

Establishment, Erasure and Synthetic Reprogramming of DNA Methylation in Mammalian Cells



Renata Z. Jurkowska and Tomasz P. Jurkowski

Contents

1	Introduction	2
2	Setting of DNA Methylation	2
2.1	Architecture of DNA Methyltransferases	4
2.2	Catalytic Properties of DNMTs	7
2.3	Intrinsic DNA Sequence Specificity of DNMTs	8
2.4	Recruitment of DNMT Enzymes to Chromatin and Replicating DNA	9
3	Erasure of DNA Methylation	12
3.1	Architecture of TET Enzymes	12
3.2	Intrinsic DNA Sequence Specificity of TET Enzymes	13
3.3	On Site and Lateral Processivity of TET Enzymes	14
3.4	Oxidation of RNA Bases	15
3.5	Recruitment of TET Enzymes	15
4	Synthetic Programming of DNA Methylation	16
4.1	Programmable Genome Targeting Modules	17
4.2	Epigenetic Effector Domains	17
4.3	Applications of Targeted DNA Methylation/Demethylation	18
4.4	Limitations of Targeted DNA Methylation and Demethylation Tools	19
	References	20

Abstract DNA methylation is a crucial epigenetic modification involved in the control of cellular function and the balance between generation of DNA methylation and its removal is important for human health. This chapter focuses on the enzymatic machinery responsible for the processes of establishment, maintenance and removal of DNA methylation patterns in mammals. We describe the biochemical, structural and enzymatic properties of DNA methyltransferases and TET DNA hydroxylases,

R. Z. Jurkowska (✉)
BioMed X Innovation Center, Heidelberg, Germany
e-mail: jurkowska@bio.mx

T. P. Jurkowski (✉)
Abteilung Biochemie, Institut für Biochemie und Technische Biochemie, Universität Stuttgart,
Stuttgart, Germany
e-mail: tomasz.jurkowski@ibc.uni-stuttgart.de

as well as their regulation in cells. We discuss how these enzymes are recruited to specific genomic loci, and how their chromatin interactions, as well as their intrinsic sequence specificities and molecular mechanisms contribute to the methylation pattern of the cell. Finally, we introduce the concept of epigenetic (re)programming, in which designer epigenetic editing tools consisting of a DNA targeting domain fused to an epigenetic editor domain can be used to edit the epigenetic state of a given locus in the genome in order to dissect the functional role of DNA methylation and demethylation. We discuss the promises of this emerging technology for studying epigenetic processes in cells and for engineering of cellular states.

Keywords DNA methylation · DNA demethylation · Synthetic epigenetics · TET · DNMT · dCas9 · Epigenetic editing

1 Introduction

Methylation of CpG sites in the DNA plays a vital role in mammalian development and has been studied extensively in the past decades. However, despite advances in understanding of the targeting and regulation of DNA methyltransferases in cells, the specific processes contributing to the generation, maintenance and erasure of DNA methylation patterns are not yet fully elucidated, and the exact molecular mechanisms leading to the aberrant methylation observed in human disease (like cancer) are only partially understood. The discovery of TET enzymes has changed the view on DNA methylation as a very stable modification, as it showed that active DNA demethylation can occur through stepwise oxidation of 5-methylcytosine (5mC) to 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC) and finally to 5-carboxylcytosine (5caC), followed by the removal of the oxidised bases by Thymine DNA glycosylase (TDG) and base excision repair mechanism. However, while the biological function of TET enzymes has been studied quite extensively, very little is known about their biochemical properties and their specificity, catalytic mechanism, as well as the contribution of different domains to enzymes targeting and regulation. In this chapter, we summarise the most important properties of both DNA methyltransferases and TET enzymes and describe some of the molecular pathways leading to their recruitment to the target sites. Finally, we introduce the concept of epigenetic editing as an elegant approach to dissect the function of DNA methylation and demethylation in a locus specific manner.

2 Setting of DNA Methylation

Since the discovery of methylated DNA bases in 1948, major advances have been made in our understanding of the biological role of DNA methylation, as well as the mechanisms regulating the function of DNA methyltransferases in cells. Through these discoveries, DNA methyltransferases emerged as key epigenetic enzymes

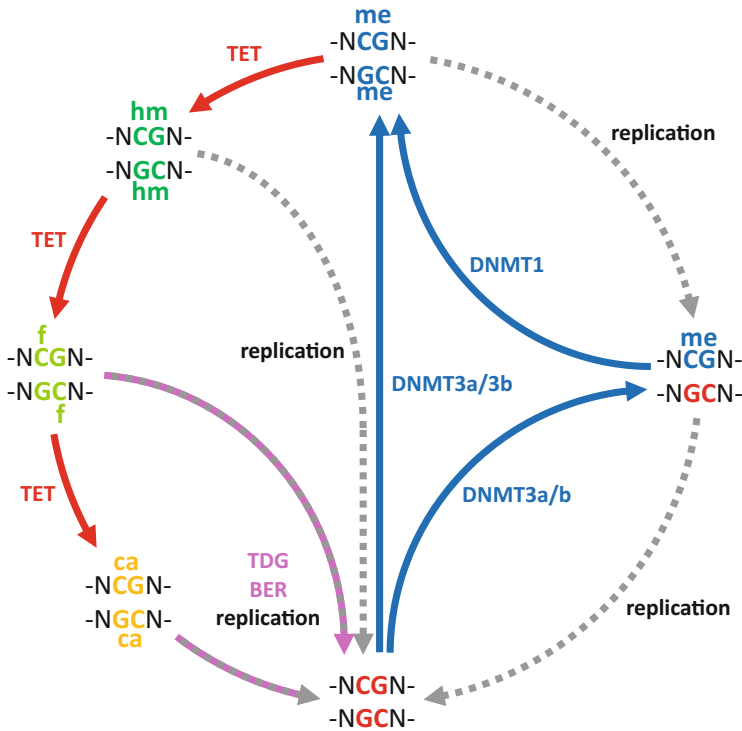


Fig. 1 The methylation cycle in mammals. DNA methylation is set on unmethylated cytosines by the combined action of DNMT3A and DNMT3B (blue) and maintained during replication by DNMT1, which has high preference for hemimethylated CpG sites. DNA methylation can be lost by a passive mechanism, getting diluted after consecutive cycles of DNA replication or through an active mechanism, involving oxidation of 5mC to 5hmC, 5fC and 5caC by TET enzymes (red). 5fC and 5caC can be recognised and excised by TDG and base excision repair enzymes, leading to the restoration of an unmethylated state (pink)

regulating mammalian development and cellular specialisation, which is clearly emphasised by the lethal phenotypes of the genetic knockouts of any of the DNA methyltransferase enzymes in mice and by the ever-growing number of diseases showing disturbed DNA methylation signatures.

