science*. 2012; 337:1360–4. [PubMed: 22984074]
48. Zilberman D, Gehring M, Tran RK, Ballinger T, Henikoff S. Genome-Wide Analysis of Arabidopsis Thaliana DNA Methylation Uncovers an Interdependence between Methylation and Transcription. *Nat Genet*. 2007; 39:61–9. [PubMed: 17128275]
49. Zhang X, Yazaki J, Sundaresan A, Cokus S, Chan SW, Chen H, Henderson IR, Shinn P, Pellegrini M, Jacobsen SE, et al. Genome-Wide High-Resolution Mapping and Functional Analysis of DNA Methylation in Arabidopsis. *Cell*. 2006; 126:1189–201. [PubMed: 16949657]
50. Penterman J, Uzawa R, Fischer RL. Genetic Interactions between DNA Demethylation and Methylation in Arabidopsis. *Plant Physiol*. 2007; 145:1549–57. [PubMed: 17951456]
51. Penterman J, Zilberman D, Huh JH, Ballinger T, Henikoff S, Fischer RL. DNA Demethylation in the Arabidopsis Genome. *Proc Natl Acad Sci U S A*. 2007; 104:6752–7. [PubMed: 17409185]
52. Choi Y, Harada JJ, Goldberg RB, Fischer RL. An Invariant Aspartic Acid in the DNA Glycosylase Domain of Demeter Is Necessary for Transcriptional Activation of the Imprinted Medea Gene. *Proc Natl Acad Sci U S A*. 2004; 101:7481–6. [PubMed: 15128940]
53. Ortega-Galisteo AP, Morales-Ruiz T, Ariza RR, Roldan-Arjona T. Arabidopsis Demeter-Like Proteins Dml2 and Dml3 Are Required for Appropriate Distribution of DNA Methylation Marks. *Plant Mol Biol*. 2008; 67:671–81. [PubMed: 18493721]
54. Agius F, Kapoor A, Zhu JK. Role of the Arabidopsis DNA Glycosylase/Lyase Ros1 in Active DNA Demethylation. *Proc Natl Acad Sci U S A*. 2006; 103:11796–801. [PubMed: 16864782]
55. Zharkov DO, Rosenquist TA, Gerchman SE, Grollman AP. Substrate Specificity and Reaction Mechanism of Murine 8-Oxoguanine-DNA Glycosylase. *J Biol Chem*. 2000; 275:28607–17. [PubMed: 10884383]

56. Hill JW, Hazra TK, Izumi T, Mitra S. Stimulation of Human 8-Oxoguanine-DNA Glycosylase by Ap-Endonuclease: Potential Coordination of the Initial Steps in Base Excision Repair. *Nucleic Acids Res.* 2001; 29:430–8. [PubMed: 11139613]
57. Vidal AE, Hickson ID, Boiteux S, Radicella JP. Mechanism of Stimulation of the DNA Glycosylase Activity of Hogg1 by the Major Human Ap Endonuclease: Bypass of the Ap Lyase Activity Step. *Nucleic Acids Res.* 2001; 29:1285–92. [PubMed: 11238994]
58. Morland I, Luna L, Gustad E, Seeberg E, Bjoras M. Product Inhibition and Magnesium Modulate the Dual Reaction Mode of Hogg1. *DNA Repair (Amst).* 2005; 4:381–7. [PubMed: 15661661]
59. Kuznetsov NA, Koval VV, Zharkov DO, Nevinsky GA, Douglas KT, Fedorova OS. Kinetics of Substrate Recognition and Cleavage by Human 8-Oxoguanine-DNA Glycosylase. *Nucleic Acids Res.* 2005; 33:3919–31. [PubMed: 16024742]
60. Dalhus B, Forsbring M, Helle IH, Vik ES, Forstrom RJ, Backe PH, Alseth I, Bjoras M. Separation-of-Function Mutants Unravel the Dual-Reaction Mode of Human 8-Oxoguanine DNA Glycosylase. *Structure.* 2011; 19:117–27. [PubMed: 21220122]
61. Lee J, Jang H, Shin H, Choi WL, Mok YG, Huh JH. Ap Endonucleases Process 5-Methylcytosine Excision Intermediates During Active DNA Demethylation in Arabidopsis. *Nucleic Acids Res.* 2014; 42:11408–18. [PubMed: 25228464]
62. Martinez-Macias MI, Qian W, Miki D, Pontes O, Liu Y, Tang K, Liu R, Morales-Ruiz T, Ariza RR, Roldan-Arjona T, et al. A DNA 3' Phosphatase Functions in Active DNA Demethylation in Arabidopsis. *Mol Cell.* 2012; 45:357–70. [PubMed: 22325353]
63. Ponferrada-Marin MI, Roldan-Arjona T, Ariza RR. Ros1 5-Methylcytosine DNA Glycosylase Is a Slow-Turnover Catalyst That Initiates DNA Demethylation in a Distributive Fashion. *Nucleic Acids Res.* 2009; 37:4264–74. [PubMed: 19443451]
64. Martinez-Macias MI, Cordoba-Canero D, Ariza RR, Roldan-Arjona T. The DNA Repair Protein Xrcc1 Functions in the Plant DNA Demethylation Pathway by Stimulating Cytosine Methylation (5-Mec) Excision, Gap Tailoring, and DNA Ligation. *J Biol Chem.* 2013; 288:5496–505. [PubMed: 23316050]
65. Mok YG, Uzawa R, Lee J, Weiner GM, Eichman BF, Fischer RL, Huh JH. Domain Structure of the Demeter 5-Methylcytosine DNA Glycosylase. *Proc Natl Acad Sci U S A.* 2010; 107:19225–30. [PubMed: 20974931]
66. Ponferrada-Marin MI, Parrilla-Doblas JT, Roldan-Arjona T, Ariza RR. A Discontinuous DNA Glycosylase Domain in a Family of Enzymes That Excise 5-Methylcytosine. *Nucleic Acids Res.* 2011; 39:1473–84. [PubMed: 21036872]
67. Brooks SC, Fischer RL, Huh JH, Eichman BF. 5-Methylcytosine Recognition by Arabidopsis Thaliana DNA Glycosylases Demeter and Dml3. *Biochemistry.* 2014; 53:2525–32. [PubMed: 24678721]
68. Boon EM, Livingston AL, Chmiel NH, David SS, Barton JK. DNA-Mediated Charge Transport for DNA Repair. *Proc Natl Acad Sci U S A.* 2003; 100:12543–7. [PubMed: 14559969]
69. Boal AK, Genereux JC, Sontz PA, Gralnick JA, Newman DK, Barton JK. Redox Signaling between DNA Repair Proteins for Efficient Lesion Detection. *Proc Natl Acad Sci U S A.* 2009; 106:15237–42. [PubMed: 19720997]
70. Sontz PA, Muren NB, Barton JK. DNA Charge Transport for Sensing and Signaling. *Acc Chem Res.* 2012; 45:1792–800. [PubMed: 22861008]
71. Sontz PA, Mui TP, Fuss JO, Tainer JA, Barton JK. DNA Charge Transport as a First Step in Coordinating the Detection of Lesions by Repair Proteins. *Proc Natl Acad Sci U S A.* 2012; 109:1856–61. [PubMed: 22308447]
72. Grodick MA, Segal HM, Zwang TJ, Barton JK. DNA-Mediated Signaling by Proteins with 4fe-4s Clusters Is Necessary for Genomic Integrity. *J Am Chem Soc.* 2014; 136:6470–8. [PubMed: 24738733]
73. Ponferrada-Marin MI, Martinez-Macias MI, Morales-Ruiz T, Roldan-Arjona T, Ariza RR. Methylation-Independent DNA Binding Modulates Specificity of Repressor of Silencing 1 (Ros1) and Facilitates Demethylation in Long Substrates. *J Biol Chem.* 2010; 285:23032–9. [PubMed: 20489198]

