Cell Research (2011) 21:1649-1651.

www.nature.com/cr

RESEARCH HIGHLIGHT

Active DNA demethylation by oxidation and repair

Zhizhong Gong[1], Jian-Kang Zhu[2, 3]

[1]State Key Laboratory of Plant Physiology and Biochemistry, College of Biological Sciences, China Agricultural University, Beijing 100193, China; [2]Center for Plant Stress Genomics and Technology, King Abdullah University of Science and Technology, Thuwal 23955-6900, Kingdom of Saudi Arabia; [3]Department of Horticulture and Landscape Architecture, Purdue University, West Lafayette, IN 47907-2010, USA

Cell Research (2011) **21**:1649-1651. doi:10.1038/cr.2011.140; published online 23 August 2011

DNA methylation and demethylation are increasingly recognized as important epigenetic factors in both plants and animals. DNA methylation, which is catalyzed by DNA methyltransferases (DNMTs), is a relatively stable and heritable modification that controls gene expression, cellular differentiation, genomic imprinting, paramutation, transposon movement, X-inactivation, and embryogenesis [1]. The methylation of cytosine to 5-methylcytosine (5mC) is an important example of DNA modification in animals and plants. This highlight concerns DNA demethylation mechanisms in mammals and whether they are similar to that in plants.

DNA demethylation can be passive or active. In both plants and animals, passive DNA demethylation occurs when cells fail to maintain the methylation during DNA replication [1]. In plants, active DNA demethylation is mainly carried out by a small group of bifunctional DNA glycosylases, including ROS1, DME, DML2 and DML3; after these DNA glycosylases remove the methylated cytosine base and create an abasic site, the gap is refilled with an unmethylated cytosine through a base-excision-repair (BER) pathway [1]. Animal genomes, however, do not contain ROS1 homologues, and active DNA demethylation in animals is proposed to result from several other mechanisms [2]. An oxidative mechanism in mammals has recently been revealed by studies on a small family of 5-methylcytosine hydroxylases (TET1, TET2, TET3; "TET" refers to Ten-Eleven-Translocation); these studies suggest that active DNA demethylation in mammals also goes through BER, but unlike that in plants, it must be preceded by oxidation and/or deamination [3-7].

5mC can be hydroxylated by TETs to become 5-hydroxymethylcytosine (5hmC). 5hmC was first reported in bacteriophage nucleic acids [8] and later in animal cells [9]. The occurrence of 5hmC in animal cells was debated until 2009, when two groups independently confirmed the presence of 5hmC in the animal genome [7, 10]. Tahiliani *et al*. found that TET1 in humans is responsible for the conversion of 5mC to 5mhC in a 2-oxoglutarate- and Fe(II)-dependent manner [7]. Later, Zhang and colleagues at the University of North Carolina found that all three mouse TET proteins can catalyze this conversion [6]. Recently, Zhang and colleagues reported that 5mC can also be converted into 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) [11]. Their data showed that TET proteins oxidize the 5mC to generate 5hmC, 5fC, and 5caC *in vitro*. Their analysis of the genomic DNA indicated that all four cytosine modifications exist in different quantities. They further showed that through increasing or decreasing the expression of TET proteins, the contents of 5hmC, 5fC, and 5caC are increased or decreased, respectively [11].

Recent studies indicate that active DNA demethylation in mammals may follow a oxidation-deamination-BER pathway [3-5]. AID (activation-induced deaminase)-dependent DNA demethylation was found during a study of nuclear reprogramming when mouse embryonic stem cells (mESCs) were fused with human fibroblasts [12]. AID is required for promoter demethylation of *OCT4* and *NANOG* genes [12]. Guo *et al*. provided further evidence that the AID/APOBEC (apolipoprotein B mRNA-editing catalytic polypeptides) family of cytidine deaminases deaminates 5hmC, but not 5mC, into 5hmU [4]. In another study, Cortellino *et al*. identified the thymine DNA glycosylase (TDG) as a key enzyme that exhibits high glycosylase activity on 5hmU:G mismatches in double-stranded DNA but not on 5hmC [5]. After TDG removes 5hmU, the DNA is repaired and an unmethylated cytosine is inserted through the BER pathway [4, 5]. Mouse primordial germ cells with knockout of AID exhibit higher genome-wide DNA methylation than the wild type, indicat-