DNA methylation in mammals occurs at the C5 position of the cytosine residues, primarily in the CpG sites, although non-CpG methylation is also present (albeit at lower levels). About 60–80% of the CpG sites in the human genome (corresponding to around 3–4% of all cytosines) are methylated in a tissue and cell type-specific pattern [reviewed in Schubeler (2015), Jurkowska et al. (2011a), Gowher and Jeltsch (2018), Ravichandran et al. (2018)]. Additionally, 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC), which arise from the step-wise oxidation of the methyl group of 5-methylcytosine, have recently been discovered in mammalian genome DNA (Fig. 1) (Tahiliani et al. 2009; He et al. 2011; Ito et al. 2011). Three methyltransferase enzymes are responsible for the generation

and maintenance of the global DNA methylation patterns in humans: DNMT3A, DNMT3B and DNMT1. DNMT3A and DNMT3B introduce DNA methylation during mammalian development and maturation of germ cells, with the assistance of a regulatory factor DNMT3L. After their establishment, DNA methylation patterns are largely preserved, with only small tissue-specific changes, but can get significantly altered in disease. During DNA replication, unmethylated DNA strands are synthesised, leading to the conversion of fully methylated CpG sites into hemimethylated sites that are then re-methylated by a maintenance methyltransferase DNMT1, which has high preference towards hemimethylated DNA (Fig. 1), and is ubiquitously and highly expressed in proliferating cells. This elegant inheritance mechanism enables DNA methylation function as a key epigenetic mark mediating long-term transcriptional silencing. In this respect, DNA methylation is involved in silencing of repetitive elements, genomic imprinting, X-chromosome inactivation and regulation of gene expression during development and cellular specialisation [reviewed in Smith and Meissner (2013), Bogdanovic and Lister (2017)]. DNA methylation can be lost by either passive mechanism, when maintenance MTase activity is absent, or via an active demethylation process (see below). Considering its important biological roles, it is not surprising that aberrant DNA methylation changes play a prominent role in the development of human diseases, including for example haematological cancer (Bergman and Cedar 2013; Yang et al. 2015).

2.1 Architecture of DNA Methyltransferases

In the structure of the mammalian DNA methyltransferases, two functional parts can be identified, a large N-terminal part and a smaller C-terminal catalytic part (Fig. 2). The N-terminal part of DNMTs contains several distinct domains with targeting and regulatory functions. The C-terminal domain contains ten catalytic amino acid motifs conserved among prokaryotic and eukaryotic C5-DNA methyltransferases and folds into a conserved structure called AdoMet-dependent MTase fold, which consists of a mixed seven-stranded beta sheet, formed by six parallel beta strands and a seventh strand in an antiparallel orientation, inserted into the sheet between strands 5 and 6. Six helices are folded around the central beta sheet (Cheng and Blumenthal 2008). This domain is involved in binding of the cofactor *S*-adenosyl-*L*-methionine (AdoMet), recognition of the DNA substrate and in catalysis. Interestingly, the spatial arrangement of the various domains in DNMTs plays a prominent role in the regulation of enzymes' activity and specificity through allosteric control of the catalytic domain [reviewed in Jeltsch and Jurkowska (2016)].

2.1.1 Domain Composition of DNMT1

DNMT1 was the first mammalian DNA methyltransferase enzyme to be cloned and biochemically characterised. It is a large protein (1616 aa in humans), containing

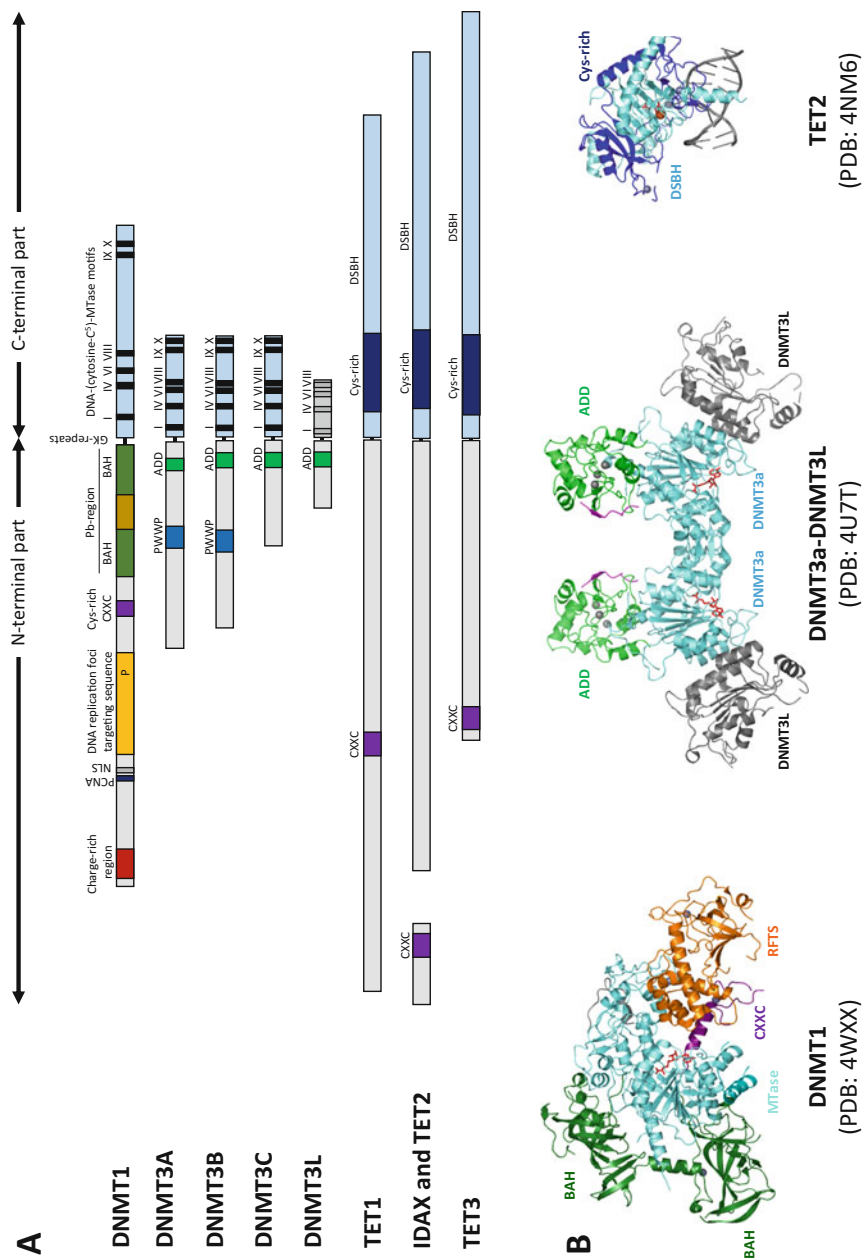


Fig. 2 Domain composition of mammalian DNA methyltransferases and TET enzymes [adapted from Ravichandran et al. (2018)]. (a) The N-terminal regulatory part of the enzymes contains various chromatin and protein interacting domains, in the C-terminal domain of DNA methyltransferases, the catalytic

several distinct domains, which are listed below from the N-terminus to the C-terminus of the protein (Jurkowska and Jeltsch 2016) (Fig. 2):

- DMAPD (DNA methyltransferase-associated protein 1 interaction domain) is involved in the targeting of Dnmt1 to replication foci.
- PBD (PCNA—proliferating cell nuclear antigen—binding domain) recruits DNMT1 to the replication fork during S phase via interaction with PCNA.
- RFTD (replication foci-targeting domain) is involved in the targeting of DNMT1 to replication foci and to centromeric chromatin.
- CXXC domain binds unmethylated DNA and might be involved in the specificity of DNMT1
- BAH1 and BAH2 (bromo-adjacent homology 1 and 2) domains are necessary for the folding of the enzyme, but their exact biological function is unknown

The catalytic domain of DNMT1 is inactive in an isolated form despite presence of all conserved methyltransferase motifs required for catalysis, demonstrating that it is controlled by the N-terminal domain of the enzyme. Indeed, structural and biochemical studies confirmed that several domains in the N-terminal part of DNMT1 directly contact the catalytic domain, providing examples of the sophisticated allosteric regulation of DNMT1 [reviewed in Jeltsch and Jurkowska (2016)].