74. Berti PJ, McCann JA. Toward a Detailed Understanding of Base Excision Repair Enzymes: Transition State and Mechanistic Analyses of N-Glycoside Hydrolysis and N-Glycoside Transfer. *Chem Rev.* 2006; 106:506–55. [PubMed: 16464017]
75. Guthrie RD, Jencks WP. Iupac Recommendations for the Representation of Reaction-Mechanisms. *Accounts Chem Res.* 1989; 22:343–349.
76. Werner RM, Stivers JT. Kinetic Isotope Effect Studies of the Reaction Catalyzed by Uracil DNA Glycosylase: Evidence for an Oxocarbenium Ion-Uracil Anion Intermediate. *Biochemistry.* 2000; 39:14054–14064. [PubMed: 11087352]
77. McCann JAB, Berti PJ. Transition-State Analysis of the DNA Repair Enzyme MutY. *J Am Chem Soc.* 2008; 130:5789–5797. [PubMed: 18393424]
78. Chen XY, Berti PJ, Schramm VL. Transition-State Analysis for Depurination of DNA by Ricin a-Chain. *J Am Chem Soc.* 2000; 122:6527–6534.
79. Schwartz PA, Vetticatt MJ, Schramm VL. Transition State Analysis of Thymidine Hydrolysis by Human Thymidine Phosphorylase. *J Am Chem Soc.* 2010; 132:13425–33. [PubMed: 20804144]
80. McCann JAB, Berti PJ. Transition State Analysis of Acid-Catalyzed Damp Hydrolysis. *J Am Chem Soc.* 2007; 129:7055–7064. [PubMed: 17497857]
81. Dinner AR, Blackburn GM, Karplus M. Uracil-DNA Glycosylase Acts by Substrate Autocatalysis. *Nature.* 2001; 413:752–755. [PubMed: 11607036]
82. Shapiro R, Danzig M. Acidic Hydrolysis of Deoxycytidine and Deoxyuridine Derivatives. The General Mechanism of Deoxyribonucleoside Hydrolysis. *Biochemistry.* 1972; 11:23–29. [PubMed: 5009434]
83. La Francois CJ, Jang YH, Cagin T, Goddard WA, Sowers LC. Conformation and Proton Configuration of Pyrimidine Deoxynucleoside Oxidation Damage Products in Water. *Chem Res Toxicol.* 2000; 13:462–470. [PubMed: 10858319]
84. 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; 135:15813–22. [PubMed: 24063363]
85. Wempen I, Fox JJ. Pyrimidines. I. The Synthesis of 6-Fluorocytosine and Related Compounds. *J Med Chem.* 1963; 122:688–93.
86. 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; 15:3442–7. [PubMed: 8670846]
87. Kwon K, Jiang YL, Stivers JT. Rational Engineering of a DNA Glycosylase Specific for an Unnatural Cytosine:Pyrene Base Pair. *Chem Biol.* 2003; 10:351–359. [PubMed: 12725863]
88. 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]
89. Drohat AC, Jagadeesh J, Ferguson E, Stivers JT. Role of Electrophilic and General Base Catalysis in the Mechanism of Escherichia Coli Uracil DNA Glycosylase. *Biochemistry.* 1999; 38:11866–11875. [PubMed: 10508389]
90. Drohat AC, Stivers JT. Escherichia Coli Uracil DNA Glycosylase: Nmr Characterization of the Short Hydrogen Bond from His187 to Uracil O2. *Biochemistry.* 2000; 39:11865–75. [PubMed: 11009598]
91. Drohat AC, Stivers JT. Nmr Evidence for an Unusually Low N1 Pka for Uracil Bound to Uracil DNA Glycosylase: Implications for Catalysis. *J Am Chem Soc.* 2000; 122:1840–1841.
92. Jiang YL, Drohat AC, Ichikawa Y, Stivers JT. Probing the Limits of Electrostatic Catalysis by Uracil DNA Glycosylase Using Transition State Mimicry and Mutagenesis. *J Biol Chem.* 2002; 277:15385–92. [PubMed: 11859082]
93. Liu P, Theruvathu JA, Darwanto A, Lao VV, Pascal T, Goddard W 3rd, Sowers LC. Mechanisms of Base Selection by the Escherichia Coli Mismatched Uracil Glycosylase. *J Biol Chem.* 2008; 283:8829–36. [PubMed: 18208817]
94. 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; 284:15835–46. [PubMed: 19324873]

