

**TET proteins and their role in regulation of DNA methylation.**Teka Obsa Feyisa<sup>1\*</sup>, Getahun Chala<sup>2</sup>, Tefera Belsty<sup>3</sup><sup>1</sup>Department of Medical Biochemistry, College of Health and Medical Sciences, Haramaya University, Harar, Ethiopia<sup>2</sup>Department of Medical Physiology, college of Health and Medical sciences, Haramaya University, Harar, Ethiopia<sup>3</sup>Department of Human Anatomy, College of Health and Medical science, Haramaya University, Harar, Ethiopia**Abstract**

Ten-Eleven Translocation (TET) proteins are enzymes that are capable oxidizing of 5-methylcytosine (5mC) into three biologically essential molecules: 5-carboxylcytosine (5caC), 5-formyl cytosine (5fC), and 5-hydroxymethyl cytosine (5hmC). In mammalian genes, TET proteins are categorized into three main members (TET1, TET2 and TET3), which basically contain a C-terminal catalytically active domain and a central region enriched with cysteine domain. Before the finding of TET proteins, DNA methylation was supposed to be irreversible epigenetic change. The discovery of the TET protein later revealed the DNA methylation process to be a reversible covalent modification catalyzed by DNA methyltransferase enzymes. In addition to regulating DNA methylation pathways, TET proteins play significant roles in controlling gene expression, epigenetic remodeling in stem cell segregation, embryogenesis, growth, and cancer. Beside the gene expression, the activities TET proteins can be altered by an array of cellular processes such as interaction with chromatin protein and other different small molecules. Here, we reviewed the current concepts regarding the structure of TET proteins, their role in DNA methylation, gene regulation, and their mutation which leads to development of certain cancer.

**Key words:** DNA Methylation, Epigenetic modification, TET proteins, 5-methylcytosine.

Accepted on 02 August, 2021

**Introduction**

Ten-eleven translocation (TET) proteins that are capable oxidizing of oxidizing 5-methylcytosine (5mC) into three biologically essential molecules: 5-hydroxymethyl cytosine (5hmC), 5-formyl cytosine (5fC) and 5-carboxylcytosine (5caC). The TET proteins were originally discovered and named as a result of the unusual ten–eleven translocation (t (10, 11), (q22; q23)) detected in acute myeloid and lymphocytic leukemia cases where by the mixed-lineage leukemia 1 (MLL1) gene located on chromosome 10 is fused with the TET1 gene positioned on chromosome 11 in humans. Three TET genes (TET1, TET2 and TET3) that have been recognized in mammals encode for proteins capable of catalyzing the oxidation of 5mC in the DNA demethylation pathways.

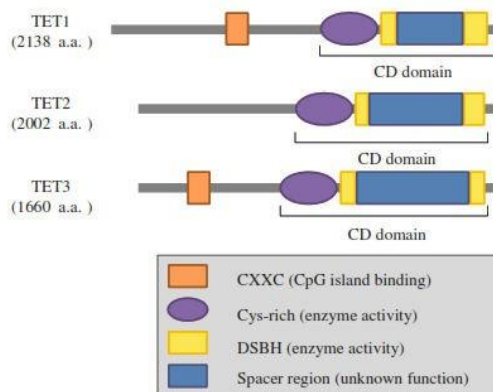
DNA methylation is defined as the covalent modification of the nitrogenous base cytosine by addition of a methyl group (-CH<sub>3</sub>) to the fifth position of 5-CpG-3 dinucleotide linkage. This covalent modification of DNA is principally recognized at CpG rich region in the mammalian genomes. The addition of -CH<sub>3</sub> to non-CpG dinucleotides is atypical, except for stem cells and mature neuron cells. DNA methylation is one of the well studied epigenetic events catalyzed by DNA methyltransferases (DNMTs) and has crucial role in normal development. Several scientific studies were conducted on the TET proteins and their role in the regulation of DNA methylation, but a

comprehensive review is limited. Thus, this review was aimed at presenting critically organized and comprehensive concepts on the role of TET proteins in the regulation of DNA methylation [1].

**The structure of TET Protein Members**

As it has been mentioned by different studies mammalian genes contain three main members of TET proteins: TET1 located on 10q21.3, TET2 located on 4q24 and TET3 located on 2p13.1 of human chromosome. The results of many studies showed catalytically active C-terminal domains, cysteine rich regions, are found in all TET proteins, as well as a double-stranded  $\beta$ -helix (DSBH) fold distinctive of the ferrous iron (Fe (II)) and 2-oxoglutarate dependent deoxygenate super family. These enzymes utilize Fe (II) that serves as a cofactor and 2-oxoglutarate as a cosubstrate for the enhancement their catalytic activities. The three TET genes identified in mammals have been proved to encode for proteins that are able to oxidize 5mC as a part of DNA demethylation pathway.

