

The Molecular Basis of DNA Demethylation

Miao Shi and Li Shen

Abstract DNA methylation is a key epigenetic modification in mammalian genomes and is dynamically regulated in development and diseases. While enzymes catalyzing DNA methylation have been well characterized, those involved in demethylation have remained elusive until recently. Mounting evidence now suggests that the TET proteins, a family of AlkB-like Fe(II)/ α -ketoglutarate-dependent dioxygenases, initiate active DNA demethylation by oxidizing 5-methylcytosine (5mC) to generate 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC), and 5-carboxylcytosine (5caC). In this chapter, we discuss the molecular basis of DNA demethylation in mammalian genomes, focusing on TET proteins and TET-mediated oxidative DNA demethylation. Other potential DNA demethylation pathways are also summarized.

Keywords DNA methylation • DNA demethylation • TET proteins • Dioxygenase • 5-methylcytosine • 5-hydroxymethylcytosine • 5-formylcytosine • 5-carboxylcytosine • TDG

1 Introduction

The chromatin of a multicellular organism stores a vast quantity of information that defines the complex gene expression patterns in diverse cell types, and is indispensable for growth and development. This information is stored both genetically in DNA sequences and epigenetically through DNA and histone modifications [1–4]. However, nearly all cells in an organism (except gametes and some immune cells) contain the same genomic sequences as the zygotic genome and therefore, it is the epigenetic information residing in chromatin that determines a cell's identity and its corresponding gene expression profiles [5]. Generally, epigenetic information is faithfully propagated to each progeny cell upon division to maintain cell identity, but epigenetic states can also undergo dynamic changes during lineage specification

M. Shi • L. Shen (✉)

Life Sciences Institute, Zhejiang University, Hangzhou 310058, China

e-mail: li_shen@zju.edu.cn

or upon certain environmental stimuli [6]. Aberrant alterations of epigenetic information, such as DNA and histone modifications, are frequently associated with the onset of various human diseases, including cancers [7].

DNA methylation, in particular, is the most common covalent modification of DNA. The best-studied form of DNA methylation is 5-methylcytosine (5mC), which is generated by S-adenosyl-L-methionine (SAM)-dependent DNA methyltransferases (DNMTs), and does not interfere with Watson-Crick base pairing [8, 9]. In mammals, this enzymatically introduced DNA methylation exists predominantly in the CpG dinucleotide context (a cytosine followed by a guanine) and carries epigenetic information typically required for long-term gene silencing. Notably, over 70% of CpGs in somatic mammalian cells are methylated, therefore 5mC has long been the focus of compelling biochemical and genomics studies. In addition to 5mC, other forms of DNA methylation also exist. However, it is important to note that not all DNA methylation is the carrier of epigenetic information. For example, methylation can be introduced by endogenous or exogenous methylation agents-mediated DNA damage [9]. These methylated bases, such as N1-methyladenosine (m1A) and N3-methylcytosine (m3C), are considered cytotoxic or mutagenic as they tend to block or alter Watson-Crick base pairing. In the following sections, we will focus on the demethylation of 5mC, and will use the term DNA methylation to refer to 5mC only.

2 DNA Methylation Machinery

In mammals, three enzymatically active DNMTs, namely DNMT1, DNMT3A and DNMT3B, catalyze the transfer of a methyl group from SAM to the carbon-5 position of cytosine residues in DNA, generating 5mC [8]. DNMT1 preferentially methylates hemimethylated DNA [10], and in the presence of its cofactor UHRF1 (also known as NP95) [11, 12], DNMT1 is mainly responsible for copying DNA methylation patterns to the daughter strands during DNA replication (maintenance methylation). In contrast, DNMT3A and DNMT3B are the main enzymes to establish initial DNA methylation patterns during early embryonic development (*de novo* methylation), and do not show any preference for hemimethylated DNA [13]. Nevertheless, both maintenance and *de novo* methylation activities are required for normal development as depletion of DNMT1 or DNMT3B in mice results in embryonic lethality, and *Dnmt3a*-knockout mice die 4–8 weeks after birth [13, 14].

Structurally, the methyl group of 5mC is located in the major groove of DNA double helix, and is involved in either attracting or repelling many DNA binding proteins [15]. For example, three of the methyl-CpG binding domain (MBD) containing proteins, MeCP2, MBD1, MBD2, and a transcriptional regulator KAISO, have been shown to preferentially bind to methylated DNA and recruit repressor complexes to methylated promoters, leading to subsequent chromatin condensation and gene silencing [16]. On the contrary, DNA methylation can also prevent binding of some transcription factors (TFs), such as YY1 and CTCF [17, 18], to their specific

recognition sites. DNA methylation has been demonstrated to play critical roles in various of cellular processes such as genomic imprinting, X-chromosome inactivation, retrotransposon silencing as well as maintenance of cell identity, supporting its general transcription repression function and heritable nature [15].

3 Passive and Active DNA Demethylation

While most histone modifications are readily reversible [19], DNA methylation has been generally viewed as a relatively stable epigenetic mark. Indeed, there is a dedicated maintenance enzyme, DNMT1, to faithfully copy DNA methylation patterns to daughter strands during DNA replication; in addition, the methyl group on 5mC is connected to the base through a C-C bond which exhibits high chemical stability under physiological conditions; furthermore, no DNA demethylase could be identified by 2009 when a large number of histone demethylases had been discovered. Nevertheless, studies in the past decade have indicated that DNA methylation is not as static as once thought. Loss of DNA methylation, or DNA demethylation, has been reported in various biological contexts and can be achieved through either passive or active mechanisms.

As illustrated in Fig. 1, passive DNA demethylation, or replication-dependent dilution of 5mC, refers to loss of 5mC instead of semi conservatively replicating methylation patterns during DNA replication. In the absence of functional maintenance methylation machinery, i.e., DNMT1 and UHRF1, successive cycles of DNA replication can result in gradual dilution of 5mC to achieve global DNA demethylation. Passive DNA demethylation has been demonstrated to play a major role in maternal-genome demethylation of zygotes [20–22], and in the whole-genome demethylation of primordial germ cells (PGCs) [23–25].

By contrast, active DNA demethylation refers to direct removal of the methyl group from 5mC, or an enzymatic process that removes or modifies 5mC with regeneration of unmodified cytosine. Processes that are initiated with active modification (AM) of 5mC can be further divided into two forms by whether the modified 5mC is converted to unmodified cytosine through passive dilution (PD) or active restoration (AR). Similar to passive DNA demethylation, the AM-PD pathway may be well suited for large-scale DNA demethylation events observed in PGCs and zygotes, which we will discuss in Sect. 4.4. However, the AM-AR pathway and direct removal of the methyl group or the 5mC base may take place rapidly, and are implicated in locus-specific demethylation which requires rapid response towards environmental stimuli. For example, rapid active DNA demethylation was observed at the interleukin-2 (IL-2) promoter-enhancer region in activated T lymphocytes within 20 minutes upon stimulation [26], at the promoter of brain-derived neurotrophic factor (BDNF) in KCl-stimulated postmitotic neurons without DNA replication [27], at several other specific genomic loci in response to nuclear hormone and growth factors [28–30]. Thus, these studies suggest that active DNA demethylation could function in the dynamic regulation of genes that require rapid responses to specific environmental stimuli.

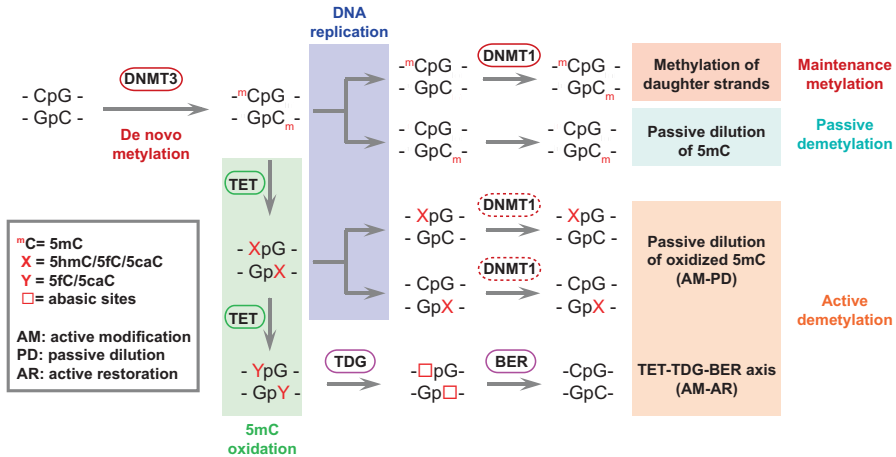


Fig. 1 Major mechanisms of passive and active DNA demethylation in the mammalian genome. DNA methylation patterns are established by DNMT3 proteins and maintained by DNMT1 during DNA replication, and passive DNA demethylation occurs when DNMT1 is inhibited. TET proteins can oxidize 5-methylcytosine (5mC) to generate 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC), and 5-carboxylcytosine (5caC), which are inefficient substrates for DNMT1 and are passively diluted during DNA replication. This form of active DNA demethylation is termed as active modification followed by passive dilution (AM-PD). Among the three 5mC oxidation derivatives, 5fC and 5caC can be excised by TDG to form abasic sites, which are further repaired by base excision repair (BER) pathway to complete DNA demethylation. This form of active DNA demethylation is termed as active modification followed by active restoration (AM-AR)