Correspondence: Zhizhong Gong
E-mail: gongzz@cau.edu.cn

ing that AID is involved in global DNA demethylation [13]. The fact that AID knockout mice are smaller but not lethal suggested that other deaminases such as APOBEC1-3 might also be involved in DNA demethylation [13]. TDG knockout or knockin mice (knockin mice are generated by introducing a N151A point mutation in TDG glycosylase domain) are embryonically lethal, indicating the importance of TDG glycosylases in DNA demethylation and early embryonic development. Interestingly, Cortellino *et al*. found that TDG physically interacts with AID and GADD45a (growth arrest and DNA damage-inducible protein 45) [5]. Although GADD45a has been implied in DNA demethylation in *Xenopus* [14], GADD45a knockout mice did not change the global or locus-specific DNA methylation pattern, indicating that it is not directly functioning in DNA demethylation in mice [15] or that there may be functionally redundant genes. More recently, Xu and colleagues also found that TET proteins oxidize 5mC and 5hmC into 5caC [3]. They provided further evidence that TDG also exhibits strong glycosylase activity on 5caC:G mismatches in double-stranded DNA. Ectopic expression of TDG or knockdown of TDG diminished or increased the amount of 5caC, respectively [3]. However, other DNA glycosylases such as the single-strand-selective monofunctional uracil DNA glycosylase 1 (SMUG1) and the uracil-DNA glycosylase or Methyl-CpG-binding domain protein 4 (MBD4) did not exhibit 5caC excision activity [3]. These studies indicate that 5hmC is a key intermediate that can be deaminated to 5hmU by AID/APOBEC deaminases, or further oxidized to 5caC by TET proteins. TDG is responsible for removing 5hmU or 5caC, leaving a gap that is subsequently filled by unmethylated cytosine through the BER pathway (Figure 1).