In the center of TET protein genes, exactly before the double-stranded  $\beta$ -helix, a cysteine rich region with unidentified function exists. The N-terminal area of TET1 and TET3 maintains the evolutionarily preserved CXXC-type domain which helps as a vital recognition in binding to unmethylated cytosine moieties of DNA. Unlike TET1 and TET3, TET2 does not contain a CXXC-type domain (Figure 1).



**Figure 1.** Structure of TET family proteins. The C-terminal central part catalytic domain found in all TET proteins consists of the DSBH, a cysteine-rich domain. The DSBH domain has a huge low complexity part varying in length along with family of TET proteins. TET1 and TET3 contain N-terminal CXXC domain that can bind directly to unmethylated region of DNA and assist in marking genomic target site.

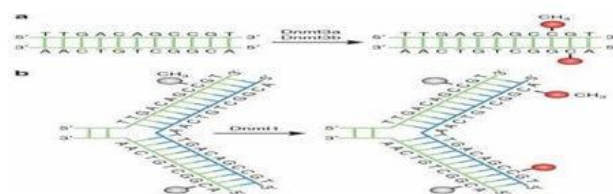
## DNA Methylation

Epigenetic is defined as the study of a gene function which can be passed from one cell to the next cell generation without an involvement of alteration in the DNA base sequence. DNA methylation and histone modification are the two major mechanisms that represent epigenetic events. Our review principally focuses on the mechanism of DNA methylation, which catalyzed by a family of three DNA methyltransferase enzymes (DNMTs): DNMT1, DNMT3A and DNMT3B. DNA methyltransferases 1, the primary enzyme, is considered a continuance methyltransferase, for the reason that it is responsible for conveying methylation pattern from one cellular generation to the next. Human genomes encompass comparatively equivalent proportions of the four DNA nitrogenous bases (A, G, C, and T). In addition to their original states, some of these bases can also be undergo reversible modifications; the most known one is being addition of -CH<sub>3</sub> to cytosine base [2].

Cytosine methylation in DNA is a common covalent modification wherein a methyl group (-CH<sub>3</sub>) is added to the aromatic ring of a cytosine base (5-mC) within a CpG dinucleotide identified by DNA methyltransferase enzymes. The unmethylated CpG dinucleotides are highly concentrated in CpG rich regions, known as CpG islands (CGIs), sequences of consisting of about 1 kb in length. In the CpG islands, the CG content accounts for greater than 55% that is typically associated to the promoter regions and to the first non coding regions (exons) of about 60% of genes.

CpG islands are evolutionarily conserved region of DNA that enhance gene expression by controlling the chromatin arrangement and transcription factor binding, thus, methylation of CpGIs is able to impair binding of transcription factor,

recruit repressive methyl-binding proteins, and firmly silence gene expression. The mechanism responsible for CGIs hypomethylation throughout the period of overall de novo methylation at early development remains unclear yet. However, some possible mechanisms were proposed which involve: (i) CGIs remain hypomethylated is related to intrinsic sequence property which can exclude the activity or association of DNA methyltransferases. (ii) CGIs attain DNA methylation generally, but are targeted by a demethylating action. (iii) The basic transcriptional apparatus (RNA Pol II and TF) and histone H3 lysine 4 trimethylation (H3K4me3) prohibits the DNMTs from transcriptional initiation sites. All of the three identified DNA methyltransferases are proteins with multidomains encompassing a large N- terminal region and a smaller C-terminal domain harboring the catalytic core. DNMT1 shows a partiality for hemi-methylated DNA more than unmethylated in human cells. Methylation occurs at cytosine residue that precedes a guanosine in the CpG dinucleotide of DNA. Unmethylated CpGs are not arbitrarily distributed within the DNA sequence, however, are often gather together in 'CpG islands that are in the promoter region of several genes (Figure 2).