2.1.2 Domain Composition of DNMT3 Family

Human DNMT3 family comprises three members: DNMT3A (912 aa), DNMT3B (853 aa) and DNMT3L (387 aa). DNMT3A and DNMT3B are enzymatically active, whereas DNMT3L does not possess methyltransferase activity, but it stimulates the activity of DNMT3A and DNMT3B (Gowher et al. 2005; Chedin et al. 2002). Of note, a novel member of the rodent Dnmt3 family, Dnmt3c that arose from duplication of the DNMT3B gene, has been identified recently (Barau et al. 2016). This male germline-specific variant is required for methylation of retrotransposons during mouse spermatogenesis (Barau et al. 2016). However, its orthologue has not been identified in humans.

In the N-terminal part of the DNMT3 proteins, which differ significantly from the N-terminal part of DNMT1, three separate regions can be distinguished (Jurkowska et al. 2011a) (Fig. 2):

- very N-terminal segment of DNMT3A and DNMT3B, which is the most variable region between both proteins, binds DNA and is important for anchoring of the enzymes to nucleosomes. It seems to also play a role in the targeting of DNMT3A to the shores of bivalent CpG promoters.

Fig. 2 (continued) motifs are indicated. For details, refer to text. **(b)** Crystal structures of the human DNMT1 (351–1600 fragment), the human DNMT3A and DNMT3L complex bound to the histone H3 tail and the human TET2 catalytic domain in complex with DNA. Distinct protein domains are color coded according to the color scheme used in **(a)**

- ADD (ATRX-DNMT3-DNMT3L) domain of DNMT3 proteins mediates the interaction with the N-terminal tails of histone H3, as well as other chromatin proteins; and is involved in the allosteric regulation of the enzymes' activity.
- PWWP domain of DNMT3A and DNMT3B interacts with histone H3 tails trimethylated at lysine 36 and is essential for targeting of the enzymes to pericentromeric chromatin and gene bodies. This domain is missing in DNMT3L.

The catalytic domains of DNMT3A, and DNMT3B share ~80% sequence identity and are active in isolated form (Gowher and Jeltsch 2002). In contrast, despite clear homology with the other family members, the C-terminal domain of DNMT3L is catalytically inactive due to amino acid exchanges and deletions within the conserved methyltransferase motifs. Structural studies of the DNMT3A and DNMT3L revealed that both proteins form a tetrameric complex, consisting of two molecules of DNMT3A in the centre and two molecules of DNMT3L at the edges of the tetramer (Jia et al. 2007; Guo et al. 2015; Zhang et al. 2018). These interfaces also support self-interaction of DNMT3A and contribute an interesting regulatory mechanism for the activity and localisation of DNMT3A [reviewed in Jeltsch and Jurkowska (2013)]. The arrangement of the two DNMT3A catalytic sites allows methylation of two adjacent CpG sites in one binding event (Jia et al. 2007; Jurkowska et al. 2008; Zhang et al. 2018). The long-awaited structure of the DNA-bound form of the complex revealed that the DNA binding interface of DNMT3A is formed by a specific loop from the target recognition domain, the catalytic loop and the homodimeric interface of DNMT3A (Zhang et al. 2018).

2.2 Catalytic Properties of DNMTs

All DNA cytosine-C5-methyltransferases share a similar catalytic mechanism, involving conserved amino acid motifs, for the transfer of the methyl group from the cofactor AdoMed to the target cytosine base [reviewed in Jurkowska et al. (2011a)]. Interestingly, they use base flipping to rotate the target cytosine out of the DNA duplex and insert it in the catalytic pocket. As in the case of other DNMTs, base flipping was observed in the structures of DNMT1 and DNMT3A with substrate DNA as well (Song et al. 2012; Zhang et al. 2018).

DNMT1 is a very processive enzyme, capable of methylating multiple CpG sites along the DNA without dissociating from the substrate (Vilkaitis et al. 2005; Hermann et al. 2004). This property fits well to the maintenance role of DNMT1 at the replication fork, as it allows very efficient methylation of the newly synthesized daughter strand before the chromatin is reassembled. The structure of DNMT1 revealed that the enzyme enwraps the DNA, enabling sliding of the protein along the substrate and catalysis of successive methylation reactions (Song et al. 2012).

In contrast to DNMT1, DNMT3A methylates DNA in a distributive manner (Norvil et al. 2018; Gowher and Jeltsch 2002), requiring enzyme dissociation after

each round of methylation. In addition, it cooperatively binds to DNA, forming multimeric protein-DNA filaments reviewed in Jeltsch and Jurkowska (2013). Cooperative binding enables methylation of multiple sites on the same DNA molecule and increases DNMT3A activity (Emperle et al. 2014), leading to an efficient spreading of DNA methylation over a larger region (Stepper et al. 2017). Interestingly, binding of several DNMT3A along the DNA leads to an 8-10 bps periodicity in the methylation pattern (Jia et al. 2007; Jurkowska et al. 2008). In contrast to DNMT3A, DNMT3B is able to processively methylate multiple CpG sites and binds to the DNA in a non-cooperative manner (Norvil et al. 2018; Gowher and Jeltsch 2002), indicating that small sequence differences in the catalytic domains of DNMT3A and DNMT3B have a profound impact on the catalytic properties of these related enzymes.

2.3 Intrinsic DNA Sequence Specificity of DNMTs

For a long time, DNA methylation in mammals was thought to be largely restricted to CpG sites, however, recent studies revealed the presence of non-CpG methylation in several cell types and tissues, both in mouse and in humans [reviewed in He and Ecker (2015)]. The original DNA methylation pattern is set by DNMT3A and DNMT3B, which are classically designated as de novo MTases, as they do not display preference between unmethylated and hemi-methylated DNA. Although both enzymes preferentially methylate CpG dinucleotides, they can also modify cytosines in a non-CpG context, with a preference for CA \gg CT > CC (Gowher and Jeltsch 2001; Zhang et al. 2018). Experiments using DNMT3s knockout in embryonic stem cells or ectopic expression of DNMT3A in cells lacking DNA methylation provided direct evidence that DNMT3 enzymes introduce methylation in non-CpG context also in vivo (Ramsahoye et al. 2000). Conversely, DNMT1 cannot efficiently methylate non-CpG sites, leading to the loss of the non-CpG methylation through cellular division in the absence of DNMT3 enzymes. Therefore, presence of the non-CpG methylation directly reflects DNMT3 enzyme activity in cells. Consistently, methylated non-CpG sites are widespread in cells and tissues, where DNMT3A and DNMT3B are highly expressed (like embryonic stem cells, induced pluripotent cells, oocytes and brain), but absent or present at only marginal levels in most somatic tissues and cells with low expression of these enzymes (Ziller et al. 2011; Varley et al. 2013).