95. Zhachkina A, Lee JK. Uracil and Thymine Reactivity in the Gas Phase: The S(N)2 Reaction and Implications for Electron Delocalization in Leaving Groups. *J Am Chem Soc.* 2009; 131:18376–18385. [PubMed: 19928991]
96. Nash HM, Lu R, Lane WS, Verdine GL. The Critical Active-Site Amine of the Human 8-Oxoguanine DNA Glycosylase, Hogg1: Direct Identification, Ablation and Chemical Reconstitution. *Chem Biol.* 1997; 4:693–702. [PubMed: 9331411]
97. Ikeda S, Biswas T, Roy R, Izumi T, Boldogh I, Kurosky A, Sarker AH, Seki S, Mitra S. Purification and Characterization of Human Nth1, a Homolog of Escherichia Coli Endonuclease Iii. Direct Identification of Lys-212 as the Active Nucleophilic Residue. *J Biol Chem.* 1998; 273:21585–93. [PubMed: 9705289]
98. Lu R, Nash HM, Verdine GL. A Mammalian DNA Repair Enzyme That Excises Oxidatively Damaged Guanines Maps to a Locus Frequently Lost in Lung Cancer. *Curr Biol.* 1997; 7:397–407. [PubMed: 9197244]
99. Kellie JL, Wilson KA, Wetmore SD. An Oniom and Md Investigation of Possible Monofunctional Activity of Human 8-Oxoguanine-DNA Glycosylase (Hogg1). *J Phys Chem B.* 2015; 119:8013–23. [PubMed: 26018802]
100. 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; 106:18497–502. [PubMed: 19841264]
101. Kuo CF, McRee DE, Fisher CL, O’Handley SF, Cunningham RP, Tainer JA. Atomic Structure of the DNA Repair [4fe-4s] Enzyme Endonuclease Iii. *Science.* 1992; 258:434–40. [PubMed: 1411536]
102. Kuo CF, McRee DE, Cunningham RP, Tainer JA. Crystallization and Crystallographic Characterization of the Iron-Sulfur-Containing DNA-Repair Enzyme Endonuclease Iii from Escherichia Coli. *J Mol Biol.* 1992; 227:347–51. [PubMed: 1522598]
103. Bruner SD, Norman DP, Verdine GL. Structural Basis for Recognition and Repair of the Endogenous Mutagen 8-Oxoguanine in DNA. *Nature.* 2000; 403:859–66. [PubMed: 10706276]
104. 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; 5:1058–64. [PubMed: 9846876]
105. Fromme JC, Banerjee A, Huang SJ, Verdine GL. Structural Basis for Removal of Adenine Mispairs with 8-Oxoguanine by Muty Adenine DNA Glycosylase. *Nature.* 2004; 427:652–656. [PubMed: 14961129]
106. Norman DP, Chung SJ, Verdine GL. Structural and Biochemical Exploration of a Critical Amino Acid in Human 8-Oxoguanine Glycosylase. *Biochemistry.* 2003; 42:1564–72. [PubMed: 12578369]
107. Hansch C, Leo A, Unger SH, Kim KH, Nikaitani D, Lien EJ. “Aromatic” Substituent Constants for Structure-Activity Correlations. *J Med Chem.* 1973; 16:1207–16. [PubMed: 4747963]
108. Riggs AD. X Inactivation, Differentiation, and DNA Methylation. *Cytogenet Cell Genet.* 1975; 14:9–25. [PubMed: 1093816]
109. Razin A, Riggs AD. DNA Methylation and Gene Function. *Science.* 1980; 210:604–10. [PubMed: 6254144]
110. Bird AP. CpG-Rich Islands and the Function of DNA Methylation. *Nature.* 1986; 321:209–13. [PubMed: 2423876]
111. Jones PA, Takai D. The Role of DNA Methylation in Mammalian Epigenetics. *Science.* 2001; 293:1068–70. [PubMed: 11498573]
112. Jones PA, Laird PW. Cancer Epigenetics Comes of Age. *Nat Genet.* 1999; 21:163–7. [PubMed: 9988266]
113. Baylin SB, Jones PA. A Decade of Exploring the Cancer Epigenome - Biological and Translational Implications. *Nat Rev Cancer.* 2011; 11:726–34. [PubMed: 21941284]
114. Bestor TH. The DNA Methyltransferases of Mammals. *Hum Mol Genet.* 2000; 9:2395–402. [PubMed: 11005794]
115. Zhu B, Zheng Y, Hess D, Anglikier 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; 97:5135–9. [PubMed: 10779566]