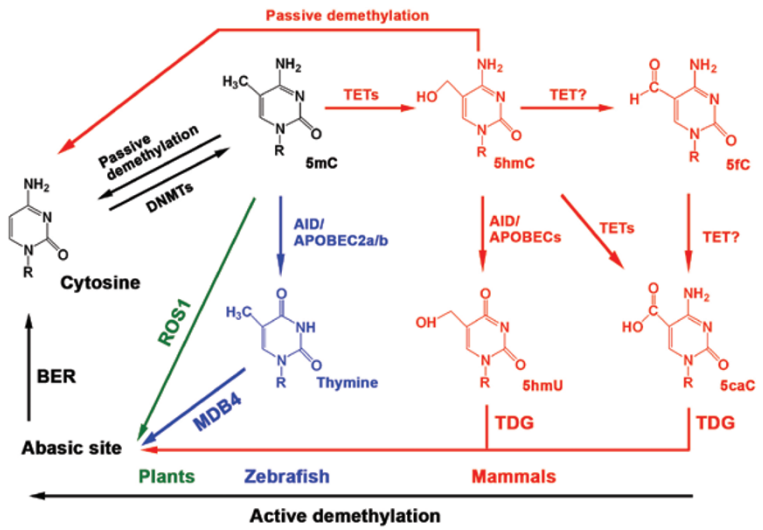

Figure 1 DNA demethylation in animals and plants. Cytosine methylation (5mC) is catalyzed by DNA methyltransferases (DNMTs). In mammals (red color), 5mC can be hydroxylated by TETs to become 5-hydroxymethylcytosine (5hmC), which can be further oxidized to produce 5fC and 5caC. 5hmC can be deaminated to 5hmU by AID/APOBEC deaminases. 5hmU or 5caC is removed by the TDG DNA glycosylase. In zebrafish (blue color), 5mC is deaminated to thymine (T) by AID/APOBEC2a/b, and then the T is removed by the MDB4 DNA glycosylase. In plants (green color), the ROS1 subfamily of DNA glycosylases directly removes 5mC. The gap can be refilled by unmethylated cytosine through the BER pathway. In mammals, 5hmC is a key intermediate for oxidative DNA demethylation. In addition, 5hmC plays important roles in passive DNA demethylation and gene regulation.

Because 5hmC modification is relatively stable during important developmental stages, researchers have begun to investigate whether this modification affects gene expression. One prediction is that TET proteins may activate the expression of their targeted genes via their role in active DNA demethylation. However, genome-wide analysis of 5hmC distribution reveals that TET1 has dual functions in transcriptional regulation because it both enhances and inhibits the expression of some genes in mESCs [16-19]. TET1 is highly expressed in mESCs that accumulate relatively high levels of 5hmC, while 5hmC levels are significantly decreased after mESC differentiation [7]. Downregulation of TET proteins by RNAi decreases the accumulation of 5hmC but increases 5mC. These results suggest a dynamic regulation of 5hmC by TET proteins [16-19]. Using various chromatin immunoprecipitation methods combined with high-throughput DNA sequencing, several laboratories analyzed the genome-wide distribution of 5hmC and TET1 binding sites [16-19]. Interestingly, TET1 is preferentially bound to CpG-rich regions at transcriptional start sites in promoters of both transcriptionally active and Polycomb-targeted repressive genes and also within exons [16-18]. TET1 depletion increases the DNA methylation levels at many TET1 binding sites but only changes the expression of a small number of targeted genes [17]. TET1 physically interacts with the SIN3A co-repressor complex that mediates histone deacetylation for transcriptional repression [18]. In addition, TET1 binding sites largely overlap with the targets of Polycomb repressive complex 2 (PRC2), and TET1 knockdown cells impaired the binding of Ezh2, a core subunit of PRC2, to the PRC2 binding sites, indicating that

TET1 plays a role in recruiting PRC2 [17]. Consistent with a role for TET1 in gene regulation, TET1 binding sites are enriched with the repressive chromatin marker H3K27me3, the active marker H3K4me3, or both [16, 17, 19]. TET1 protein contains a CXXC motif that can bind to unmodified C- or 5mC- and 5hmC-modified CpG-rich DNA [16]. In addition to contributing to active DNA demethylation in mammals, TET1 may also facilitate passive DNA demethylation because the binding of TET1 to the target sites would inhibit the recruitment of DNMTs to methylate these CpG sites. It is also likely that the DNMTs cannot maintain DNA methylation during DNA replication because DNMTs do not recognize 5hmC, which would result in passive DNA demethylation. These results suggest that, in mammals, TET proteins play an important role in both active and passive DNA demethylation and that TET proteins can repress gene expression independent of their hydroxylase activity in mESCs.