Currently, it is unclear what is the biological role of the non-CpG methylation. It has been viewed as a by-product of the hyperactivity and low specificity of DNMT3 enzymes. However, depending on the experimental system, there is also evidence of its potential role in gene repression or expression. Most insights about the potential biological role of non-CpG methylation came from studies on brain [reviewed in Kinde et al. (2015), Jang et al. (2017)], where non-CpG methylation occurs at high levels and contributes to neuronal maturation and specification of brain cells.

However, further studies are required to elucidate the exact biological function of the non-CpG methylation.

Although DNMT3A and DNMT3B do not seem to have strong sequence specificity beyond CpG dinucleotides, both enzymes are sensitive to the sequences flanking their target cytosines. For example, DNMT3A prefers purine bases at the 5' end of the CpG, whereas pyrimidines are favoured at their 3' end (Handa and Jeltsch 2005; Jurkowska et al. 2011b). Interestingly, experimental flanking sequence preferences of DNMT3s correlate with the methylation level of CpG sites found in the human genome (Handa and Jeltsch 2005), suggesting that the inherent sequence preferences of de novo methyltransferases contribute to the selection of their target regions in the genome.

The mechanistic understanding of the flanking sequence preferences and specificity of DNMT3 enzymes towards CpG sites has long awaited the availability of the structure with bound substrate DNA, which has been obtained recently (Zhang et al. 2018). The structure of the DNMT3A-DNMT3L complex bound to DNA revealed that the guanine base of the target CpG site is accurately recognised by the R836 residue of DNMT3A, mutation of which results in a reduced preference of the enzyme for CpG methylation (Gowher et al. 2006; Zhang et al. 2018). Additionally, several residues directly contact the bases flanking the CpG dinucleotide, explaining the strong flanking sequence preferences of DNMT3A. Notably, in the structure, no protein contact with the cytosine base of the opposite strand was observed, explaining the lack of discrimination of DNMT3 enzymes between unmethylated and hemi-methylated DNA (Zhang et al. 2018).

In contrast to DNMT3 enzymes, DNMT1 shows strong preference towards hemi-methylated DNA over unmethylated substrate (Song et al. 2012; Goyal et al. 2006; Bashtrykov et al. 2012), which enables its function as the methylation copy machine at the replication fork. The structure of DNMT1 bound to hemi-methylated DNA provided molecular explanation for this preference and revealed that the methyl group of the cytosine is recognised by a hydrophobic pocket in the catalytic domain of DNMT1 and that both the 5mC and the corresponding G in the target DNA strand are recognized accurately (Song et al. 2012). This observation also explains the high specificity of DNMT1 towards CpG sites over non-CpG sites mentioned above.

2.4 Recruitment of DNMT Enzymes to Chromatin and Replicating DNA

Correct establishment and maintenance of DNA methylation patterns is crucial for human development and health, therefore mechanisms contributing to these processes have been extensively studied over the past decades. Several synergistic models, including both the inherent specificity of the methyltransferases, as well as the role of other proteins and chromatin modifications, have been proposed to

explain how specific DNA methylation patterns are established [reviewed in Jurkowska et al. (2011a)].

2.4.1 Interaction of DNMT3s with Chromatin Marks

As DNA methylation is embedded in multifaced epigenetic network, direct interaction with specific chromatin marks has been proposed as a general mechanism involved in the recruitment of DNA methyltransferases to specific genomic regions.

The ADD domain, which is present in all DNMT3 proteins, interacts specifically with H3 tails unmethylated at lysine 4, modification (e.g. acetylation or di/trimethylation) of which prevents ADD binding (Noh et al. 2015; Otani et al. 2009; Zhang et al. 2010b). Importantly, binding of H3 tails to the ADD domain also allosterically activates DNMT3A (Guo et al. 2015; Li et al. 2011), thereby stimulating methylation of chromatin-bound DNA by DNMT3A. Since (tri)methylation of H3K4 is associated with active genes, its presence would repel DNA methyltransferases and prevent DNA methylation of active regions. Indeed, a strong inverse genome-wide correlation of DNA methylation and H3K4me3 modification was observed (Hodges et al. 2009; Meissner et al. 2008) and demethylation of K4 of H3 at enhancers of pluripotency genes was required for localization of DNMT3 enzymes in embryonic stem cells (Petell et al. 2016). The crucial role of the ADD domain in the targeting of DNMT3A to chromatin *in vivo* was further confirmed by an elegant study, which showed that engineering of the ADD domain of DNMT3A led to aberrant DNA methylation patterns in cells and disturbed differentiation programs of embryonic cells (Noh et al. 2015). Besides interacting with histone H3 tails, the ADD domain is a platform involved in DNMT3A interaction with other proteins, including transcription factors, histone methyltransferases and other chromatin proteins [reviewed in Ravichandran et al. (2018)]. Importantly, as ADD domain is involved in the allosteric regulation of DNMT3A, interaction with this domain may directly influence DNMT3A activity, as shown for H3 (Guo et al. 2015) and MeCP2 (Rajavelu et al. 2018).

The PWWP domain of DNMT3A and DNMT3B, which specifically recognizes H3 tails tri-methylated at K36 (H3K36me3) (Dhayalan et al. 2010), is the second DNMT3 domain directly contributing to the recruitment of methyltransferases to specific genomic regions, including pericentromeric chromatin and gene bodies. Strong correlation of both H3K36me3 and DNA methylation was observed in the body of active genes and at exon-intron boundaries (Vakoc et al. 2006; Kolasinska-Zwierz et al. 2009; Baubec et al. 2015). The central role of H3K36me3 recognition in targeting of DNA methylation was experimentally confirmed in a variety of cellular systems (Neri et al. 2017; Morselli et al. 2015; Baubec et al. 2015). For example, H3K36me3-dependent intragenic DNA methylation by DNMT3B is crucial to protect gene bodies from cryptic transcription initiation (Neri et al. 2017). Furthermore, a subset of heterochromatic repeats shows strong enrichment in H3K36me3, explaining the role of the DNMT3A PWWP domain in the heterochromatic localization of the enzyme (Ernst et al. 2011).

Besides interacting with histone tails, the PWWP domain of DNMT3 enzymes can also bind DNA (Qiu et al. 2002). A recent model for the methylation of nucleosomal DNA by DNMT3A suggested that the targeting occurs through a specific binding of H3K36me3 by the PWWP domain of DNMT3A, which is followed by an activation of the catalytic domain mediated by the binding of H3 tails to the ADD domain, resulting in the methylation of nearby cytosines by the catalytic domain (Rondelet et al. 2016).