4 TET-Mediated Oxidative DNA Demethylation

4.1 TET Family Dioxygenases

While passive DNA demethylation has long been understood and accepted, the mechanism of active DNA demethylation was not understood until recently, following the discovery that TET (ten-eleven translocation) proteins can convert 5mC to its oxidized forms, namely 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC), and 5-carboxylcytosine (5caC). Interestingly, the names of TET genes trace back to the involvement of the human *TET1* gene in the ten-eleven translocation [t(10;11)(q22;q23)] in rare cases of acute myeloid leukemia (AML), which fuses the *TET1* gene on chromosome 10 with the mixed-lineage leukemia gene (*MLL*; also known as *KMT2A*) on chromosome 11 [31, 32]. Along with *TET1*, two additional genes in this protein family, *TET2* and *TET3* are also identified based on their sequence homology. Sequence comparison and structural studies have shown that TET proteins are a distinct family of the Fe(II)/ α -ketoglutarate (α KG)-dependent dioxygenase superfamily [33, 34], members of which also include JmjC domain containing histone demethylases and AlkB family DNA/RNA repair proteins.

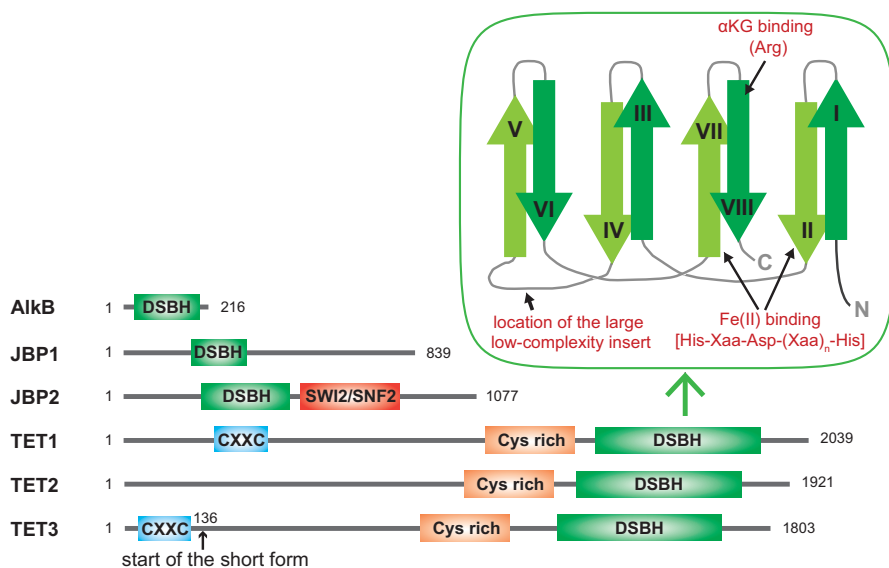


Fig. 2 Schematic diagrams of the TET proteins. Three conserved domains are indicated in mouse TET proteins, including CXXC zinc finger, cysteine-rich region (Cys-rich), and the double-stranded β -helix (DSBH) fold of the Fe(II)/ α -ketoglutarate (α KG)-dependent dioxygenases. Locations of Fe(II) and the α KG-binding sites in the conserved DSBH fold are shown in the topology diagram. Also presented are the domain structures of three other enzymes in the super family: *Trypanosoma brucei* JBP1, JBP2, and *Escherichia coli* AlkB. Note that TET3 has a shorter form, which starts at amino acid 136, that does not contain the CXXC domain

Similar to most Fe(II)/ α KG-dependent dioxygenase superfamily members, TET proteins share a conserved DSBH fold (or jelly-roll fold) in their catalytic domains, consisting of eight antiparallel β -strands (I–VIII) and an iron-binding motif (Fig. 2). Unique characteristics are also present in the catalytic domain of TET proteins, such as a cysteine-rich domain adjacent to the N terminus of the DSBH fold, and a large non-conserved low-complexity region between conserved β -strands IV and V [33, 35, 36]. Although the low-complexity region's function is not clear, the cysteine-rich domain has been shown to stabilize substrate DNA by wrapping around the DSBH core, and is essential for the enzymatic activity [37]. Outside of the catalytic domain, a CXXC (Cysteine-X-X-cysteine) domain is present at the N terminus of both TET1 and TET3. Indeed, TET1 is also known as CXXC6 (CXXC zinc finger 6) and LCX (leukemia-associated CXXC protein). Both in vitro DNA binding assay and structural analyses have revealed that TET proteins' CXXC domain strongly binds to unmethylated DNA [38]. However, the catalytic domains of TET proteins alone can also bind to DNA and oxidize 5mC without the help of a CXXC domain [39]. Therefore, the catalytic domains of TET proteins may possess a non-sequence-specific DNA-binding capacity whereas the CXXC domain may increase the sequence selectivity to facilitate and regulate binding of TET proteins to their genomic targets [38, 40, 41]. Surprisingly, TET2 does not possess a CXXC

domain, which was suggested to be lost during evolution, and is now encoded by a separate, neighboring gene IDAX (also known as CXXC4)[41]. It is worth noting that TET3 has three isoforms, among which only the full-length form contains the CXXC domain [42]. The full-length TET3 was reported to bind to 5caC at CCG sequences through its CXXC domain, and promote DNA demethylation by acting as a regulator of 5caC removal by base excision repair [42].

4.2 TET-Mediated Iterative Oxidation of 5mC

The finding that TET proteins can convert 5mC of DNA to 5hmC by oxidation was a great advance in understanding the mechanisms of DNA demethylation [34]. This finding was initially inspired by the biosynthesis of glucosylated 5-hydroxymethyluracil (base J) in the genome of *Trypanosoma brucei*, a parasite causing African sleeping sickness. Base J synthesis involves oxidation of thymine to 5-hydroxymethyluracil (5hmU) by J-binding proteins 1 and 2 (JBP1 and JBP2), two enzymes of the Fe(II)/ α KG-dependent dioxygenase superfamily [43]. Because of the structural similarity between 5mC and thymine, mammalian homologs of JBP proteins were thought to possess 5mC oxidation activity, and TET family proteins were identified as the mammalian homologs of JBP proteins [33, 34]. Interestingly, the presence of TET genes in animals seems to coincide with the presence of 5mC in the genome [33, 36]. It was then convincingly demonstrated by in vitro biochemical experiments that TET proteins can oxidize 5mC to 5hmC [34]. Moreover, 5hmC is relatively abundant in mouse embryonic stem cells (ESCs) where both TET1 and TET2 are highly expressed, and its presence is TET dependent [34, 39], providing in vivo evidence that 5hmC is generated by TET-mediated oxidation of 5mC.

Fe(II)/ α KG-dependent dioxygenase-mediated oxidation reactions typically consist of two stages: dioxygen activation and substrate oxidation (Fig. 3). The dioxygen activation stage is a four-electron process, where Fe(II) and α KG may each contribute two electrons to activate a dioxygen molecule first into bridged peroxo and then into the Fe(IV)-oxo intermediate. In the following substrate oxidation stage, the inert C-H bond of the substrate can be oxidized by the highly active Fe(IV)-oxo species, and finally Fe(IV) is reduced back into Fe(II) to complete the catalytic cycle [44]. During the whole process, four electrons, two from α KG and another two from the substrate C-H bond, are consumed to fully reduce a dioxygen molecule. The two oxygen atoms of the dioxygen molecule are incorporated into the succinate (the oxidized and decarboxylated product of α KG) and the oxidized product (Fig. 3).