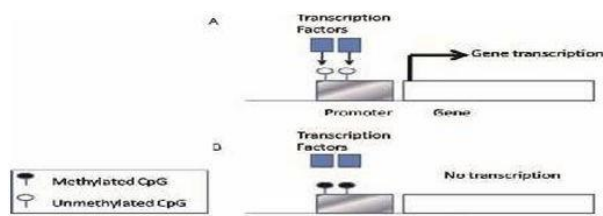


**Figure 2.** DNA methylation pathways: (a) DNMT3a and DNMT3b represent the de novo DNMTs and transfer methyl groups (red) from methyl group donor, S-adenyl methionine (SAM) onto the 5' carbon of cytosine residue of bare DNA to form 5mC. (b) DNMT1 is preservative DNMT and maintaining DNA methylation pattern at DNA replication level. When the DNA molecule undergoes semi-conservative replication, the parental DNA template maintains the original DNA methylation model (gray).

## DNA Methylation and control of Gene Expression

DNA methylation is a steady process; however, reversible covalent modifications that may change the arrangement of chromatin structure and gene regulation. Various organisms use this type of DNA modifications to regulate transcriptional process. The status of DNA methylation affects several biological processes during mammalian developmental stages and is well-known to be enormously aberrant in cancerous cells. DNA methylation influences activity of gene directly through inhibition of transcription factors binding and indirectly by employing chromatin associated proteins with suppressive properties. In addition, it contributes to cell lineage restriction and genetic imprinting, for instance silencing of the X-chromosome in female mammals, during developmental period [3].

Methylation of 5'-carbon of cytosine in the CpG dinucleotide sites is an evolutionarily conserved epigenetic modification which partakes in regulation of gene expression without alterations in basic DNA sequences. For example, the loss of DNA-methyltransferase1 activity, well demonstrated event in knockout mice, results in the abnormal differentiation of hematopoietic stem cells (HSCs) into myeloid progenitor cells (MPCs) (Figure 3).



**Figure 3.** DNA methylation and regulation of gene expression. (A) Unmethylated CpG island promoter lets binding of transcriptional factors, which is necessary for initiation of transcription. (B) The methylated CpG island promoter keeps away transcriptional factors from binding of and leads to gene silencing.

### The Role of TET Proteins in Regulation of DNA Methylation

Until the finding of TET1, DNA methylation was supposed to be an irreversible epigenetic change associated with gene repression, which might merely be assuaged through DNA replication. The discovery of the Ten-Eleven Translocations (TET), hydroxylases, that catalyzed DNA demethylation has basically changed the conception of how DNA methylation is regulated. Consequently, the discovery of TET family proteins capable of oxidizing 5mC to 5hmC, 5fC and 5caC in DNA molecule has lead to deep progress in perception of the basic mechanism underlying DNA demethylation. The TET proteins, DNA demethylases, actively downregulate the DNA methylation level. On the one hand, TET proteins regulate DNA methylation through binding to CpG-rich regions of DNA to prevent unnecessary DNA methyltransferase activity; on the other hand, they regulate DNA methylation by changing 5mC to 5hmC via their hydroxylase activities [4].

Several studies have suggested that there might be multiple mechanisms by which 5hmC and TET proteins regulate DNA methylation and gene expression processes. The conversion of 5mC to 5hmC by TET proteins, leads to failure of DNA methylation during cell division. The derivative of 5mC that are produced by TET proteins are probable to serve like intermediates in DNA demethylation pathway. So far, two principal DNA demethylation processes have been proposed: The passive and active DNA demethylation pathways.

### The Passive DNA Demethylation Pathway

The passive DNA demethylation pathway refers to the failure to retain DNA methylation pattern across cell divisions, and is supposed to be the outcome of replication dependent 5-methylcytosine dilution. Beside downregulating of the DNA

methylation machinery, the deposition of 5-methyl cytosine can have a function in provoking passive DNA demethylation. The passive loss of 5mC can hence be accomplished through consecutive cycles of DNA replication in the absence of functional DNMT1/UHRF1, passive dilution of 5mC. The passive demethylation most probably occurs during mammalian developmental stage and cell differentiation, mostly in the maternal genome at some stages of preimplantation growth.

Conversion of 5-methylcytosine to 5-hydroxymethylcytosine might promote passive demethylation as a result of replication by inhibiting the activity of DNMT1 to distinguish 5hmCpGs region as observed *in vitro* experiments. The TET-mediate 5hmC deposition may thus elicit passive replication dependent DNA demethylation on the opposite DNA template and be essential to counteract irregular assembly of atypically methylated pattern. However, the compel DNMT1 collaborator UHRF1 has been reported to bind 5-mC and 5-hmC with comparable affinities, suggesting that further investigation is required.