In summary, some 5mC in mammals is oxidized by TETs to 5hmC, which can be deaminated to 5hmU by AID/APOBEC deaminases or further oxidized to 5fC and then to 5caC. The 5hmU or 5caC is then removed by the DNA glycosylase TDG, and the gap is refilled by unmethylated C through the BER pathway (Figure 1). The most recent discovery of 5fC and 5caC and of the 5caC DNA glycosylase activity of TDG [3, 11] not only provides further support of the important role of TETs and BER in active DNA demethylation in mammals, but also suggests increasing complexities of the process of active DNA demethylation (Figure 1). In zebrafish, 5mC was reported to be deaminated to thymine (T) by AID/APOBEC2a/b, and T is subsequently converted to C by the MDB4 glycosylase (a T:G mismatch-specific thymine glycosylase)-mediated BER pathway [2]. In plants, the ROS1 family of DNA glycosylases directly removes 5mC to initiate a BER pathway for active DNA demethylation [1]. These results suggest that BER-based demethylation mechanisms are conserved from plants to animals (Figure 1) [1]. The importance of the newly discovered 5hC and 5caC in mammals requires further studies, especially considering that they occur at very low levels compared to 5mC and 5hmC. It is likely that the presence of multiple oxidative intermediates of 5mC may afford more points of regulation of active DNA demethylation in mammals even before BER starts. It is also possible that 5caC may be converted to C by an as yet unknown decarboxylase enzyme, thus bypassing BER to achieve DNA demethylation [11]. This latter model would be particularly attractive for rapid, global DNA demethylation that occurs in the paternal pronucleus in some mammalian zygotes. Future studies will determine whether these different demethylation pathways may all operate in one cell type in mammals or whether one main mechanism is required for achieving DNA demethylation in a particular cell.

References

1 Zhu JK. Active DNA demethylation mediated by DNA glycosylases. *Annu Rev Genet* 2009; **43**:143-166.
2 Wu SC, Zhang Y. Active DNA demethylation: many roads lead to Rome. *Nat Rev Mol Cell Biol* 2010; **11**:607-620.
3 He YF, Li BZ, Li Z, *et al*. Tet-mediated formation of 5-Carboxylcytosine and its excision by TDG in mammalian DNA. *Science* 2011 Aug 4. doi:10.1126/science.1210944
4 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.
5 Cortellino S, Xu J, Sannai M, *et al*. Thymine DNA glycosylase is essential for active DNA demethylation by linked deamination-base excision repair. *Cell* 2011; **146**:67-79.
6 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**:1129-1133.
7 Tahiliani M, Koh KP, Shen Y, *et al*. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* 2009; **324**:930-935.
8 Wyatt GR, Cohen SS. A new pyrimidine base from bacteriophage nucleic acids. *Nature* 1952; **170**:1072-1073.
9 Penn NW, Suwalski R, O'Riley C, Bojanowski K, Yura R. The presence of 5-hydroxymethylcytosine in animal deoxyribonucleic acid. *Biochem J* 1972; **126**:781-790.
10 Kriaucionis S, Heintz N. The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. *Science* 2009; **324**:929-930.
11 Ito S, Shen L, Dai Q, *et al*. Tet proteins can convert 5-Methylcytosine to 5-Formylcytosine and 5-Carboxylcytosine. *Science* 2011 Jul 21. doi:10.1126/science.1210597
12 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.
13 Popp C, Dean W, Feng S, *et al*. Genome-wide erasure of DNA methylation in mouse primordial germ cells is affected by AID deficiency. *Nature* 2010; **463**:1101-1105.
14 Barreto G, Schafer A, Marhold J, *et al*. Gadd45a promotes epigenetic gene activation by repair-mediated DNA demethylation. *Nature* 2007; **445**:671-675.
15 Engel N, Tront JS, Erinle T, *et al*. Conserved DNA methylation in Gadd45a(-/-) mice. *Epigenetics* 2009; **4**:98-99.
16 Xu Y, Wu F, Tan L, *et al*. Genome-wide regulation of 5hmC, 5mC, and gene expression by Tet1 hydroxylase in mouse embryonic stem cells. *Mol Cell* 2011; **42**:451-464.
17 Wu H, D'Alessio AC, Ito S, *et al*. Dual functions of Tet1 in transcriptional regulation in mouse embryonic stem cells. *Nature* 2011; **473**:389-393.
18 Williams K, Christensen J, Pedersen MT, *et al*. TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity. *Nature* 2011; **473**:343-348.
19 Pastor WA, Pape UJ, Huang Y, *et al*. Genome-wide mapping of 5-hydroxymethylcytosine in embryonic stem cells. *Nature* 2011; **473**:394-397.