Interestingly, some Fe(II)/ α KG-dependent dioxygenase are capable to iteratively oxidize the substrate methyl group to carboxyl group. For instance, the thymine-7-hydroxylase, a Fe(II)/ α KG-dependent dioxygenase in the thymidine salvage pathway, is known to catalyze a three-step oxidation of thymine to generate 5-carboxyluracil (isoorotate), where the methyl group of thymine is sequentially oxidized to hydroxymethyl group, formyl group, and finally to the carboxyl group,

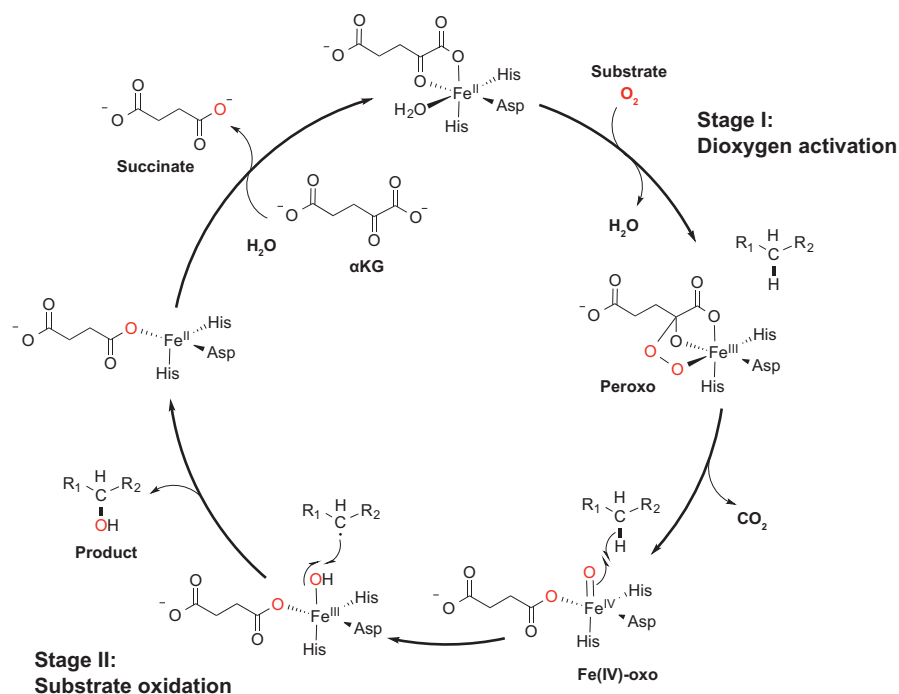


Fig. 3 The chemical mechanism underlying TET-mediated 5mC oxidation. In the dioxygen activation stage, Fe(II) and α KG activate dioxygen to form a highly active Fe(IV)-oxo species; and in the substrate oxidation stage, Fe(IV)-oxo inserts the oxygen atom into the C-H bond of the substrate, and Fe(IV) is reduced back to Fe(II) to complete the catalytic cycle

which is subsequently removed by an isoorotate decarboxylase (IDase) to generate uracil [45]. Enlightened by this example, it was proposed that TET proteins might oxidize 5mC not only to 5hmC, but also to 5fC and 5caC [46]. This hypothesis was soon experimentally proved both in vitro and in vivo [47, 48], and was further supported by a structural study that TET2's active cavity could recognize CpG dinucleotide regardless of its methylation/oxidation status [37].

4.3 TDG-Mediated Excision of 5fC/5caC

Because 5mC can be converted to 5hmC, 5fC and 5caC, these modified bases are naturally considered to be involved in DNA demethylation. However, unlike the N-methyl group in m1A, m3C and methylated histones, which is unstable on the C-N bond and go through spontaneous hydrolytic deformylation upon enzymatic oxidation (i.e., direct removal of the oxidized methyl group) [19], the methyl group of 5mC is connected through a highly stable C-C bond to the rest of the

base, and therefore the oxidized 5-substituents remain stable under physiological conditions [9]. Interestingly, although the 5-substituents seem not to be directly removed from 5mC oxidation derivatives, emerging evidence suggests that once converted to 5fC and 5caC, the modified cytosine base can be entirely removed from DNA by thymine-DNA glycosylase (TDG) [48, 49]. DNA demethylation is then completed by replacing the resulting abasic site with unmodified cytosine through the base excision repair (BER) pathway, similar to the active DNA demethylation mechanism in plants [50]. This TET-TDG-BER-mediated DNA demethylation process may take place rapidly, and seems a perfect candidate for locus-specific demethylation which requires rapid response towards environmental stimuli.

TDG belongs to the uracil-DNA glycosylase (UDG) superfamily. It has been well established that TDG can excise pyrimidine moiety from G/U and G/T mispairs in dsDNA by a base-flipping mechanism [51]. Interestingly, TDG also excises properly base-paired cytosine bases with 5-position substituents that destabilize the base-sugar bond (*N*-glycosidic bond), such as 5-fluorocytosine, indicating that the stability of the *N*-glycosidic bond contributes to TDG's substrate specificity [52]. More recent studies further demonstrated that TDG can recognize and remove 5fC and 5caC, but not 5mC, 5hmC, and unmodified cytosine, from DNA duplex when paired with guanine [48, 49]. Indeed, computational analyses suggest that 5fC and 5caC form a more labile *N*-glycosidic bond compared to unmodified cytosine, 5mC, 5hmC, and even 5-fluorocytosine [53]. Consistently, TDG has a slightly higher binding affinity towards G/5fC and G/5caC pairs than to G/U and G/T mismatches [49]. When co-overexpressed in HEK293 cells [48, 54], TDG efficiently depletes TET-generated 5fC and 5caC; and in contrast, TDG knockdown in mouse ESCs results in a 5–10-fold increase of endogenous 5fC and 5caC [55, 56], providing *in vivo* evidence that TDG is responsible for 5fC/5caC removal.

Intriguingly, among the four enzymes with UDG activity in mammals (i.e., TDG, UNG, MBD4, and SMUG1), only TDG is required during mouse embryonic development [57–60], where global DNA methylation reprogramming takes place, implicating that the DNA glycosylase activity of TDG is essential for DNA demethylation. Compared with the other UDGs, the active site of TDG is indeed uniquely configured to accommodate 5fC/5caC and facilitate its cleavage, as revealed by the crystal structure of human TDG in complex with 5caC-containing dsDNA [61]. Consistently, both *Tdg*-null-mutant and *Tdg*-catalytic-mutant mice exhibit abnormal DNA methylation and die around embryonic day (E)12.5 [57, 58], confirming a crucial role of the TET-TDG-BER axis in DNA demethylation during embryonic development.

4.4 Replication-Dependent Dilution of 5mC Oxidation Derivatives

TET mediated 5mC oxidation not only initiates the TET-TDG-BER demethylation pathway, but also generates DNA demethylation intermediates (i.e., 5hmC, 5fC, and 5caC) that can be passively diluted in a replication-dependent manner. Mechanistically, hemi-modified CpGs carrying 5hmC, 5fC or 5caC (XG:GC, where X = 5hmC/5fC/5caC) have been demonstrated to be significantly less efficient in being methylated by DNMT1 compared with hemimethylated CpGs (i.e., 5mCG:GC) [62–64], therefore, TET-mediated 5mC oxidation can block the maintenance methylation machinery, facilitating replication-dependent DNA demethylation. Because this DNA demethylation process starts with active modification of 5mC, it has been suggested to be regarded as active DNA demethylation (Fig. 1) [65]. The replication-dependent active DNA demethylation has been observed in the paternal genome (to a less extent in the maternal genome) of zygotes and in developing PGCs [21, 22, 66, 67]. However, it is worth noting that global DNA demethylation in zygotes could be largely achieved without 5mC oxidation due to the inhibition of DNMT1 at this stage. Therefore, 5mC oxidation probably only facilitates, but is not indispensable for, replication-dependent whole genome DNA demethylation. The extent to which 5mC oxidation is required for demethylation may depend on the genomic context of the DNA sequence.

4.5 Other Potential TET-Initiated Active DNA Demethylation Pathways

In addition to the two major TET-mediated DNA demethylation pathways discussed above (i.e., the TET-TDG-BER axis and the replication-dependent dilution of 5mC oxidation derivatives) which have been extensively supported by recent biochemical and genetic studies [65], evidence for the existence of other TET-initiated DNA demethylation pathways, in which 5mC oxidation derivatives act as demethylation intermediates, have also been reported.

Firstly, the 5-carboxyl group on 5caC might be removed by a putative decarboxylase to complete DNA demethylation. This mechanism was proposed under the inspiration of the thymidine salvage pathway that we discussed above [46], where the thymine-7-hydroxylase oxidizes the methyl group of thymine to a carboxyl group that is subsequently removed by an isoorotate decarboxylase to convert thymine to uracil [45]. Although the idea of decarboxylation is more energy-efficient compared with the TET-TDG-BER pathway, only one study reported weak 5caC decarboxylase activity in mouse ESC extracts [68]. In addition, the pronounced increase of endogenous 5caC upon TDG depletion has already indicated that TDG is the major enzyme for 5caC removal [55, 56]. Therefore, whether a 5caC decarboxylase exists remains to be explored.

Secondly, the 5-position substituents may be directly removed by DNMTs. It has been reported that both bacteria and mammalian DNMTs could remove the 5-hydroxymethyl group of 5hmC and the 5-carboxyl group of 5caC *in vitro* to generate unmodified cytosine in the absence of SAM, the methyl donor in a DNMT-mediated DNA methylation reaction [69–71]. However, given that SAM is present in all cell types as a general methyl donor of many other essential biochemical reactions, whether DNMT-mediated 5-position substituent removal of 5hmC/5caC can take place *in vivo* remains questionable.

Thirdly, 5hmC deamination followed by BER has also been implicated in active DNA demethylation. AID (activation-induced deaminase)/APOBEC (apolipoprotein B mRNA editing enzyme, catalytic polypeptide) family of cytidine deaminases typically target unmodified cytosine in single-strand DNA or RNA to generate mutations, which are required for the generation of antibody diversity in B cells, RNA editing, and retroviral defense [72]. Interestingly, one study showed that AID/APOBEC deaminases might deaminate 5hmC to produce 5hmU in HEK293 cells and in the mouse brain [73], indicating a TET-AID/APOBEC-BER axis for active DNA demethylation. But this potential pathway has been questioned due to the following reasons: 1) AID only acts robustly on single-strand DNA but not on double-strand DNA [74]; 2) AID/APOBEC deaminases exhibit no detectable *in vitro* deamination activity on 5hmC [54, 75]. Therefore, further evidence is required to support this potential active DNA demethylation mechanism.