### The Active DNA Demethylation

Active DNA demethylation is defined as an enzymatic reaction in which 5-mC base, in its oxidized form, is substituted by unmodified cytosine in a replication- independent mode. In the active DNA demethylation pathway, multienzymes including TET1, TET2 and TET3, which catalyze the conversion of 5mC to 5hmC, thymine DNA glycosylase (TDG), and methyl-CpG-binding domain protein 4 (MBD4), the RNA/DNA editing enzyme activation induced cytidine deaminase (AID), and growth restrain and DNA damage 45 (GADD45a, b and c), which employs nucleotide and /or base excision repair (BER) factors to gene specific loci and serve as adapter molecules between the repair factors and chromatin are involved.

A number of mechanisms regarding to active DNA demethylation pathways have been hypothesized; however, base excision and DNA base excision repair (BER) catalyzed by thymine-DNA-glycosylase (TDG) are currently assumed to be the major imposer. Thymine DNA glycosylase is a DNA mismatch repair enzyme which binds and excises mispaired pyrimidines in G:U and G:T base pairs. The DNA mismatch repair enzyme, TDG, was initially identified as recognizing and repairing T:G mismatches in DNA base pairing. However, TDG also capably excises 5fC and 5caC that are correctly paired to G in double strand DNA, and depletion of TDG leads to increased levels of 5fC and 5caC at a particular gene location.

Biochemically, the TET catalyzed DNA demethylation pathway should be irreversible, because no replication dependent means is found to resynthesize 5mC oxidation products. However, note that, the levels of both 5hmC and 5mC are comparatively stable in normal mammalian tissues. Most evidences exist for mechanisms by which TET proteins convert 5mC into 5hmC, 5fC and 5caC through three successive oxidation reactions. Furthermore, evidence exists for a mechanism by which AID/APOBEC proteins deaminate

5hmC to 5hmU followed by TDG mediated base excision repairs.

### Expression TET Proteins in Different Tissues

All TET family members possess the ability to oxidize 5mC to 5hmC, and 5fC5ac, however, they are not equally expressed in different cells or tissues. For instance, the expression of TET1 and TET2 is mostly occurred in mouse Embryonic Stem Cell (ESC), while TET3 is mostly expressed in oocytes and one cell zygotes. Even though they were recognized only a few years ago, increasing evidence indicates that TET proteins have critical roles in epigenetic modification of stem cell differentiation, development of embryo, growth, and tumor cell formation.

On the contrary to the association of 5mC oxidation with regional restoration of an unmethylated form during reprogramming, in somatic cells like stem cells and neuron cells, 5mC oxidation usually appears to contribute to the continuance of methylation fidelity by counteracting aberrant de novo methylation process. In embryonic stem cells, TET1 keeps CpG-rich promotes hypomethylated by getting rid of erroneously placed methylation. Existing evidence suggested that oxidation of 5-mC in the paternal genome in fertilized eggs by TET3 starts DNA demethylation and stimulates the activation of the paternal replica in the early embryonic genes [5].

### Physiological Functions of TET Proteins

The physiological importances of the TET proteins have been investigated using genetic knockout mouse models. Constituent deletion of TET3 results in neonatal lethality with 100% penetrance, while the deletions of either TET1 or TET2 do not result in any noticeable developmental phenotype. This implies that TET3 has an exceptional role during embryogenesis that cannot be remunerated for by other TET enzyme family.

### Mechanisms of TET protein Regulation

The demonstration revealed that TET proteins are subjected to diversity of post-translational modifications that typically occur at specific regulatory regions (N- terminus and low-complexity regions found between the domains of dioxygenases). Their results provided a new possible means for TET protein regulation based on a dynamic interchange of phosphorylation and O-linked GlcNAcylation at specific regulatory regions. Away from gene expression, the activity of TET is influenced by a variety of cellular processes, including alteration in levels of Reactive Oxygen Species (ROS), Krebs cycle intermediates (succinate and fumarate) compete with TET for its cofactor  $\alpha$  - ketoglutarate, and proliferation.

Many proteins including components of chromatin, subcellular localization and chromatin-binding interact with TET proteins and involved in the regulation of their activities. Additionally, many other molecules were also identified to be involved in the regulation of TET enzyme activity. Ascorbic acid (vitamin C) and ATP, which not directly participate in the hydroxylation of

5hmC are the two most common molecules demonstrated both *in vitro* and cultured cell to have triggering effect on enzymatic activity of TET proteins. However, in 2013, Dr. Yin confirmed that ascorbic acid can directly involved in the regulation of TET proteins by enhancing their activity through unclear mechanism.