2.4.2 Recruitment of DNMT1 to Replicating Chromatin

Several targeting mechanisms ensure the correct localization of DNMT1 to replicating DNA. First one involves PCNA, a component of the replication machinery that interacts and co-localizes with DNMT1 *in vivo* (Iida et al. 2002), indicating that it might directly recruit the methyltransferase to the replication fork and load it onto DNA. The PCNA-DNMT1 interaction contributes to the efficiency of DNA re-methylation in cells, but it is not essential for this process (Egger et al. 2006). Second factor essential for the recruitment of DNMT1 and the maintenance of DNA methylation patterns in mammals is UHRF1 (Sharif et al. 2007; Bostick et al. 2007). UHRF1 specifically binds to hemi-methylated DNA via its SET and RING-associated (SRA) domain (Hashimoto et al. 2008; Bostick et al. 2007; Arita et al. 2008) and recognizes histone H3 tails methylated at lysine 9 (H3K9me2/me3) via cooperative binding of its tandem Tudor domain (TTD) and its plant homeodomain (PHD) (Rothbart et al. 2012; Nady et al. 2011). The chromatin interactions of UHRF1 are necessary for the recruitment of DNMT1 to replicating chromatin, since UHRF1 mutations preventing histone binding abolished DNA methylation by DNMT1 in cells (Rothbart et al. 2012; Nady et al. 2011). Similarly, UHRF1 knockout in mice results in a genome-wide loss of DNA methylation (Bostick et al. 2007; Sharif et al. 2007). In addition to its role in targeting of DNMT1, UHRF1 was also shown to stimulate the catalytic activity of DNMT1 through a direct interaction (Bashtrykov et al. 2014).

A model of a direct recruitment of DNMT1 by histone marks is also plausible, as the methyltransferase preferentially associates with H3 tails ubiquitinated at K18 and K23 (Qin et al. 2015; Nishiyama et al. 2013). This interaction is mediated by the replication foci-targeting (RFTS) domain of DNMT1 and leads to the recruitment of the enzyme to newly replicated DNA and its simultaneous activation, providing another beautiful example of allosteric regulation of DNMTs. The ubiquitination of the H3 tail is introduced by UHRF1 and is stimulated UHRF1 binding to hemi-methylated DNA (Harrison et al. 2016). Ubiquitinated H3 accumulates during S-phase, leading to the recruitment of DNMT1 to newly replicated DNA (Qin et al. 2015; Nishiyama et al. 2013; Harrison et al. 2016). These data indicate an important additional connection between DNMT1 and UHRF1 chromatin interactions, which is essential for an efficient maintenance of DNA methylation.

3 Erasure of DNA Methylation

Since its discovery, 5mC was considered a very stable modification due to the chemical strength of the C-C bond. Therefore, DNA demethylation was expected only to occur passively, through a replication-dependent dilution in the absence or inhibition of the maintenance methylation machinery (Fig. 1). However, global genome-wide loss of DNA methylation occurring in a DNA replication-independent manner was observed in mouse zygotes (Mayer et al. 2000; Oswald et al. 2000) and during specification of primordial germ cells (Hajkova et al. 2002; Yamazaki et al. 2003), pointing towards existence of an active demethylation machinery. Furthermore, active DNA demethylation has also been observed at specific loci in T cells, neurons and other cells (Bruniquel and Schwartz 2003; Martinowich et al. 2003).

Despite the discovery of biological processes where active DNA demethylation occurs in the absence of DNA replication, the enzymatic machinery responsible for this process in mammals remained enigmatic until 2009, when a group of enzymes called Ten-Eleven Translocation (TET) was shown to oxidize 5mC to 5-hydroxymethylcytosine (5hmC) both *in vitro* and in mouse embryonic stem cells (Tahiliani et al. 2009). Moreover, TET enzymes are able to further oxidize 5hmC to 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) (He et al. 2011; Ito et al. 2011). These oxidized bases are recognized and excised by the thymine DNA glycosylase (TDG) triggering base excision repair pathway (BER) to replace the abasic site by an unmodified cytosine (He et al. 2011; Maiti and Drohat 2011), thereby completing the DNA demethylation cycle (Fig. 1). This finding suggested a plausible pathway for active DNA demethylation and opened a new dynamic field of research (Tahiliani et al. 2009).

3.1 Architecture of TET Enzymes

The mammalian TET family comprises three paralogous members (TET1, 2136 aa in human, TET2, 2002 aa and TET3, 1776 aa), which share similar domain architecture (Fig. 2). They all are big proteins harboring a large, mostly unstructured N-terminal part and the C-terminal catalytic domain. The core catalytic domain is composed of a cysteine-rich region and a following double-stranded β helix domain (DSBH) characteristic for $\text{Fe}^{2+}/\alpha\text{KG}$ dioxygenases and (Hu et al. 2013, 2015). In metazoan TETs, the DSBH domain is interrupted by a large unstructured region, which is believed to engage in protein-protein interactions. In their N-termini, TET1 and TET3 contain a CXXC domain, which interacts with DNA (Xu et al. 2012; Jin et al. 2016). The CXXC domain of TET2 was lost during evolution after gene duplication and inversion, and is now encoded as a separate protein IDAX (inhibition of the dvl and axin complex) (Iyer et al. 2009).

The recently solved crystal structure of the human TET2 catalytic domain in complex with 5mC containing DNA substrate (PDB ID: 4NM6) revealed that the

DSBH core domain forms a globular structure, which is stabilized by the Cys-rich region enwrapping the DSBH core (Hu et al. 2013). The Cys-rich region is crucial for the stability of the DSBH domain and consequently for catalysis (Hu et al. 2013). DNA is bound above the DSBH core in a groove that is enriched in basic and hydrophobic amino acids. Similar to DNMTs and DNA repair enzymes (Klimasauskas et al. 1994), TET enzymes utilize a base flipping mechanism to position the target base in the catalytic pocket for the oxidation reaction. Once the base is located in the catalytic pocket, the methyl-group is oriented towards the catalytic iron and α -KG, which facilitate the catalytic turnover (Hu et al. 2013). TET enzymes follow a conserved catalytic mechanism that is characteristic for other known $\text{Fe}^{2+}/\alpha\text{KG}$ -dependent dioxygenases, like histone lysine demethylases (JMJC-family) [reviewed in Hausinger and Schofield (2015)].

3.2 Intrinsic DNA Sequence Specificity of TET Enzymes

Most of the studies on TET enzymes were focused on elucidating their biological role and their reaction products; however, the intrinsic biochemical properties of TET enzymes that govern their function remain not well investigated. Recent reports showed that TET-dependent demethylation in zygotes represents only a small fraction of all demethylation events and that TET-associated demethylation seems to be locus specific (Guo et al. 2014; von Meyenn et al. 2016). Additionally, fine mapping of the genomic location of 5hmC using SCL-exo protocol showed that 5hmC is highly enriched within defined sequence context (Serandour et al. 2016), suggesting that TET enzymes could display some sequence preferences. It is however still unknown what is the molecular reason granting this sequence selectivity. Both DNMTs and TET enzymes modify CpG dinucleotides, yet DNMT3 enzymes can also efficiently methylate non-CpG sites (as discussed above). Unfortunately, little work was contributed to investigate the intrinsic preference of TET enzymes towards non-CpG sites. In the initial report that identified TET enzymes as 5mC hydroxylases, the authors showed that these enzymes are capable of oxidation of 5mC embedded in a CpG site, yet non-CpG substrates were not tested (Tahiliani et al. 2009). Later on, it has been showed that 5mCpA and 5mCpC sites were poor substrates for TET2, with conversion efficiencies of $<2\%$ and $<5\%$, respectively, as opposed to $>85\%$ for 5mCG sites in the same sequence context (Hu et al. 2013). Like TET2, TET1 preferentially oxidizes 5mCpG with some incidence of oxidation of 5mCpC sites. Structural studies provided molecular explanation for the observed preference of TET enzymes towards the CpG sites (Hu et al. 2013). The TET2-DNA crystal structure showed that the target 5mC is specifically recognized by two hydrogen bonds formed by the side chains of H1904 and N1387 and base endocyclic nitrogen atoms N3 and N4, respectively. The base-stacking interaction between TET2 Y1902 residue and the pyrimidine base of the 5mC additionally supports this recognition. Furthermore, base-stacking interaction between Y1294 residue and the G:5mC base pair in the DNA provides specific recognition of the following G:C