5 Potential TET-Independent Active DNA Demethylation Mechanisms

While TET proteins have been widely accepted as major players for active DNA demethylation, many other proteins were also historically proposed to play direct roles in demethylating DNA [46, 76]. Here we list some of the putative mechanisms that are independent of TET proteins, however, due to lack of direct evidence or conflicting observations, these demethylation mechanisms must be reexamined to confirm their biological relevance.

Firstly, despite the difficult nature of breaking a C-C bond, enzymatic removal of the 5-methyl group from 5mC is the simplest way to achieve DNA demethylation. The first protein reported to possess this activity is the methylated DNA binding protein MBD2. It was shown that MBD2-mediated 5-methyl group excision could take place *in vitro* without any cofactors [77]. However, this observation could not be reproduced by other laboratories and MBD-null mice were viable with normal DNA methylation patterns [78], raising the question whether MBD2 could serve as a functional DNA demethylase *in vivo*. In addition to MBD2, elongator complex protein3 (ELP3) was also proposed to achieve DNA demethylation by breaking the C-C bond through a radical SAM mechanism [46]. While ELP3 bears a Fe-S radical SAM domain, and was reported to play a role in the paternal genome demethylation

in mouse zygotes [79], direct biochemical evidence demonstrating its enzymatic activity is still lacking. Interestingly, an *in vitro* study has showed that, in the absence of SAM, the mammalian DNMTs (i.e., DNMT3A, DNMT3B, and DNMT1) themselves could also remove the 5-methyl group from 5mC [80], but the physiological relevance of this observation remains unclear due to the widespread presence of SAM in all cell types as discussed above.

Secondly, the entire 5mC base can be erased by a DNA glycosylase to form an abasic site, followed by BER DNA repair pathway to complete active DNA demethylation. In plants, compelling biochemical and genetic evidence has validated this mechanism with the discovery of a family of specialized DNA glycosylases responsible for 5mC excision, namely Demeter (Dme) family proteins [81]. While no obvious mammalian orthologues of Dme family proteins have been identified, two mammalian DNA glycosylases, TDG and MBD4, were reported to have incision activity against 5mC [82, 83]. However, 5mC incision activity of the two enzymes is about 30 times lower compared with that against G/T mismatches. In addition, *Mbd4*-null mice were viable and exhibit normal DNA methylation patterns [59]. Although *Tdg*-deficient mice exhibit abnormal DNA methylation and die around E12.5 [57, 58], the phenotype is more likely to be attributed to the loss of 5fC/5caC incision activity of TDG required in the TET/TDG-mediated DNA demethylation as discussed in Sect. 4.3. Thus, whether BER of 5mC by a DNA glycosylase can contribute to DNA demethylation in mammals has yet to be determined.

Thirdly, active DNA demethylation may also be achieved through deamination of 5mC to generate thymine, followed by BER to replace this mismatched thymine to unmodified cytosine. As discussed earlier, AID/APOBEC family proteins show no detectable *in vitro* deamination activity on 5hmC, however, these deaminases do deaminate 5mC in the context of single-strand DNA *in vitro*, despite at a 10-fold slower rate compared with that towards their canonical substrate cytosine [54, 75, 84]. Indeed, several lines of evidence has suggested that AID/APOBEC family deaminases play a role in active DNA demethylation, including studies in zebrafish embryos [85], in mouse PGCs [86], in promoting pluripotency in somatic nuclei after fusion with ESCs [87], and in reprogramming of somatic cells to induced pluripotent stem cells (iPSCs) [88]. Nevertheless, due to conflicting observations and the fact that these deaminases only act robustly on single-strand DNA, further mechanistic studies are required to clarify the function of AID/APOBEC proteins in active DNA demethylation. Interestingly, DNMTs, in addition to AID/APOBEC proteins, have also been shown to deaminate 5mC in the absence of SAM *in vitro* [29]. But again, the physiological relevance of DNMTs' *in vitro* deamination activity is uncertain as depletion of SAM is unlikely in living cells.

Fourthly, nucleotide excision repair (NER), which typically repairs bulky DNA lesions generated by exposure to chemicals and radiation, has also been implicated in active DNA demethylation. Multiple lines of evidence have shown that GADD45 (growth arrest and DNA-damage-inducible 45) family proteins could stimulate active DNA demethylation via NER in frog, zebrafish, and mammals [85, 89–91]. However, evidence to the contrary also exists [92, 93]. More importantly, the exact underlying mechanism is still unclear. Therefore, the role of GADD45 family proteins and NER in DNA demethylation remains to be elucidated.

6 Regulation of TET-Mediated DNA Demethylation

Precise control of DNA methylation is critical to the maintenance of genome stability as well as cell-type- and developmental-stage-specific gene expression. Therefore, proper regulation of both DNA methylation and demethylation is required in many biological processes, such as development and the onset of diseases. Compared with the regulation of DNA methylation, which has been extensively studied [8, 15], the regulation of DNA demethylation has just begun to be understood. Because TET-mediated 5mC oxidation has become the most accepted mechanism of active DNA demethylation [9, 94], we only discuss factors involved in the regulation of TET-mediated DNA demethylation in the following sections.

6.1 Regulation of TET Expression

Regulation of enzyme abundance is a common way to control its activity in cells. In mouse E10.5 PGCs, transient conversion of 5mC to 5hmC can be readily detected together with a dramatic TET1 upregulation [67, 95, 96], indicating that regulation of TET expression is an important way to regulate DNA demethylation. Indeed, the three TET genes exhibit different expression patterns in a cell-type- and developmental-stage-specific manner: TET1 shows a high-level of expression specifically in mouse E10.5–12.5 PGCs, the inner cell mass (ICM) of blastocysts, as well as ESCs [39, 95]; TET2 is highly expressed in ESCs, and is broadly expressed in various mouse adult tissues [39]; TET3 is the only TET family member that is highly expressed in mouse oocytes and zygotes [97, 98], although it also shows a broad expression pattern in mouse adult tissues.

The regulation of TET expression has been reported at different levels. At the transcriptional level, cell-type-specific transcription factors (TFs) may play a major role. For example, a large cluster of binding sites for core pluripotency TFs are present in the upstream promoter of mouse *Tet1* gene [99], in line with the rapid reduction of TET1 expression upon ESC differentiation [47, 100]. At the posttranscriptional level, it has been shown that an oncogenic microRNA miR-22 could negatively regulate TET family proteins in breast cancer development and in hematopoietic stem cell transformation [101, 102]. In addition, one study reported that the CXXC domain-containing protein IDAX could directly interact with the catalytic domain of TET2 to downregulate TET2 protein through caspase-mediated degradation [41]. Moreover, all three TET proteins are direct substrates of calpains, a family of calcium-dependent proteases. Specifically, calpain1 mediates TET1 and TET2 turnover in mouse ESCs, and calpain2 regulates TET3 level during differentiation [103]. Such multiple layers of regulation on TET expression provides a robust control of TET activity in cells.

6.2 Regulation of TET Activity by Metabolites and Nutrients

Given the importance of precise regulation of DNA methylation in various biological processes, it is not surprising that TET activity can be regulated in multiple ways, including the metabolic states and the milieu of cells (e.g., nutritional and developmental signals, stress, and chemical exposure). For example, both adenosine-5'-triphosphate (ATP) and hydroquinone were reported to stimulate TET-mediated 5mC oxidation [48, 104]. More importantly, the five-carbon dicarboxylic acid α KG, which is part of the tricarboxylic acid (TCA) cycle, is an essential co-factor for TET-mediated 5mC oxidation (as discussed in Sect. 4.2.). Therefore, metabolic states of cells, which affect intracellular α KG levels, may influence TET activity. Indeed, it has been reported that global 5hmC levels were rapidly increased together with α KG levels in mouse livers within 30 min after glucose, glutamine or glutamate injection [105].

In contrast, another five-carbon dicarboxylic acid, 2-hydroxyglutarate (2HG), which is chemically analogous to α KG, has been shown to inhibit TET activity by competing with α KG [106, 107]. Cellular accumulation of 2HG is often caused by tumor-associated mutations in the NADP⁺-dependent isocitrate dehydrogenase genes (*IDH1/IDH2*), which encode enzymes that normally produce α KG in the cell. These tumor-associated *IDH1/IDH2* mutations (R132 of *IDH1* and R140/R172 of *IDH2*) impair α KG production, and obtain an enzymatic activity to convert α KG to 2HG [108, 109], which inhibits TET activity. Consistently, co-expression of mutant *IDH* enzymes and TET proteins inhibits TET-mediated 5mC to 5hmC conversion [106, 107]. It was hypothesized that the substitution of the keto group on α KG to a hydroxyl group on 2HG might interfere with Fe(II) binding and stabilize the reaction intermediate. In line with this hypothesis, another two metabolites, fumarate and succinate, which also share structural similarity with α KG, both function as competitive inhibitors of Fe(II)/ α KG-dependent dioxygenases, including TET proteins. Similar to 2HG, these metabolites are accumulated in a subset of human cancers with inactivation mutations of fumarate hydratase (FH) and succinate dehydrogenase (SDH), respectively [110]. Thus, multiple intracellular metabolites may regulate TET-mediated oxidative DNA demethylation, at least under certain pathological conditions.