### TET Proteins and Cancer

DNA methylation pathways has appeared to be an important epigenetic event that plays crucial roles in cell growth, apoptosis, aging and diseases such as cancer. The elevated rate of mutagenesis of 5mC compared to other nucleotides has been broadly documented in the literatures. The mutation rate of CpG dinucleotides has been predicted to be about 10 to 40-fold compared to other dinucleotide mutation rate. Many research results have also showed 5hmC as a potential biomarker whose reduction is associated with tumor development and spread. Changes in DNA methylation model have been reported in several malignant cells and revealed to play a vital role in cancer differentiation and progression processes. The balance between DNA methylation and demethylation pathways is controlled by many proteins and cofactors. In the case of cancer, this balance is recurrently deregulated and leads to altered DNA methylation patterns, which in turn results in the repression of tumor suppressor gene or activation oncogene. As the loss of 5mC is a feature of nearly all cancer types, active DNA demethylation pathway plays a significant role in the preservation of DNA methylation patterns, and perhaps, of other epigenetic change mark.

The TET proteins whose name was first derived from the TET1 characterization as a fusing partner in Mixed-Lineage Leukemia (MLL), which has been identified in both lymphoid and myeloid cancer cases. The most common cytogenetic findings of Acute Myeloid Leukemia (AML) showed aberrations of chromosomal gene involving the MLL at 11q23. The (t (q22; q23)) has been frequently reported in several cases of AML. In 2009, TET2 aberration was found in various hematological malignancies. The TET2 gene was found to be affected by deletion, splice site, and point mutations such as missense and nonsense. In 2013, Asmar and his colleagues found that TET2 mutations in 12% of diffuse large B-cell lymphoma (DLBCL), with 5% carrying missense and 7% carrying nonsense, frame-shift or splice-site mutations with a mutation detecting analysis focused on TET2 gene.

The gain of function of oncogene in IDH mutation is an initiation event that identifies main clinical and prognostic classifications of gliomas. Mutant IDH proteins produce a new oncometabolite, 2-hydroxyglutarate that competes with  $\alpha$ -ketoglutarate for active site of the DNA modifying enzymes of the TET family and interrupts TET function.

### Conclusion

In summary, the basic concepts regarding to the structure and the regulation role of the TET proteins in DNA methylation as well as in development of cancers have been covered in this review article. Several studies revealed that TET proteins play

a crucial regulatory role in DNA methylation. The TET proteins, DNA methyltransferases, catalyze DNA methylation through reversible covalent biochemical modification by addition of a methyl group (-CH<sub>3</sub>) to the aromatic ring of cytosine residues (5mC) within a CpG dinucleotides. Nowadays, the discovery of the TET proteins has dramatically changed the assumption that belief the DNA methylation to be an irreversible epigenetic event associated with gene suppression and revealed the mechanisms by which they regulate DNA methylation. In addition regulation of DNA methylation, recent findings showed that TET proteins play crucial roles in epigenetic remodeling in stem cell differentiation, embryogenesis and control of cancers. In general, since TET proteins are the key regulators of DNA demethylation pathways and gene expression, the mutation in the TET protein genes or the dysregulation of their activities or alteration in their expression levels can lead to the development of different cancers. Therefore, the future research findings should focus on how to target TET proteins as potential biomarkers of tumor cells at early development in order to control cancers that are developed as a result of TET proteins mutation.

## References

1. Shen L, Zhang Y. Enzymatic Analysis of Tet Proteins: Key Enzymes in the Metabolism of DNA Methylation. *Methods in Enzymol.* 2012;512:93-105.
2. Xu GL, Walsh CP. Enzymatic DNA oxidation: mechanisms and biological significance. *BMB Rep.* 2014;47:609-18.
3. An J, Pastor WA. TET proteins and 5-methylcytosine oxidation in hematological cancers. *Immunol Rev.* 2015;263:6-21.
4. Delatte B, Fuks F. TET proteins: on the frenetic hunt for new cytosine modifications. *Brief Funct Genomics.* 2013;12:191–204.
5. Pastor WA, Aravind L, Rao A. TETonic shift: biological roles of TET proteins in DNA demethylation and transcription. *Nat Rev Mol Cell Biol.* 2014;14:341-56.

## \*Correspondence to

**Teka Obsa Feyisa**

Department of Medical Biochemistry,  
College of Health and Medical Sciences,  
Haramaya University,  
Harar,  
Ethiopia  
Email: takobsi2008@gmail.com