base pair within the CpG dinucleotide. Intriguingly, TET2 enzyme does not make any contact with the methyl group of the target cytosine, suggesting that it could generate oxidation of 5hmC to 5fC and 5caC (Hu et al. 2013).

In the TET2:DNA co-crystal structure no protein-base specific contacts outside of the CpG site were observed, suggesting that the enzyme has weak or no flanking sequence specificity (Hu et al. 2013, 2015). Nevertheless, the bound DNA is strongly bent and distorted, giving the possibility of indirect readout of DNA sequence as observed with numerous other DNA binding proteins, restriction enzymes and bacterial MTases (Jurkowski et al. 2007; Little et al. 2008). Whether TET enzymes use indirect readout for sequence recognition remains to be addressed.

In addition, TET enzymes are also able to oxidize the methyl group of thymine (T) to 5-hydroxymethyl uracil (5hmU) (Pffaffeneder et al. 2014), however the efficiency of this reaction is rather low, and its physiological relevance still needs to be uncovered.

3.3 *On Site and Lateral Processivity of TET Enzymes*

Processivity of TET enzymes can be regarded in two different ways. First, as a serial oxidation of 5mC to 5hmC, 5fC and 5caC on a single CpG site without the enzyme dissociating from that site, which could be regarded as “on-site processivity”. Second, which could be called “lateral processivity”, is the consecutive oxidation of numerous CpGs on a single DNA molecule.

The isolated catalytic domain of human TET2 efficiently oxidizes 5mC to 5hmC, yet, further oxidation steps are inefficient, leading to the reaction stalling at the 5hmC state (Hu et al. 2015). Conversely, numerous reports showed that TET enzymes are capable of efficient conversion of 5mC to 5caC without being blocked at the 5hmC state (Tamanaha et al. 2016; Liu et al. 2017; Crawford et al. 2016). Moreover, the same group which reported stalling of the oxidation reaction at the 5hmC state, also showed that TET2 could convert 5mC all the way to 5caC (Hu et al. 2013). It is likely that the contradictory conclusions of the studies that investigated mouse TET2 “on-site” processivity could be potentially explained by differences in the reaction conditions and experimental setup (Tamanaha et al. 2016; Crawford et al. 2016).

The catalytic domains of TET1 and TET2 show no preference to modify neighboring CpG sites on the same DNA molecule, suggesting that TET enzymes are not laterally processive (Tamanaha et al. 2016). However, both TET1 and TET3 full-length enzymes contain an additional DNA binding domain, namely the CXXC domain, which can modify the enzymes behavior on DNA. Strikingly, TET3 CXXC domain preferentially binds 5caCpG sites, which represent the final TET reaction product. This observation led to proposal that the TET3 CXXC—5caCpG interaction could stimulate processive activity of the enzyme and consequently lead to spreading of the 5caC from the first oxidized CpG site. In the proposed model, the first 5mCpG site that is oxidized to 5caCpG gets bound by the CXXC domain of the

enzyme, therefore keeping the catalytic domain in close proximity and promoting oxidation of nearby 5mCpGs (Jin et al. 2016). It is a very interesting hypothesis, which still requires further experimental validation.

3.4 Oxidation of RNA Bases

Mammalian TET enzymes were initially found to oxidize 5mC in genomic DNA, but have since been also shown to oxidize 5mC in RNA (Basanta-Sanchez et al. 2017; Fu et al. 2014). Moreover, the presence of TET homologues in organisms that do not possess any active DNA methyltransferase, like *D. melanogaster*, suggested that other substrate than 5mC in dsDNA could be processed by the enzyme (Dunwell et al. 2013). Indeed, *Drosophila* TET is responsible (at least in part) for the formation of 5hmC in the fly mRNAs, particularly in mRNAs involved in neuronal development. Consequently, blocking of the TET enzyme causes brain defects and is lethal. In vivo, RNA hydroxymethylation promotes mRNA translation (Delatte et al. 2016).

DNA and RNA represent different structural configurations, which impact the way TET enzymes can interact with them. DeNizio and colleagues performed a systematic survey aimed to compare the activity of TET2 CD on ds- and ss- DNA and RNA, as well as DNA/RNA hybrids. They discovered that 5mC in dsDNA is the most proficient substrate, ssRNA and ssDNA are well tolerated, whereas dsRNA is a very poor substrate for TET2 (DeNizio et al. 2018).

3.5 Recruitment of TET Enzymes

The mechanisms of locus specific recruitment and regulation of TET enzymes is much less understood than the genomic distribution and physiological relevance of the oxidized-5mC derivatives. The CXXC domains located in the N-termini of TET1 and TET3 are thought to be at least in part responsible for the targeting of the enzymes to the CpG-rich regions (CpG islands), as the CXXC domain has been shown to recruit DNMT1, MLL1, CFP1 to unmethylated CpG sites (Stroynowska-Czerwinska et al. 2018; Xu et al. 2018). Consistently, DNA binding studies showed that the CXXC domain of TET1 is able to bind to CpG-rich DNA irrespective of its modification state (C, 5mC or 5hmC) (Zhang et al. 2010a), whereas the CXXC domain of TET3 from *Xenopus* binds unmodified cytosines in both CpG and non-CpG context, with a slightly higher preference for CpG (Xu et al. 2012; Jin et al. 2016). Another interesting study demonstrated that the CXXC domain of TET3 can bind 5caCpG and that full-length TET3 preferentially binds to the transcriptional start sites (TSS) of genes involved in base excision repair (Jin et al. 2016). This suggests that TET3 may be specifically targeted to these loci through the CXXC domain or by other interacting proteins (Jin et al. 2016).