Ascorbate (also known as vitamin C), an essential nutrient for humans and certain other animal species, has also been demonstrated to positively regulate TET activity [111–113]. In wild-type, but not *Tet1/Tet2* deficient mouse ESCs, ascorbate significantly increases the levels of all 5mC oxidation products, particularly 5fC and 5caC by more than an order of magnitude, leading to a global loss of 5mC (~40%) [111, 113]. Ascorbate uniquely interacts with the catalytic domain of TET enzymes, enhancing their catalytic activities likely by promoting their folding and/or recycling of Fe(II) [111]. Intriguingly, ascorbate-induced demethylation has stronger effect on the DNA sequences that gain methylation in cultured ES cells compared to blastocysts, which are typically methylated only after implantation in vivo [113]. These studies suggest that ascorbate is a positive modulator of TET activity and may play a critical role in regulating DNA methylation during development. Further studies are needed to elucidate the sequence specificity of ascorbate-mediated stimulation of TET activity.

6.3 Regulation by TET-Interacting Proteins and DNA-Binding Proteins

In addition to the overall TET activity in cells, the specific targeting of TET proteins, and the regulation of their processivity (i.e., why TET-mediated 5mC oxidation tends to stall at 5hmC, and only proceed to 5fC and 5caC at specific loci) provide another important layer of demethylation control. Emerging evidence indicates that the genomic targeting, activity, and processivity of TET enzymes can be modulated by their interacting proteins and some DNA-binding proteins. The *O*-linked N-acetylglucosamine (O-GlcNAc) transferase OGT, has been reported to directly interact with, and also GlcNAcylate TET proteins [114–118]. Although OGT binding and GlcNAcylation appear not to regulate the enzymatic activity of TET proteins [115], OGT regulates the subcellular location of TET3 by promoting its nuclear export in high-glucose conditions [118]. Moreover, depletion of OGT in mouse ESCs decreases the association of TET1 with chromatin and alters 5hmC enrichment at certain loci [116, 117], suggesting that OGT plays a specific role in targeting and stabilizing TET proteins to the chromatin. In addition to OGT, the CXXC domain protein IDAX was shown to interact with TET2 and was suggested to recruit TET2 to promoters and CpG islands [41]. Recently, a sequence specific transcription factor WT1 (Wilms tumor protein 1) has also been shown to physically interact with and recruits TET2 to its target genes to activate their expression [119]. Furthermore, PGC7 (also known as STELLA or DPPA3), a maternal factor essential for early development, was demonstrated in one-cell zygotes to protect maternal genome and the imprinting control regions (ICRs) in paternal genome by inhibiting TET activity through direct interaction [97, 120, 121]. These findings suggest an important role of TET interacting partners in targeting and restricting TET activity in the cell.

In addition to TET interacting proteins, some DNA binding proteins can also regulate DNA demethylation. For example, one study showed that knockdown of methyl-CpG binding domain protein 3 (MBD3), which also binds to 5hmC, caused a strong reduction in global 5hmC level in mouse ESCs [122]. In another study, UHRF2 was identified as a 5hmC-specific binding protein in neuronal progenitor cells, and was shown to be capable of stimulating the processivity of TET1 when co-overexpressed with the catalytic domain of TET1 in HEK293T cells [123]. These observations demonstrate that proteins bound to the substrate DNA of TET enzymes may regulate the enzymes' activity and processivity, therefore implying a role of DNA binding proteins in controlling DNA demethylation.

7 Concluding Remarks

Ever since the discovery of 5mC oxidation by TET family proteins, there has been tremendous progress in understanding the molecular basis of DNA demethylation. Accumulating biochemical and genetic studies have demonstrated that TET family proteins play a critical role in active DNA demethylation during dynamic regulation of DNA methylation patterns in development and diseases. While the TET-TDG-BER pathway and replication-dependent dilution of 5hmC/5fC/5caC have been generally accepted as the major forces of active DNA demethylation, other potential active DNA demethylation mechanisms have also been reported. It is worth noting that most of those observations were made before knowing the existence of 5hmC/5fC/5caC, and by immunostaining of 5mC or bisulfite sequencing that do not distinguish 5mC from 5hmC or unmodified cytosine from 5fC and 5caC [9]. Therefore, with many new technologies recently developed to map various new modifications in the genome, historically reported active DNA demethylation pathways should be revisited to further advance this exciting field by revealing a more comprehensive understanding on how DNA methylation is dynamically regulated.

References

1. Bird A (2002) DNA methylation patterns and epigenetic memory. *Genes Dev* 16(1):6–21
2. Jenuwein T, Allis CD (2001) Translating the histone code. *Science* 293(5532):1074–1080
3. Holliday R, Pugh JE (1975) DNA modification mechanisms and gene activity during development. *Science* 187(4173):226–232
4. Riggs AD (1975) X inactivation, differentiation, and DNA methylation. *Cytogenet Cell Genet* 14(1):9–25
5. Henikoff S, Grewal JM (2016) Epigenetics, cellular memory and gene regulation. *Curr Biol* 26(14):R644–R648
6. Bonasio R, Tu S, Reinberg D (2010) Molecular signals of epigenetic states. *Science* 330(6004):612–616
7. Easwaran H, Tsai HC, Baylin SB (2014) Cancer epigenetics: tumor heterogeneity, plasticity of stem-like states, and drug resistance. *Mol Cell* 54(5):716–727
8. Goll MG, Bestor TH (2005) Eukaryotic cytosine methyltransferases. *Annu Rev Biochem* 74:481–514
9. Shen L, Song CX, He C, Zhang Y (2014) Mechanism and function of oxidative reversal of DNA and RNA methylation. *Annu Rev Biochem* 83:585–614
10. Hermann A, Goyal R, Jeltsch A (2004) The Dnmt1 DNA-(cytosine-C5)-methyltransferase methylates DNA processively with high preference for hemimethylated target sites. *J Biol Chem* 279(46):48350–48359
11. Sharif J, Muto M, Takebayashi S, Suetake I, Iwamatsu A, Endo TA, Shinga J, Mizutani-Koseki Y, Toyoda T, Okamura K, Tajima S, Mitsuya K, Okano M, Koseki H (2007) The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. *Nature* 450(7171):908–912
12. Bostick M, Kim JK, Esteve PO, Clark A, Pradhan S, Jacobsen SE (2007) UHRF1 plays a role in maintaining DNA methylation in mammalian cells. *Science* 317(5845):1760–1764
13. Okano M, Bell DW, Haber DA, Li E (1999) DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* 99(3):247–257