116. 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; 28:4157–65. [PubMed: 11058112]
117. Cortellino S, Xu J, Sannai M, Moore R, Caretti E, Cigliano A, Le Coz M, Devarajan K, Wessels A, Soprano D, et al. Thymine DNA Glycosylase Is Essential for Active DNA Demethylation by Linked Deamination-Base Excision Repair. *Cell*. 2011; 146:67–79. [PubMed: 21722948]
118. Cortazar D, Kunz C, Selfridge J, Lettieri T, Saito Y, Macdougall E, Wirz A, Schuermann D, Jacobs AL, Siegrist F, et al. Embryonic Lethal Phenotype Reveals a Function of Tdg in Maintaining Epigenetic Stability. *Nature*. 2011; 470:419–423. [PubMed: 21278727]
119. Shen L, Song CX, He C, Zhang Y. Mechanism and Function of Oxidative Reversal of DNA and Rna Methylation. *Annu Rev Biochem*. 2014; 83:585–614. [PubMed: 24905787]
120. Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, Agarwal S, Iyer LM, Liu DR, Aravind L, et al. Conversion of 5-Methylcytosine to 5-Hydroxymethylcytosine in Mammalian DNA by Mll Partner Tet1. *Science*. 2009; 324:930–5. [PubMed: 19372391]
121. 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]
122. 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]
123. Ito S, D'Alessio AC, Taranova OV, Hong K, Sowers LC, Zhang Y. Role of Tet Proteins in 5mc to 5hmc Conversion, Es-Cell Self-Renewal and Inner Cell Mass Specification. *Nature*. 2010; 466:1129–33. [PubMed: 20639862]
124. Weber AR, Krawczyk C, Robertson AB, Kusnierczyk A, Vagbo CB, Schuermann D, Klungland A, Schar P. Biochemical Reconstitution of Tet1-Tdg-Ber-Dependent Active DNA Demethylation Reveals a Highly Coordinated Mechanism. *Nat Commun*. 2016; 7:10806. [PubMed: 26932196]
125. Morgan HD, Dean W, Coker HA, Reik W, Petersen-Mahrt SK. Activation-Induced Cytidine Deaminase Deaminates 5-Methylcytosine in DNA and Is Expressed in Pluripotent Tissues: Implications for Epigenetic Reprogramming. *J Biol Chem*. 2004; 279:52353–52360. [PubMed: 15448152]
126. Metivier R, Gallais R, Tiffocche C, Le Peron C, Jurkowska RZ, Carmouche RP, Ibberson D, Barath P, Demay F, Reid G, et al. Cyclical DNA Methylation of a Transcriptionally Active Promoter. *Nature*. 2008; 452:45–50. [PubMed: 18322525]
127. 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; 135:1201–12. [PubMed: 19109892]
128. Kim MS, Kondo T, Takada I, Youn MY, Yamamoto Y, Takahashi S, Matsumoto T, Fujiyama S, Shirode Y, Yamaoka I, et al. DNA Demethylation in Hormone-Induced Transcriptional Derepression. *Nature*. 2009; 461:1007–12. [PubMed: 19829383]
129. Popp C, Dean W, Feng S, Cokus SJ, Andrews S, Pellegrini M, Jacobsen SE, Reik W. Genome-Wide Erasure of DNA Methylation in Mouse Primordial Germ Cells Is Affected by Aid Deficiency. *Nature*. 2010; 463:1101–1105. [PubMed: 20098412]
130. Bhutani N, Brady JJ, Damian M, Sacco A, Corbel SY, Blau HM. Reprogramming Towards Pluripotency Requires Aid-Dependent DNA Demethylation. *Nature*. 2010; 463:1042–1047. [PubMed: 20027182]
131. Guo JU, Su Y, Zhong C, Ming GL, Song H. Hydroxylation of 5-Methylcytosine by Tet1 Promotes Active DNA Demethylation in the Adult Brain. *Cell*. 2011; 145:423–434. [PubMed: 21496894]
132. 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]

133. 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; 40:8276–8284. [PubMed: 22740654]
134. Hashimoto H, Hong S, Bhagwat AS, Zhang X, Cheng X. Excision of 5-Hydroxymethyluracil and 5-Carboxylcytosine by the Thymine DNA Glycosylase Domain: Its Structural Basis and Implications for Active DNA Demethylation. *Nucleic Acids Res.* 2012; 40:10203–14. [PubMed: 22962365]
135. Boorstein RJ, Cummings A Jr, Marenstein DR, Chan MK, Ma Y, Neubert TA, Brown SM, Teebor GW. Definitive Identification of Mammalian 5-Hydroxymethyluracil DNA N-Glycosylase Activity as Smug1. *J Biol Chem.* 2001; 276:41991–7. [PubMed: 11526119]
136. Matsubara M, Masaoka A, Tanaka T, Miyano T, Kato N, Terato H, Ohyama Y, Iwai S, Ide H. Mammalian 5-Formyluracil-DNA Glycosylase. 1. Identification and Characterization of a Novel Activity That Releases 5-Formyluracil from DNA. *Biochemistry.* 2003; 42:4993–5002. [PubMed: 12718542]
137. Liu P, Burdzy A, Sowers LC. Repair of the Mutagenic DNA Oxidation Product, 5-Formyluracil. *DNA Repair (Amst).* 2003; 2:199–210. [PubMed: 12531390]
138. Masaoka A, Matsubara M, Hasegawa R, Tanaka T, Kurisu S, Terato H, Ohyama Y, Karino N, Matsuda A, Ide H. Mammalian 5-Formyluracil-DNA Glycosylase. 2. Role of Smug1 Uracil-DNA Glycosylase in Repair of 5-Formyluracil and Other Oxidized and Deaminated Base Lesions. *Biochemistry.* 2003; 42:5003–12. [PubMed: 12718543]
139. Matsubara M, Tanaka T, Terato H, Ohmae E, Izumi S, Katayanagi K, Ide H. Mutational Analysis of the Damage-Recognition and Catalytic Mechanism of Human Smug1 DNA Glycosylase. *Nucleic Acids Res.* 2004; 32:5291–302. [PubMed: 15466595]
140. 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]
141. Valinluck V, Liu P, Kang JI Jr, Burdzy A, Sowers LC. 5-Halogenated Pyrimidine Lesions within a Cpg Sequence Context Mimic 5-Methylcytosine by Enhancing the Binding of the Methyl-Cpg-Binding Domain of Methyl-Cpg-Binding Protein 2 (Mecp2). *Nucl Acids Res.* 2005; 33:3057–3064. [PubMed: 15917437]
142. Miyabe I, Zhang QM, Kino K, Sugiyama H, Takao M, Yasui A, Yonei S. Identification of 5-Formyluracil DNA Glycosylase Activity of Human Hnht1 Protein. *Nucleic Acids Res.* 2002; 30:3443–8. [PubMed: 12140329]
143. Chelico L, Pham P, Goodman MF. Stochastic Properties of Processive Cytidine DNA Deaminases Aid and Apobec3g. *Philos Trans R Soc Lond B Biol Sci.* 2009; 364:583–93. [PubMed: 19022738]
144. 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]
145. 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]
146. Wijesinghe P, Bhagwat AS. Efficient Deamination of 5-Methylcytosines in DNA by Human Apobec3a, but Not by Aid or Apobec3g. *Nucleic Acids Res.* 2012; 40:9206–17. [PubMed: 22798497]
147. Carpenter MA, Li M, Rathore A, Lackey L, Law EK, Land AM, Leonard B, Shandilya SM, Bohn MF, Schiffer CA, et al. Methylcytosine and Normal Cytosine Deamination by the Foreign DNA Restriction Enzyme Apobec3a. *J Biol Chem.* 2012; 287:34801–8. [PubMed: 22896697]
148. Siriwardena SU, Guruge TA, Bhagwat AS. Characterization of the Catalytic Domain of Human Apobec3b and the Critical Structural Role for a Conserved Methionine. *J Mol Biol.* 2015; 427:3042–55. [PubMed: 26281709]
149. Conticello SG. The Aid/Apobec Family of Nucleic Acid Mutators. *Genome Biol.* 2008; 9:229.