TET2, which lacks the CXXC domain may be more depend on other proteins, for example transcription factors (TFs), for locus specific recruitment. Supporting this idea, TET2 interacts with the transcription factor Wilms tumor (WT) and Early B cell factor 1 (EBF1), which modulate TET2 activity and target gene expression [reviewed in (Ravichandran et al.)]. Recently, several TFs important for cellular differentiation were reported to induce DNA demethylation by interacting with TET proteins. For example, RUNX1, an essential master transcription factor in hematopoietic development and an important regulator of immune functions, was shown to recruit TET2 and induce local DNA demethylation at its binding regions (Suzuki et al. 2017). Likewise, NANOG-dependent recruitment of TET1 and TET2 promotes expression of genes involved in reprogramming and lineage commitment (Costa et al. 2013). Furthermore, a study by Perera and colleagues in mouse retinal cells demonstrated that RE1-silencing transcription factor (REST) recruits an isoform of TET3 lacking the CXXC domain along with the histone methyltransferase NSD3 to activate its target genes (Perera et al. 2015). TET enzymes were shown to interact with proteins involved in base excision repair pathway such as TDG, PARP1, MBD4, NEIL (Muller et al. 2014). Furthermore, all three TET enzymes associate with O-linked β -D-N-acetylglucosamine (O-GlcNAc) transferase (OGT). It has been suggested that TETs recruit OGT to the chromatin and that TET-OGT interaction promotes the OGT activity (Vella et al. 2013; Chen et al. 2013). In summary, it is increasingly clear that TET enzymes do not function alone but interact with multiple other proteins in a contextual manner and through this cooperation modulate gene expression.

4 Synthetic Programming of DNA Methylation

Rapid development of next-generation based sequencing technologies enabled genome-wide interrogation of cytosine methylation at single-base resolution, providing invaluable insights into the frequency and genomic distribution of 5mCs, as well as into the interplay between DNA methylation and other epigenetic mechanisms. Yet, the lack of tools for locus-specific manipulation of cytosine status has hampered the functional understanding of the role of DNA methylation and demethylation. Recent progress in programmable DNA binding domains has open new synthetic ways to study epigenetic regulation (Jurkowski et al. 2015), and in particular DNA methylation and demethylation. Fusing an active DNA methyltransferase or demethylase (or any other epigenetic enzyme) to a customizable DNA binding domain enables targeting of the methylation or demethylation functionality to selected places in the genome. From a mechanistic point of view this powerful technology permits not only to study the principles of how the enzymes set up or remove the methylation mark, but also to directly probe and dissect the epigenetic mechanisms and transcriptional consequences of DNA methylation or demethylation at a given genomic locus. On the application side, it allows verifying consequentiality of disease associated epigenetic changes or even their repair as a potential therapeutic strategy.

4.1 Programmable Genome Targeting Modules

Three different classes of programmable DNA binding domains have been employed so far in epigenetic editing. The C2H2 zinc fingers were the first example of predictable DNA interaction domains amenable to rational protein design [reviewed in Wolfe et al. (2000), Pabo et al. (2001)] and were first used for programmable sequence specific genome targeting of fused epigenetic enzymes (Xu and Bestor 1997). More recently, two additional programmable genome binders were discovered: the TAL effector arrays (TALE) (Boch et al. 2009) and CRISPR/Cas9 systems (Jinek et al. 2012). The transcription activating-like effectors (TALEs) are important virulence factors initially isolated from the bacterial plant pathogen *Xanthomonas* (Boch and Bonas 2010) and are composed of tandemly arranged 34 amino acid long highly similar repeats (Scholze and Boch 2010).

The newest and most exciting addition to the genome targeting toolbox repository is the CRISPR/Cas9 system (Hsu et al. 2014). CRISPR (clustered regularly interspaced short palindromic repeats) functions as a prokaryotic adaptive immune system that confers resistance to exogenous genetic elements such as plasmids and phages (Mojica et al. 2005). CRISPR/Cas9 proteins recognize their targets based on Watson/Crick base pairing and rely on complementarity of the recognized DNA and the guide RNA sequences which are used for targeting. Therefore, retargeting of the Cas9 protein to specific genomic location requires only a gRNA component specific for the desired target. However, because Cas9 is an active nuclease, for targeting of the epigenetic enzymes, a catalytically inactive Cas9 variant is used. It still recognizes and binds the target sequence, yet does not cleave it (Qi et al. 2013). Whereas each of the available programmable genome targeting domains offers unique advantages and disadvantages, due to the simplicity of target design and the possibility for multiplexing CRISPR/Cas9 system seems the most attractive.

4.2 Epigenetic Effector Domains

The epigenetic editing activity is provided by fusing active DNA methylating or demethylating enzymes to the targeting domain. Until now different DNA methyltransferases have been used [reviewed in Lau and Suh (2018), Lei et al. (2018)], which include bacterial CpG specific methyltransferases M.SssI (Xiong et al. 2017) or MQ1 (a modified CpG methyltransferase derived from *Mollicutes Spiroplasma*) (Lei et al. 2017), the catalytic domains or full-length mammalian Dnmt3a (Vojta et al. 2016; Liu et al. 2016) or Dnmt3b (Lin et al. 2018) proteins, as well as an engineered Dnmt3a-Dnmt3L fusion protein, which in addition to the Dnmt3a methyltransferase contains the co-activator protein Dnmt3L (Stepper et al. 2017; Saunderson et al. 2017). For targeted DNA demethylation, all three mammalian TET enzymes have been used (Liu et al. 2016), yet TET1 CD is the most commonly used version. Interestingly, a direct removal of methylated cytosine has

also been achieved recently by employing plant ROS1 DNA glycosylase, leading to the transcriptional increase of the target locus (Parrilla-Doblas et al. 2017).

4.3 Applications of Targeted DNA Methylation/ Demethylation

Targeted DNA modification (both methylation and demethylation) has the potential to answer so far unapproachable questions in basic and translational research. It allows mechanistic dissection of the epigenetic signaling cascades and validation of the causality of epigenetic changes in diseases. It can widen our understanding of epigenetic dynamics and the basis of stability of DNA methylation signal, but also address the contribution of epigenetic changes to etiology of complex and simple diseases, through discovery and validation of disease-promoting epimutations and provide means for reverting them (Fig. 3).

4.3.1 How Do Epigenetic Changes Contribute to Disease Etiology?

As discussed above, widespread changes in DNA methylation patterns are commonly observed in diseases (Egger et al. 2004; Koch et al. 2018), including cancer, chronic or acute diseases. However, it is hard to evaluate whether these epigenetic changes are causal for the disease progression or are merely by-standers, reflecting the overall epigenetic dysregulation caused by the disease.

Epigenome-wide association studies (EWAS) are commonly used to derive associations between epigenetic variation and a particular identifiable phenotype (Birney et al. 2016). When epigenetic patterns, such as DNA methylation, change at specific loci, discriminating the phenotypically affected cases from the control individuals, this is considered an indication that epigenetic perturbation has taken place that is associated either causally or consequentially with the studied phenotype. However, EWAS results do not allow discriminating causal from consequential epigenetic changes, just merely their correlation with the screened phenotype. In such cases, targeted DNA methylation/demethylation could be used to study the causality of the observed changes towards the phenotype.

Aberrant promoter methylation is a well-recognized hallmark of cancer; however, it is unclear whether epigenetic changes are enough to drive cellular transformation. Sanderson (Saunderson et al. 2017) used CRISPR-based targeted DNA methylation to stably methylate and repress the CDKN2A, HIC1, PTEN and RASSF1 tumor suppressor genes in healthy primary breast cells. Furthermore, they show that targeted *de novo* methylation of the CDKN2A p16 transcript promoter prevented cells from entering senescence arrest, thus possibly facilitating tumor initiation.