14. Li E, Bestor TH, Jaenisch R (1992) Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* 69(6):915–926
15. Li E, Zhang Y (2014) DNA methylation in mammals. *Cold Spring Harb Perspect Biol* 6(5):a019133
16. Liu Y, Zhang X, Blumenthal RM, Cheng X (2013) A common mode of recognition for methylated CpG. *Trends Biochem Sci* 38(4):177–183
17. Bell AC, Felsenfeld G (2000) Methylation of a CTCF-dependent boundary controls imprinted expression of the *Igf2* gene. *Nature* 405(6785):482–485
18. Kim J, Kollhoff A, Bergmann A, Stubbs L (2003) Methylation-sensitive binding of transcription factor YY1 to an insulator sequence within the paternally expressed imprinted gene, *Peg3*. *Hum Mol Genet* 12(3):233–245
19. Mosammaparast N, Shi Y (2010) Reversal of histone methylation: biochemical and molecular mechanisms of histone demethylases. *Annu Rev Biochem* 79:155–179
20. Rougier N, Bourc'his D, Gomes DM, Niveleau A, Plachot M, Paldi A, Viegas-Pequignot E (1998) Chromosome methylation patterns during mammalian preimplantation development. *Genes Dev* 12(14):2108–2113
21. Shen L, Inoue A, He J, Liu Y, Lu F, Zhang Y (2014) Tet3 and DNA replication mediate demethylation of both the maternal and paternal genomes in mouse zygotes. *Cell Stem Cell* 15(4):459–470
22. Guo F, Li X, Liang D, Li T, Zhu P, Guo H, Wu X, Wen L, Gu TP, Hu B, Walsh CP, Li J, Tang F, Xu GL (2014) Active and passive demethylation of male and female pronuclear DNA in the mammalian zygote. *Cell Stem Cell* 15(4):447–458
23. Seisenberger S, Andrews S, Krueger F, Arand J, Walter J, Santos F, Popp C, Thienpont B, Dean W, Reik W (2012) The dynamics of genome-wide DNA methylation reprogramming in mouse primordial germ cells. *Mol Cell* 48(6):849–862
24. Kagiwada S, Kurimoto K, Hirota T, Yamaji M, Saitou M (2013) Replication-coupled passive DNA demethylation for the erasure of genome imprints in mice. *EMBO J* 32(3):340–353
25. Vincent JJ, Huang Y, Chen PY, Feng S, Calvopina JH, Nee K, Lee SA, Le T, Yoon AJ, Faull K, Fan G, Rao A, Jacobsen SE, Pellegrini M, Clark AT (2013) Stage-specific roles for tet1 and tet2 in DNA demethylation in primordial germ cells. *Cell Stem Cell* 12(4):470–478
26. Bruniquel D, Schwartz RH (2003) Selective, stable demethylation of the interleukin-2 gene enhances transcription by an active process. *Nat Immunol* 4(3):235–240
27. Martinowich K, Hattori D, Wu H, Fouse S, He F, Hu Y, Fan G, Sun YE (2003) DNA methylation-related chromatin remodeling in activity-dependent BDNF gene regulation. *Science* 302(5646):890–893
28. Kangaspeska S, Stride B, Metivier R, Polycarpou-Schwarz M, Ibberson D, Carmouche RP, Benes V, Gannon F, Reid G (2008) Transient cyclical methylation of promoter DNA. *Nature* 452(7183):112–115
29. Metivier R, Gallais R, Tiffoche C, Le Peron C, Jurkowska RZ, Carmouche RP, Ibberson D, Barath P, Demay F, Reid G, Benes V, Jeltsch A, Gannon F, Salbert G (2008) Cyclical DNA methylation of a transcriptionally active promoter. *Nature* 452(7183):45–50
30. Thillainadesan G, Chitilian JM, Isovich M, Ablack JN, Mymryk JS, Tini M, Torchia J (2012) TGF-beta-dependent active demethylation and expression of the p15^{ink4b} tumor suppressor are impaired by the ZNF217/CoREST complex. *Mol Cell* 46(5):636–649
31. Ono R, Taki T, Taketani T, Taniwaki M, Kobayashi H, Hayashi Y (2002) LCX, leukemia-associated protein with a CXXC domain, is fused to MLL in acute myeloid leukemia with trilineage dysplasia having t(10;11)(q22;q23). *Cancer Res* 62(14):4075–4080
32. Lorschach RB, Moore J, Mathew S, Raimondi SC, Mukatira ST, Downing JR (2003) TET1, a member of a novel protein family, is fused to MLL in acute myeloid leukemia containing the t(10;11)(q22;q23). *Leuk Off J Leuk Soc Am Leuk Res Fund UK* 17(3):637–641
33. Iyer LM, Tahiliani M, Rao A, Aravind L (2009) Prediction of novel families of enzymes involved in oxidative and other complex modifications of bases in nucleic acids. *Cell Cycle* 8(11):1698–1710

34. Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, Agarwal S, Iyer LM, Liu DR, Aravind L, Rao A (2009) Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* 324(5929):930–935
35. McDonough MA, Loenarz C, Chowdhury R, Clifton IJ, Schofield CJ (2010) Structural studies on human 2-oxoglutarate dependent oxygenases. *Curr Opin Struct Biol* 20(6):659–672
36. Iyer LM, Abhiman S, Aravind L (2011) Natural history of eukaryotic DNA methylation systems. *Prog Mol Biol Transl Sci* 101:25–104
37. Hu L, Li Z, Cheng J, Rao Q, Gong W, Liu M, Shi YG, Zhu J, Wang P, Xu Y (2013) Crystal structure of TET2-DNA complex: insight into TET-mediated 5mC oxidation. *Cell* 155(7):1545–1555
38. Xu Y, Xu C, Kato A, Tempel W, Abreu JG, Bian C, Hu Y, Hu D, Zhao B, Cerovina T, Diao J, Wu F, He HH, Cui Q, Clark E, Ma C, Barbara A, Veenstra GJ, Xu G, Kaiser UB, Liu XS, Sugrue SP, He X, Min J, Kato Y, Shi YG (2012) Tet3 CXXC domain and dioxygenase activity cooperatively regulate key genes for *Xenopus* eye and neural development. *Cell* 151(6):1200–1213
39. Ito S, D'Alessio AC, Taranova OV, Hong K, Sowers LC, Zhang Y (2010) Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature* 466(7310):1129–1133
40. Liu N, Wang M, Deng W, Schmidt CS, Qin W, Leonhardt H, Spada F (2013) Intrinsic and extrinsic connections of Tet3 Dioxygenase with CXXC zinc finger modules. *PLoS One* 8(5):e62755
41. Ko M, An J, Bandukwala H, Chavez L (2013) Modulation of TET2 expression and 5-methylcytosine oxidation by the CXXC domain protein IDAX. *Nature* 497:122–126
42. Jin SG, Zhang ZM, Dunwell TL, Harter MR, Wu X, Johnson J, Li Z, Liu J, Szabo PE, Lu Q, Xu GL, Song J, Pfeifer GP (2016) Tet3 reads 5-Carboxylcytosine through its CXXC domain and is a potential Guardian against Neurodegeneration. *Cell Rep* 14(3):493–505
43. Borst P, Sabatini R (2008) Base J: discovery, biosynthesis, and possible functions. *Annu Rev Microbiol* 62:235–251
44. Krebs C, Galonic Fujimori D, Walsh CT, Bollinger JM Jr (2007) Non-heme Fe(IV)-oxo intermediates. *Acc Chem Res* 40(7):484–492
45. Smiley JA, Kundracik M, Landfried DA, Barnes VR Sr, Axhemi AA (2005) 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* 1723(1–3):256–264
46. Wu SC, Zhang Y (2010) Active DNA demethylation: many roads lead to Rome. *Nat Rev Mol Cell Biol* 11(9):607–620
47. Ito S, Shen L, Dai Q, Wu SC, Collins LB, Swenberg JA, He C, Zhang Y (2011) Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science* 333(6047):1300–1303
48. He YF, Li BZ, Li Z, Liu P, Wang Y, Tang Q, Ding J, Jia Y, Chen Z, Li L, Sun Y, Li X, Dai Q, Song CX, Zhang K, He C, Xu GL (2011) Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science* 333(6047):1303–1307
49. Maiti A, Drohat AC (2011) Thymine DNA glycosylase can rapidly excise 5-formylcytosine and 5-carboxylcytosine: potential implications for active demethylation of CpG sites. *J Biol Chem* 286(41):35334–35338
50. Zhang H, Zhu JK (2012) Active DNA demethylation in plants and animals. *Cold Spring Harb Symp Quant Biol* 77:161–173
51. Stivers JT, Jiang YL (2003) A mechanistic perspective on the chemistry of DNA repair glycosylases. *Chem Rev* 103(7):2729–2759
52. Bennett MT, Rodgers MT, Hebert AS, Ruslander LE, Eisele L, Drohat AC (2006) Specificity of human thymine DNA glycosylase depends on N-glycosidic bond stability. *J Am Chem Soc* 128(38):12510–12519