150. Song CX, Szulwach KE, Dai Q, Fu Y, Mao SQ, Lin L, Street C, Li Y, Poidevin M, Wu H, et al. Genome-Wide Profiling of 5-Formylcytosine Reveals Its Roles in Epigenetic Priming. *Cell*. 2013; 153:678–91. [PubMed: 23602153]
151. 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]
152. 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
153. 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 Rep*. 2015
154. Zhang L, Chen W, Iyer LM, Hu J, Wang G, Fu Y, Yu M, Dai Q, Aravind L, He C. A Tet Homologue Protein from *Coprinopsis cinerea* (Cctet) That Biochemically Converts 5-Methylcytosine to 5-Hydroxymethylcytosine, 5-Formylcytosine, and 5-Carboxylcytosine. *J Am Chem Soc*. 2014; 136:4801–4. [PubMed: 24655109]
155. Crawford DJ, Liu MY, Nabel CS, Cao XJ, Garcia BA, Kohli RM. Tet2 Catalyzes Stepwise 5-Methylcytosine Oxidation by an Iterative and De Novo Mechanism. *J Am Chem Soc*. 2016; 138:730–3. [PubMed: 26734843]
156. Simmons JM, Muller TA, Hausinger RP. Fe(II)/Alpha-Ketoglutarate Hydroxylases Involved in Nucleobase, Nucleoside, Nucleotide, and Chromatin Metabolism. *Dalton Trans*. 2008:5132–42. [PubMed: 18813363]
157. Hu L, Lu J, Cheng J, Rao Q, Li Z, Hou H, Lou Z, Zhang L, Li W, Gong W, et al. Structural Insight into Substrate Preference for Tet-Mediated Oxidation. *Nature*. 2015; 527:118–22. [PubMed: 26524525]
158. Hashimoto H, Pais JE, Dai N, Correa IR Jr, Zhang X, Zheng Y, Cheng X. Structure of Naegleria Tet-Like Dioxygenase (Ngtet1) in Complexes with a Reaction Intermediate 5-Hydroxymethylcytosine DNA. *Nucleic Acids Res*. 2015; 43:10713–21. [PubMed: 26323320]
159. Hashimoto H, Pais JE, Zhang X, Saleh L, Fu ZQ, Dai N, Correa IR Jr, Zheng Y, Cheng X. Structure of a Naegleria Tet-Like Dioxygenase in Complex with 5-Methylcytosine DNA. *Nature*. 2014; 506:391–5. [PubMed: 24390346]
160. Hu L, Li Z, Cheng J, Rao Q, Gong W, Liu M, Shi YG, Zhu J, Wang P, Xu Y. Crystal Structure of Tet2-DNA Complex: Insight into Tet-Mediated 5mc Oxidation. *Cell*. 2013; 155:1545–55. [PubMed: 24315485]
161. Smiley JA, Kundracik M, Landfried DA, Barnes VR Sr, Axhemi AA. Genes of the Thymidine Salvage Pathway: Thymine-7-Hydroxylase from a *Rhodotorula glutinis* Cdna Library and Iso-Orotate Decarboxylase from *Neurospora crassa*. *Biochim Biophys Acta*. 2005; 1723:256–64. [PubMed: 15794921]
162. Schiesser S, Hackner B, Pfaffeneder T, Muller M, Hagemeyer C, Truss M, Carell T. Mechanism and Stem-Cell Activity of 5-Carboxycytosine Decarboxylation Determined by Isotope Tracing. *Angew Chem Int Ed Engl*. 2012; 51:6516–20. [PubMed: 22644704]
163. Xu S, Li W, Zhu J, Wang R, Li Z, Xu GL, Ding J. Crystal Structures of Isoorotate Decarboxylases Reveal a Novel Catalytic Mechanism of 5-Carboxyl-Uracil Decarboxylation and Shed Light on the Search for DNA Decarboxylase. *Cell Res*. 2013; 23:1296–309. [PubMed: 23917530]
164. Liutkeviciute Z, Kriukiene E, Licyte J, Rudyte M, Urbanaviciute G, Klimasauskas S. Direct Decarboxylation of 5-Carboxylcytosine by DNA C5-Methyltransferases. *J Am Chem Soc*. 2014; 136:5884–7. [PubMed: 24716540]
165. Schiesser S, Pfaffeneder T, Sadeghian K, Hackner B, Steigenberger B, Schroder AS, Steinbacher J, Kashiwazaki G, Hofner G, Wanner KT, et al. Deamination, Oxidation, and C-C Bond Cleavage Reactivity of 5-Hydroxymethylcytosine, 5-Formylcytosine, and 5-Carboxycytosine. *J Am Chem Soc*. 2013; 135:14593–9. [PubMed: 23980549]
166. Liutkeviciute Z, Lukinavicius G, Masevicius V, Daujotyte D, Klimasauskas S. Cytosine-5-Methyltransferases Add Aldehydes to DNA. *Nat Chem Biol*. 2009; 5:400–2. [PubMed: 19430486]