53. Williams RT, Wang Y (2012) A density functional theory study on the kinetics and thermodynamics of N-glycosidic bond cleavage in 5-substituted 2'-deoxycytidines. *Biochemistry* 51(32):6458–6462
54. Nabel CS, Jia H, Ye Y, Shen L, Goldschmidt HL, Stivers JT, Zhang Y, Kohli RM (2012) AID/APOBEC deaminases disfavor modified cytosines implicated in DNA demethylation. *Nat Chem Biol* 8(9):751–758
55. Shen L, Wu H, Diep D, Yamaguchi S, D'Alessio AC, Fung HL, Zhang K, Zhang Y (2013) Genome-wide analysis reveals TET- and TDG-dependent 5-methylcytosine oxidation dynamics. *Cell* 153(3):692–706
56. Song C-X, Szulwach KE, Dai Q, Fu Y, Mao S-Q, Lin L, Street C, Li Y, Poidevin M, Wu H, Gao J, Liu P, Li L, Xu G-L, Jin P, He C (2013) Genome-wide profiling of 5-Formylcytosine reveals its roles in epigenetic priming. *Cell* 153:678–691
57. Cortellino S, Xu J, Sannai M, Moore R, Caretti E, Cigliano A, Le Coz M, Devarajan K, Wessels A, Soprano D, Abramowitz LK, Bartolomei MS, Rambow F, Bassi MR, Bruno T, Fanciulli M, Renner C, Klein-Szanto AJ, Matsumoto Y, Kobi D, Davidson I, Alberti C, Larue L, Bellacosa A (2011) Thymine DNA glycosylase is essential for active DNA demethylation by linked deamination-base excision repair. *Cell* 146(1):67–79
58. Cortazar D, Kunz C, Selfridge J, Lettieri T, Saito Y, MacDougall E, Wirz A, Schuermann D, Jacobs AL, Siegrist F, Steinacher R, Jiricny J, Bird A, Schar P (2011) Embryonic lethal phenotype reveals a function of TDG in maintaining epigenetic stability. *Nature* 470(7334):419–423
59. Millar CB, Guy J, Sansom OJ, Selfridge J, MacDougall E, Hendrich B, Keightley PD, Bishop SM, Clarke AR, Bird A (2002) Enhanced CpG mutability and tumorigenesis in MBD4-deficient mice. *Science* 297(5580):403–405
60. Kemmerich K, Dingler FA, Rada C, Neuberger MS (2012) Germline ablation of SMUG1 DNA glycosylase causes loss of 5-hydroxymethyluracil- and UNG-backup uracil-excision activities and increases cancer predisposition of Ung^{-/-}Msh2^{-/-} mice. *Nucleic Acids Res* 40(13):6016–6025
61. Zhang L, Lu X, Lu J, Liang H, Dai Q, Xu GL, Luo C, Jiang H, He C (2012) Thymine DNA glycosylase specifically recognizes 5-carboxylcytosine-modified DNA. *Nat Chem Biol* 8(4):328–330
62. Hashimoto H, Liu Y, Upadhyay AK, Chang Y, Howerton SB, Vertino PM, Zhang X, Cheng X (2012) Recognition and potential mechanisms for replication and erasure of cytosine hydroxymethylation. *Nucleic Acids Res* 40(11):4841–4849
63. Valinluck V, Sowers LC (2007) Endogenous cytosine damage products alter the site selectivity of human DNA maintenance methyltransferase DNMT1. *Cancer Res* 67(3):946–950
64. Ji D, Lin K, Song J, Wang Y (2014) Effects of Tet-induced oxidation products of 5-methylcytosine on Dnmt1- and DNMT3a-mediated cytosine methylation. *Mol BioSyst* 10(7):1749–1752
65. Kohli RM, Zhang Y (2013) TET enzymes, TDG and the dynamics of DNA demethylation. *Nature* 502(7472):472–479
66. Inoue A, Zhang Y (2011) Replication-dependent loss of 5-hydroxymethylcytosine in mouse preimplantation embryos. *Science* 334(6053):194
67. Hackett JA, Sengupta R, Zyllicz JJ, Murakami K, Lee C, Down TA, Surani MA (2013) Germline DNA demethylation dynamics and imprint erasure through 5-hydroxymethylcytosine. *Science* 339(6118):448–452
68. Schiesser S, Hackner B, Pfaffeneder T, Muller M, Hagemeyer C, Truss M, Carell T (2012) Mechanism and stem-cell activity of 5-carboxycytosine decarboxylation determined by isotope tracing. *Angew Chem Int Ed* 51(26):6516–6520
69. Chen CC, Wang KY, Shen CK (2012) The mammalian de novo DNA methyltransferases DNMT3A and DNMT3B are also DNA 5-hydroxymethylcytosine dehydroxymethylases. *J Biol Chem* 287(40):33116–33121
70. Liutkeviciute Z, Lukinavicius G, Masevicius V, Daujotyte D, Klimasauskas S (2009) Cytosine-5-methyltransferases add aldehydes to DNA. *Nat Chem Biol* 5(6):400–402

71. Liutkeviciute Z, Kriukiene E, Licyte J, Rudyte M, Urbanaviciute G, Klimasauskas S (2014) Direct decarboxylation of 5-carboxylcytosine by DNA C5-methyltransferases. *J Am Chem Soc* 136(16):5884–5887
72. Conticello SG (2008) The AID/APOBEC family of nucleic acid mutators. *Genome Biol* 9(6):229
73. Guo JU, Su Y, Zhong C, Ming GL, Song H (2011) Hydroxylation of 5-methylcytosine by TET1 promotes active DNA demethylation in the adult brain. *Cell* 145(3):423–434
74. Bransteitter R, Pham P, Scharff MD, Goodman MF (2003) Activation-induced cytidine deaminase deaminates deoxycytidine on single-stranded DNA but requires the action of RNase. *Proc Natl Acad Sci U S A* 100(7):4102–4107
75. Rangam G, Schmitz KM, Cobb AJ, Petersen-Mahrt SK (2012) AID enzymatic activity is inversely proportional to the size of cytosine C5 orbital cloud. *PLoS One* 7(8):e43279
76. Ooi SK, Bestor TH (2008) The colorful history of active DNA demethylation. *Cell* 133(7):1145–1148
77. Bhattacharya SK, Ramchandani S, Cervoni N, Szyf M (1999) A mammalian protein with specific demethylase activity for mCpG DNA. *Nature* 397(6720):579–583
78. Santos F, Hendrich B, Reik W, Dean W (2002) Dynamic reprogramming of DNA methylation in the early mouse embryo. *Dev Biol* 241(1):172–182
79. Okada Y, Yamagata K, Hong K, Wakayama T, Zhang Y (2010) A role for the elongator complex in zygotic paternal genome demethylation. *Nature* 463(7280):554–558
80. Chen CC, Wang KY, Shen CK (2013) DNA 5-methylcytosine demethylation activities of the mammalian DNA methyltransferases. *J Biol Chem* 288(13):9084–9091
81. Zhu JK (2009) Active DNA demethylation mediated by DNA glycosylases. *Annu Rev Genet* 43:143–166
82. Zhu B, Zheng Y, Hess D, Angliker H, Schwarz S, Siegmann M, Thiry S, Jost JP (2000) 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* 97(10):5135–5139
83. Zhu B, Zheng Y, Angliker H, Schwarz S, Thiry S, Siegmann M, Jost JP (2000) 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* 28(21):4157–4165
84. Morgan HD, Dean W, Coker HA, Reik W, Petersen-Mahrt SK (2004) Activation-induced cytidine deaminase deaminates 5-methylcytosine in DNA and is expressed in pluripotent tissues: implications for epigenetic reprogramming. *J Biol Chem* 279(50):52353–52360
85. Rai K, Huggins IJ, James SR, Karpf AR, Jones DA, Cairns BR (2008) DNA demethylation in zebrafish involves the coupling of a deaminase, a glycosylase, and gadd45. *Cell* 135(7):1201–1212
86. Popp C, Dean W, Feng S, Cokus SJ, Andrews S, Pellegrini M, Jacobsen SE, Reik W (2010) Genome-wide erasure of DNA methylation in mouse primordial germ cells is affected by AID deficiency. *Nature* 463(7284):1101–1105
87. Bhutani N, Brady JJ, Damian M, Sacco A, Corbel SY, Blau HM (2010) Reprogramming towards pluripotency requires AID-dependent DNA demethylation. *Nature* 463(7284):1042–1047
88. Kumar R, DiMenna L, Schrode N, Liu TC, Franck P, Munoz-Descalzo S, Hadjantonakis AK, Zarrin AA, Chaudhuri J, Elemento O, Evans T (2013) AID stabilizes stem-cell phenotype by removing epigenetic memory of pluripotency genes. *Nature* 500(7460):89–92
89. Barreto G, Schafer A, Marhold J, Stach D, Swaminathan SK, Handa V, Doderlein G, Maltry N, Wu W, Lyko F, Niehrs C (2007) Gadd45a promotes epigenetic gene activation by repair-mediated DNA demethylation. *Nature* 445(7128):671–675
90. Ma DK, Jang MH, Guo JU, Kitabatake Y, Chang ML, Pow-Anpongkul N, Flavell RA, Lu B, Ming GL, Song H (2009) Neuronal activity-induced Gadd45b promotes epigenetic DNA demethylation and adult neurogenesis. *Science* 323(5917):1074–1077