167. Chen CC, Wang KY, Shen CK. The Mammalian De Novo DNA Methyltransferases Dnmt3a and Dnmt3b Are Also DNA 5-Hydroxymethylcytosine Dehydroxymethylases. *J Biol Chem.* 2012; 287:33116–21. [PubMed: 22898819]
168. Xue JH, Xu GF, Gu TP, Chen GD, Han BB, Xu ZM, Bjoras M, Krokan HE, Xu GL, Du YR. Uracil-DNA Glycosylase Ung Promotes Tet-Mediated DNA Demethylation. *J Biol Chem.* 2016; 291:731–8. [PubMed: 26620559]
169. Muller 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]
170. Sabag O, Zamir A, Keshet I, Hecht M, Ludwig G, Tabib A, Moss J, Cedar H. Establishment of Methylation Patterns in Es Cells. *Nat Struct Mol Biol.* 2014; 21:110–2. [PubMed: 24336222]
171. Zhang RP, Shao JZ, Xiang LX. Gadd45a Protein Plays an Essential Role in Active DNA Demethylation During Terminal Osteogenic Differentiation of Adipose-Derived Mesenchymal Stem Cells. *J Biol Chem.* 2011; 286:41083–94. [PubMed: 21917922]
172. Jin SG, Guo C, Pfeifer GP. Gadd45a Does Not Promote DNA Demethylation. *PLoS Genet.* 2008; 4:e1000013. [PubMed: 18369439]
173. Kienhofer S, Musheev MU, Stapf U, Helm M, Schomacher L, Niehrs C, Schafer A. Gadd45a Physically and Functionally Interacts with Tet1. *Differentiation.* 2015; 90:59–68. [PubMed: 26546041]
174. Li Z, Gu TP, Weber AR, Shen JZ, Li BZ, Xie ZG, Yin R, Guo F, Liu X, Tang F, et al. Gadd45a Promotes DNA Demethylation through Tdg. *Nucleic Acids Res.* 2015; 43:3986–97. [PubMed: 25845601]
175. Schomacher L, Han D, Musheev MU, Arab K, Kienhofer S, von Seggern A, Niehrs C. Neil DNA Glycosylases Promote Substrate Turnover by Tdg During DNA Demethylation. *Nat Struct Mol Biol.* 2016; 23:116–24. [PubMed: 26751644]
176. McLaughlin D, Coey CT, Yang WC, Drohat AC, Matunis MJ. Characterizing Requirements for Sumo Modification and Binding on Base Excision Repair Activity of Thymine DNA Glycosylase in Vivo. *J Biol Chem.* 2016
177. Sassa A, Caglayan M, Dyrkheeva NS, Beard WA, Wilson SH. Base Excision Repair of Tandem Modifications in a Methylated Cpg Dinucleotide. *J Biol Chem.* 2014; 289:13996–4008. [PubMed: 24695738]
178. 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]
179. 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]
180. 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; 289:15810–9. [PubMed: 24753249]
181. 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; 21:1456–64. [PubMed: 11889051]
182. Steinacher R, Schar P. Functionality of Human Thymine DNA Glycosylase Requires Sumo-Regulated Changes in Protein Conformation. *Curr Biol.* 2005; 15:616–623. [PubMed: 15823533]
183. Baba D, Maiti N, Jee JG, Uchimura Y, Saitoh H, Sugawara 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]
184. 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]

185. 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]
186. 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]
187. Drohat AC, Xiao G, Tordova M, Jagadeesh J, Pankiewicz KW, Watanabe KA, Gilliland GL, Stivers JT. Heteronuclear Nmr and Crystallographic Studies of Wild-Type and H187q Escherichia Coli Uracil DNA Glycosylase: Electrophilic Catalysis of Uracil Expulsion by a Neutral Histidine 187. *Biochemistry*. 1999; 38:11876–86. [PubMed: 10508390]
188. 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]
189. 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; 66:7686–93. [PubMed: 16885370]
190. Sumino M, Ohkubo A, Taguchi H, Seio K, Sekine M. Synthesis and Properties of Oligodeoxynucleotides Containing 5-Carboxy-2'-Deoxycytidines. *Bioorg Med Chem Lett*. 2008; 18:274–7. [PubMed: 18023346]
191. Dai Q, Sanstead PJ, Peng CS, Han D, He C, Tokmakoff A. Weakened N3 Hydrogen Bonding by 5-Formylcytosine and 5-Carboxylcytosine Reduces Their Base-Pairing Stability. *ACS Chem Biol*. 2016; 11:470–7. [PubMed: 26641274]
192. Hansch C, Leo A, Taft RW. A Survey of Hammett Substituent Constants and Resonance and Field Parameters. *Chem Rev*. 1991; 91:165–195.
193. Malik SS, Coey CT, Varney KM, Pozharski E, Drohat AC. Thymine DNA Glycosylase Exhibits Negligible Affinity for Nucleobases That It Removes from DNA. *Nucleic Acids Res*. 2015
194. Hashimoto H, Zhang X, Cheng X. Selective Excision of 5-Carboxylcytosine by a Thymine DNA Glycosylase Mutant. *J Mol Biol*. 2013; 425:971–976. [PubMed: 23337108]
195. LaFrancois CJ, Fujimoto J, Sowers LC. Synthesis and Characterization of Isotopically Enriched Pyrimidine Deoxynucleoside Oxidation Damage Products. *Chem Res Toxicol*. 1998; 11:75–83. [PubMed: 9477229]
196. Raiber EA, Murat P, Chirgadze DY, Beraldi D, Luisi BF, Balasubramanian S. 5-Formylcytosine Alters the Structure of the DNA Double Helix. *Nat Struct Mol Biol*. 2015; 22:44–9. [PubMed: 25504322]
197. Szulik MW, Pallan PS, Nocek B, Voehler M, Banerjee S, Brooks S, Joachimiak A, Egli M, Eichman BF, Stone MP. Differential Stabilities and Sequence-Dependent Base Pair Opening Dynamics of Watson-Crick Base Pairs with 5-Hydroxymethylcytosine, 5-Formylcytosine, or 5-Carboxylcytosine. *Biochemistry*. 2015; 54:1294–305. [PubMed: 25632825]
198. Munzel M, Lischke U, Stathis D, Pfaffeneder T, Gnerlich FA, Deiml CA, Koch SC, Karaghiosoff K, Carell T. Improved Synthesis and Mutagenicity of Oligonucleotides Containing 5-Hydroxymethylcytosine, 5-Formylcytosine and 5-Carboxylcytosine. *Chemistry*. 2011; 17:13782–8. [PubMed: 22069110]
199. Ngo TT, Yoo J, Dai Q, Zhang Q, He C, Aksimentiev A, Ha T. Effects of Cytosine Modifications on DNA Flexibility and Nucleosome Mechanical Stability. *Nat Commun*. 2016; 7:10813. [PubMed: 26905257]

Biographies

Chris Coey has been a Ph.D. candidate at the University of Maryland, Baltimore since 2014, having joined the Drohat laboratory in 2013. His research interests focus on the roles of post-translational modifications on enzymes involved in DNA repair and demethylation.

Alex Drohat is an Associate Professor in the Department of Biochemistry and Molecular Biology at the University of Maryland School of Medicine. He obtained a Ph.D. in Biochemistry and Molecular Biology at the University of Maryland in 1997 and conducted postdoctoral training at Johns Hopkins University School of Medicine and at the Center for Advanced Research in Biotechnology (CARB; now the Institute for Bioscience and Biotechnology Research). His research group focuses on enzymes that act on DNA, including those involved in DNA repair and active DNA demethylation, and on regulation of these enzymes by SUMO binding and conjugation.

Figure 1.
Basic steps of the BER pathway. While there are several variations with regard to the detailed chemistry and the relevant enzymes, the overall process is conserved.

Figure 2.

DNA glycosylases use nucleotide flipping (or base flipping) to extrude nucleotides from the DNA helix into their active site, providing access to the scissile *N*-glycosyl bond. Shown is a crystal structure of thymine DNA glycosylase (TDG), with the target nucleotide (dUrd, red) flipped into its active site (PDB ID: 3UFJ). The dGua (blue) that had been base-paired with dUrd (G/U mismatch) remains stacked in the DNA helix.

Figure 3.

Two types of DNA glycosylases. Monofunctional DNA glycosylases catalyze hydrolysis of the *N*-glycosyl bond to produce an AP site. Bifunctional DNA glycosylases can cleave the *N*-glycosyl bond using an enzyme-derived amine nucleophile, generating covalent intermediate, and then cleave a phosphodiester bond in the DNA backbone via β -elimination. Some bifunctional enzymes also cleave a second phosphodiester bond via δ -elimination.

Figure 4.

Overview of enzymes and pathways that regulated the levels of cytosine methylation in DNA. Dnmts (DNA methyltransferases) convert C to 5mC. The 5mC mark is erased by BER alone in plants. In animals, 5mC must be enzymatically modified via deamination or oxidation, giving a 5mC derivative that is converted to C via BER.

Figure 5.

BER-mediated DNA demethylation in plants. The DME-ROS1 enzymes are bifunctional, and can catalyze β -elimination, and potentially δ -elimination, in addition to base excision. APE1L and ZDP are “cleanup” enzymes that generate the requisite 3'-OH and 5'-phosphate termini for DNA synthesis and ligation reactions. We note that the DME enzymes could also perform only the base excision reaction (monofunctional), giving an AP site that could be acted on by the AP endonuclease (not shown).

Figure 6. Structure of DME-ROS1 DNA glycosylases. (A) Primary structure of the four DME-type enzymes in *Arabidopsis thaliana*. The conserved domains are: Domain A (blue), Glycosylase (green), and Domain B (orange). (B) Homology model of DME includes a C-terminal portion of “Domain A” (blue) and the Glycosylase domain (green). Regions of the DME primary construct used for modeling are shown above the model. Some catalytic residues are shown. The DNA is from a structure of DNA-bound Nth (PDB ID: 1P59), which was used for modelling. The junction between Domain A and the Glycosylase domain is indicated (dotted circle). For additional details, see the original paper, REF 67. Adapted from REF 67. Copyright 2014 American Chemical Society.