91. Schmitz KM, Schmitt N, Hoffmann-Rohrer U, Schafer A, Grummt I, Mayer C (2009) TAF12 recruits Gadd45a and the nucleotide excision repair complex to the promoter of rRNA genes leading to active DNA demethylation. *Mol Cell* 33(3):344–353
92. Engel N, Tront JS, Erinle T, Nguyen N, Latham KE, Sapienza C, Hoffman B, Liebermann DA (2009) Conserved DNA methylation in Gadd45a(-/-) mice. *Epigenetics* 4(2):98–99
93. Jin SG, Guo C, Pfeifer GP (2008) GADD45A does not promote DNA demethylation. *PLoS Genet* 4(3):e1000013
94. Wu H, Zhang Y (2014) Reversing DNA methylation: mechanisms, genomics, and biological functions. *Cell* 156(1–2):45–68
95. Yamaguchi S, Hong K, Liu R, Shen L, Inoue A, Diep D, Zhang K, Zhang Y (2012) Tet1 controls meiosis by regulating meiotic gene expression. *Nature* 492(7429):443–447
96. Yamaguchi S, Hong K, Liu R, Inoue A, Shen L, Zhang K, Zhang Y (2013) Dynamics of 5-methylcytosine and 5-hydroxymethylcytosine during germ cell reprogramming. *Cell Res* 23(3):329–339
97. Wossidlo M, Nakamura T, Lepikhov K, Marques CJ, Zakhartchenko V, Boiani M, Arand J, Nakano T, Reik W, Walter J (2011) 5-hydroxymethylcytosine in the mammalian zygote is linked with epigenetic reprogramming. *Nat Commun* 2:241
98. Iqbal K, Jin SG, Pfeifer GP, Szabo PE (2011) Reprogramming of the paternal genome upon fertilization involves genome-wide oxidation of 5-methylcytosine. *Proc Natl Acad Sci U S A* 108(9):3642–3647
99. Ficiz G, Branco MR, Seisenberger S, Santos F, Krueger F, Hore TA, Marques CJ, Andrews S, Reik W (2011) Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation. *Nature* 473(7347):398–402
100. Koh KP, Yabuuchi A, Rao S, Huang Y, Cunniff K, Nardone J, Laiho A, Tahiliani M, Sommer CA, Mostoslavsky G, Lahesmaa R, Orkin SH, Rodig SJ, Daley GQ, Rao A (2011) Tet1 and Tet2 regulate 5-hydroxymethylcytosine production and cell lineage specification in mouse embryonic stem cells. *Cell Stem Cell* 8(2):200–213
101. Song SJ, Poliseno L, Song MS, Ala U, Webster K, Ng C, Beringer G, Brikbak NJ, Yuan X, Cantley LC, Richardson AL, Pandolfi PP (2013) MicroRNA-antagonism regulates breast cancer Stemness and metastasis via TET-family-dependent chromatin remodeling. *Cell* 154(2):311–324
102. Song Su J, Ito K, Ala U, Kats L, Webster K, Sun Su M, Jongen-Lavrencic M, Manova-Todorova K, Teruya-Feldstein J, Avigan David E, Delwel R, Pandolfi Pier P (2013) The oncogenic microRNA miR-22 targets the TET2 tumor suppressor to promote hematopoietic stem cell self-renewal and transformation. *Cell Stem Cell* 13:87–101
103. Wang Y, Zhang Y (2014) Regulation of TET protein stability by Calpains. *Cell Rep* 6(2):278–284
104. Coulter JB, O'Driscoll CM, Bressler JP (2013) Hydroquinone increases 5-hydroxymethylcytosine formation through ten eleven translocation 1 (Tet1) 5-methylcytosine dioxygenase. *J Biol Chem* 288(40):28792–28800
105. Yang H, Lin H, Xu H, Zhang L, Cheng L, Wen B, Shou J, Guan K, Xiong Y, Ye D (2014) TET-catalyzed 5-methylcytosine hydroxylation is dynamically regulated by metabolites. *Cell Res* 24(8):1017–1020
106. Figueroa ME, Abdel-Wahab O, Lu C, Ward PS, Patel J, Shih A, Li Y, Bhagwat N, Vasanthakumar A, Fernandez HF, Tallman MS, Sun Z, Wolniak K, Peeters JK, Liu W, Choe SE, Fantin VR, Paietta E, Lowenberg B, Licht JD, Godley LA, Delwel R, Valk PJ, Thompson CB, Levine RL, Melnick A (2010) Leukemic IDH1 and IDH2 mutations result in a hypermethylation phenotype, disrupt TET2 function, and impair hematopoietic differentiation. *Cancer Cell* 18(6):553–567
107. Xu W, Yang H, Liu Y, Yang Y, Wang P, Kim SH, Ito S, Yang C, Xiao MT, Liu LX, Jiang WQ, Liu J, Zhang JY, Wang B, Frye S, Zhang Y, Xu YH, Lei QY, Guan KL, Zhao SM, Xiong Y (2011) Oncometabolite 2-hydroxyglutarate is a competitive inhibitor of alpha-ketoglutarate-dependent dioxygenases. *Cancer Cell* 19(1):17–30

108. Ward PS, Patel J, Wise DR, Abdel-Wahab O, Bennett BD, Collier HA, Cross JR, Fantin VR, Hedvat CV, Perl AE, Rabinowitz JD, Carroll M, Su SM, Sharp KA, Levine RL, Thompson CB (2010) The common feature of leukemia-associated IDH1 and IDH2 mutations is a neomorphic enzyme activity converting alpha-ketoglutarate to 2-hydroxyglutarate. *Cancer Cell* 17(3):225–234
109. Dang L, White DW, Gross S, Bennett BD, Bittinger MA, Driggers EM, Fantin VR, Jang HG, Jin S, Keenan MC, Marks KM, Prins RM, Ward PS, Yen KE, Liao LM, Rabinowitz JD, Cantley LC, Thompson CB, Vander Heiden MG, Su SM (2009) Cancer-associated IDH1 mutations produce 2-hydroxyglutarate. *Nature* 462(7274):739–744
110. Xiao M, Yang H, Xu W, Ma S, Lin H, Zhu H, Liu L, Liu Y, Yang C, Xu Y, Zhao S, Ye D, Xiong Y, Guan KL (2012) Inhibition of alpha-KG-dependent histone and DNA demethylases by fumarate and succinate that are accumulated in mutations of FH and SDH tumor suppressors. *Genes Dev* 26(12):1326–1338
111. Yin R, Mao S-Q, Zhao B, Chong Z, Yang Y, Zhao C, Zhang D, Huang H, Gao J, Li Z, Jiao Y, Li C, Liu S, Wu D, Gu W, Yang Y-G, Xu G-L, Wang H (2013) Ascorbic acid enhances Tet-mediated 5-methylcytosine oxidation and promotes DNA demethylation in mammals. *J Am Chem Soc* 135(28):10396–10403
112. Minor EA, Court BL, Young JI, Wang G (2013) Ascorbate induces ten-eleven translocation (Tet) methylcytosine dioxygenase-mediated generation of 5-hydroxymethylcytosine. *J Biol Chem* 288(19):13669–13674
113. Blaschke K, Ebata KT, Karimi MM, Zepeda-Martínez JA, Goyal P, Mahapatra S, Tam A, Laird DJ, Hirst M, Rao A, Lorincz MC, Ramalho-Santos M (2013) Vitamin C induces Tet-dependent DNA demethylation and a blastocyst-like state in ES cells. *Nature* 500(7461):222–226
114. Deplus R, Delatte B, Schwinn MK, Defrance M, Méndez J, Murphy N, Ma D, Volkmar M, Putmans P, Calonne E, Shih AH, Levine RL, Bernard O, Mercher T, Solary E, Uhr M, Daniels DL, Fuks F (2013) TET2 and TET3 regulate GlcNAcylation and H3K4 methylation through OGT and SET1/COMPASS. *EMBO J* 32:645–655
115. Chen Q, Chen Y, Bian C, Fujiki R, Yu X (2013) TET2 promotes histone O-GlcNAcylation during gene transcription. *Nature* 493:561–564
116. Vella P, Scelfo A, Jammula S, Chiacchiera F, Williams K, Cuomo A, Roberto A, Christensen J, Bonaldi T, Helin K, Pasini D (2013) Tet proteins connect the O-linked N-acetylglucosamine transferase Ogt to chromatin in embryonic stem cells. *Mol Cell* 49:645–656
117. Shi FT, Kim H, Lu W, He Q, Liu D, Goodell MA, Wan M, Songyang Z (2013) Ten-eleven translocation 1 (tet1) is regulated by o-linked N-acetylglucosamine transferase (ogt) for target gene repression in mouse embryonic stem cells. *J Biol Chem* 288(29):20776–20784
118. Zhang Q, Liu X, Gao W, Li P, Hou J, Li J, Wong J (2014) Differential regulation of the ten-eleven translocation (TET) family of dioxygenases by O-linked beta-N-acetylglucosamine transferase (OGT). *J Biol Chem* 289(9):5986–5996
119. Wang Y, Xiao M, Chen X, Chen L, Xu Y, Lv L, Wang P, Yang H, Ma S, Lin H, Jiao B, Ren R, Ye D, Guan KL, Xiong Y (2015) WT1 recruits TET2 to regulate its target gene expression and suppress leukemia cell proliferation. *Mol Cell* 57(4):662–673
120. Nakamura T, Liu YJ, Nakashima H, Umehara H, Inoue K, Matoba S, Tachibana M, Ogura A, Shinkai Y, Nakano T (2012) PGC7 binds histone H3K9me2 to protect against conversion of 5mC to 5hmC in early embryos. *Nature* 486(7403):415–419
121. Bian C, Yu X (2014) PGC7 suppresses TET3 for protecting DNA methylation. *Nucleic Acids Res* 42(5):2893–2905
122. Yildirim O, Li R, Hung JH, Chen PB, Dong X, Ee LS, Weng Z, Rando OJ, Fazzio TG (2011) Mbd3/NURD complex regulates expression of 5-hydroxymethylcytosine marked genes in embryonic stem cells. *Cell* 147(7):1498–1510
123. Spruijt CG, Gnerlich F, Smits AH, Pfaffeneder T, Jansen PWTC, Bauer C, Münzel M, Wagner M, Müller M, Khan F, Eberl HC, Mensinga A, Brinkman AB, Lephikov K, Müller U, Walter J, Boelens R, van Ingen H, Leonhardt H, Carell T, Vermeulen M (2013) Dynamic readers for 5-(hydroxy)methylcytosine and its oxidized derivatives. *Cell* 152:1146–1159