Figure 7.

Two general mechanisms for hydrolysis of *N*-glycosyl bonds. A stepwise reaction features two transition states (TS), one for leaving-group departure and the second for nucleophile addition, while a concerted mechanism has a single TS.

Figure 8.

Acid catalysis of leaving-group departure and its role in *N*-glycoside hydrolysis of dCyd.

(A) The departing anionic base can be highly unstable; leaving-group quality can be greatly increased by protonation of the nucleobase, typically at a ring nitrogen, to give a neutral departing species. (B) *N*-glycoside hydrolysis of dCyd is acid catalyzed via N3 protonation.

Figure 9.

N-glycoside hydrolysis of dUrd. (A) The enzymatic and non-enzymatic reactions feature departure of a Ura anion. (B) Resonance stabilization of the Ura anion. (C) Ura protonation is highly unfavorable ($pK_a - 3.4$), thus acid catalysis of Ura departure is not observed at neutral pH.

Figure 10.

Potential stepwise mechanisms for excision of 5mC by DME-ROS1 enzymes. (A) Two potential stepwise mechanisms include the expulsion of 5mC as an anion (upper) or acid-catalyzed departure of a neutral leaving group through protonation of 5mC (at N3). In both models, the excised base deprotonates the Lys nucleophile. The C5 substituent (R) is CH₃ for 5mC and Br for 5BrC. (B) Stepwise mechanism involving a water molecule, rather than a Lys side chain, as the nucleophile. Acid catalyzed base expulsion is indicated, though departure of an anionic leaving group is also a possibility (not shown). For both pathways (Lys or water nucleophile), the potential for monofunctional activity, leading to an AP site that is processed by an AP endonuclease and then downstream BER enzymes, is indicated.

Figure 11.

Potential pathways for active 5mC demethylation. DNA methyltransferase (DNMT) enzymes convert C to 5mC. In the TET-TDG-BER pathway, TET enzymes oxidize 5mC to give 5hmC, 5fC, or 5caC, TDG removes 5fC or 5caC, and the resulting abasic sites are converted to cytosine via BER. Other possibilities include active deamination of 5mC or 5hmC to give T or 5hmU, which are excised by a DNA glycosylase and converted to C via BER. Finally, it is possible that 5caC could be directly converted to C by a decarboxylase, though no enzyme possessing this activity has been discovered to date. For clarity, passive DNA demethylation mechanisms are not shown.

Figure 12.
Potential mechanism for decarboxylation of 5caC in DNA.

Figure 13.

Brønsted-type linear free energy relationship (LFER) for non-enzymatic *N*-glycoside hydrolysis of C5-substituted dUrd nucleosides. The dependence of $\log k_{\text{obs}}$ on N1 acidity ($\text{p}K_{\text{a}}$) gives a slope of $\beta_{\text{g}} = -0.87 \pm 0.03$. Details of the LFER have been described.^{32,88} The inset shows a putative loose transition state for a concerted reaction, with negative charge developing on the departing base.⁸¹

Figure 14.

Brønsted-type LFER for TDG-catalyzed excision of 5-substituted uracils and cytosines from DNA. The dependence of $\log k_{\text{obs}}$ on N1 pK_a for the base gives $\beta_{\text{lg}} = \text{minus}; 1.6 \pm 0.2$; BrU, IU, and BrC ($\square$) were excluded from data fitting because $k_{\max}$ values suggested these bases have limited access to the TDG active site. Structures of Ura and Cyt show the location of C5 substituents. Adapted from REF 88. Copyright 2006 American Chemical Society.

Figure 15.

Resonance stabilization and N1 acidity for pyrimidines. (A) Resonance stabilization is more extensive for anion of 5fC relative to 5FC. (B) Calculated N1 acidities, reported as the free energy (ΔG , kcal mol⁻¹) of deprotonation in water, where a lower value indicates greater acidity. Adapted from REF 84. Copyright 2013 American Chemical Society.

Figure 16.

TDG uses acid catalysis for excision of 5caC but not 5fC. Fitting the 5caC activity to a simple model for ionization of a single essential protonated group (dotted line) gives a $pK_a = 5.80 \pm 0.03$, but the fitting is poor. Fitting to a model with one essential protonated group and a second, nonessential group that gives higher activity in its deprotonated state (solid line) gives $pK_a^1 = 5.75 \pm 0.03$ and $pK_a^2 = 8.2 \pm 0.7$, with a rate enhancement factor $(1 + \alpha)$ of 4.5. Adapted from REF 84. Copyright 2013 American Chemical Society.

Figure 17.

Structure of TDG (catalytic domain) with a d5caCyd analog (non-cleavable) flipped into the active site (PDBID: 3UOB).¹⁸⁸ Hydrogen bonds (dashed lines) and van der Waals contacts (dotted lines, $d \leq 3.7$ Å) are shown. Similar interactions are found for the N140A-TDG variant (catalytic domain) bound to an A-caC mismatch (PDBID: 3UO7).¹⁸⁸ Adapted from REF 84. Copyright 2013 American Chemical Society.

Figure 18.

Tautomeric forms of fC and caC and their guanine base pairs. (A) Imino and amino tautomers of fC and caC. (B) Shown is a canonical Watson-Crick type G:caC base pair involving the imino tautomer of caC, which has been observed in crystal structures; an the wobble G/caC pair predicted, but not yet observed, for Gua pairing with the amino tautomer of caC. Similarly, Watson-Crick type G-fC base pairs involving the fC imino tautomer have been observed in crystal structures but the wobble G/fC pair predicted for the amino fC tautomer has not been experimentally observed. For comparison, the experimentally established wobble structure for a G/T mismatched pair is also shown.