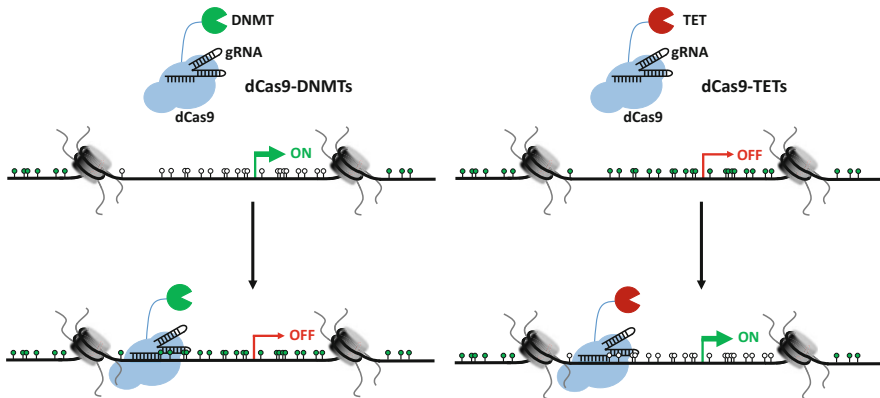


Fig. 3 Targeted DNA methylation and demethylation controls gene expression. Locus-specific targeting of DNA methyltransferases (DNMTs) represses the target gene, whereas demethylation of a methylated promoter after targeting of TETs leads to gene activation. Green lollipops indicate methylated CpG sites, white lollipops—unmethylated CpG sites

4.3.2 Repair of Aberrant, Disease Causing Epigenetic States

Fragile X syndrome (FXS) is the most frequent form of inherited mental retardation (Sutcliffe et al. 1992). FMR-1 gene found in fragile X patients shows an increase in the number of CGG repeats and an abnormal methylation of a CpG island 250 bp proximal to this repeat. Liu et al., used the dCas9-TET1 CD construct to demethylate CGG repeats in FXS induced pluripotent stem cells (iPS) and reactivate the silenced FMR1 gene by demethylating and activating its promoter (Liu et al. 2018).

4.4 Limitations of Targeted DNA Methylation and Demethylation Tools

Despite being such a very powerful technology, epigenetic editing has also its pitfalls and limitations. The promise of epigenetic regulation is that once DNA methylation is established or removed, cellular epigenetic mechanisms will maintain the new state of the locus, such that it can be inherited after semiconservative DNA replication (Jeltsch and Jurkowska 2014). Therefore, targeted DNA methylation or demethylation could provide a unique opportunity to heritably switch off gene expression (loss-of-function) (Siddique et al. 2013; Nunna et al. 2014; Stolzenburg et al. 2012). However, recent reports indicate that DNA methylation deposited at active gene promoters is not necessarily stably maintained and consequently gets diluted with DNA replication and cell division (Vojta et al. 2016; Kungulovski et al. 2015). Nevertheless, stable epigenetic reprogramming has also been achieved (Saunderson et al. 2017; Amabile et al. 2016), yet the epigenetic mechanisms which grant this stability are not well understood. It is plausible that the stability

of the introduced DNA methylation will depend on the local environment of the targeted locus. As DNA methylation is just one layer of epigenetic regulation, targeting of multiple epigenetic marks simultaneously might improve the stability of introduced DNA methylation.

Unintended, off-target epigenetic modification can lead to misinterpretation of the epigenetic editing experiments in regard to the observed biological effects, therefore specificity of introduced epigenetic modification is of principal importance. The precision of targeting is even more important for potential therapeutic epigenetic interventions, as mistargeted modifications can dysregulate other genes and cause additional diseases. Because of this, numerous studies have addressed this issue [reviewed in Lei et al. (2018)]. Two types of off-targeting can be distinguished: first one stems from the misrecognition of the targeting module (i.e. binding promiscuity of the dCas9 part to other near-cognate sequences in the genome) and the second one is the unintended modification by the epigenetic domains used.

In targeting experiments, dCas9 ChIP-seq coupled with bisulfite sequencing has been used to investigate the off-target methylation, and showed that even at the top ranking dCas9 binding sites dCas9-DNMT3a only marginally increased DNA methylation relative to the methylation observed at the intentionally targeted loci (Liu et al. 2016), suggesting that mis-targeting of dCas9 is not contributing strongly to off-target methylation. Other studies applied genome-wide sequencing technologies, including reduced-representation bisulfite sequencing (RRBS) and whole genome bisulfite sequencing to assess potential side effects of various methylation tools and reported no detectable off-target hypermethylation (Huang et al. 2017; Lei et al. 2017). Similarly, few off-target effects have been reported with demethylation tools.

In contrast, numerous studies also reported significant off-target methylation when targeting dCas9-DNMT3a CD. The extent of the off-targets effects varied vastly (Huang et al. 2017; Lei et al. 2017; McDonald et al. 2016). A recent study showed presence of extensive off-target genome-wide methylation in mouse ES cells (mESC) and somatic cells (Galonska et al. 2018) regardless of whether or not sgRNA was used for targeting. Expression level of the constructs might greatly influence the extent of off-targets, as once the “true” binding sites are occupied, the rest of the produced targeting constructs will be available to modify unspecific sites.

Acknowledgements We apologize to the authors whose work could not be cited due to the space limitations. Work in the authors laboratory has been supported by Boehringer Ingelheim (R.Z.J) and DFG SPP1784 (T.P.J).

References

- Amabile A, Migliara A, Capasso P et al (2016) Inheritable silencing of endogenous genes by hit-and-run targeted epigenetic editing. *Cell* 167:219–232, e214
- Arita K, Ariyoshi M, Tochio H et al (2008) Recognition of hemi-methylated DNA by the SRA protein UHRF1 by a base-flipping mechanism. *Nature* 455:818–821

- Barau J, Teissandier A, Zamudio N et al (2016) The DNA methyltransferase DNMT3C protects male germ cells from transposon activity. *Science* 354:909–912
- Basanta-Sanchez M, Wang R, Liu Z et al (2017) TET1-mediated oxidation of 5-formylcytosine (5fC) to 5-carboxycytosine (5caC) in RNA. *Chembiochem* 18:72–76
- Bashtrykov P, Jankevicius G, Smarandache A et al (2012) Specificity of Dnmt1 for methylation of hemimethylated CpG sites resides in its catalytic domain. *Chem Biol* 19:572–578
- Bashtrykov P, Jankevicius G, Jurkowska RZ et al (2014) The UHRF1 protein stimulates the activity and specificity of the maintenance DNA methyltransferase DNMT1 by an allosteric mechanism. *J Biol Chem* 289:4106–4115
- Baubec T, Colombo DF, Wirbelauer C et al (2015) Genomic profiling of DNA methyltransferases reveals a role for DNMT3B in genic methylation. *Nature* 520:243–247
- Bergman Y, Cedar H (2013) DNA methylation dynamics in health and disease. *Nat Struct Mol Biol* 20:274–281
- Birney E, Smith GD, Grealley JM (2016) Epigenome-wide association studies and the interpretation of disease -omics. *PLoS Genet* 12:e1006105
- Boch J, Bonas U (2010) Xanthomonas AvrBs3 family-type III effectors: discovery and function. *Annu Rev Phytopathol* 48:419–436
- Boch J, Scholze H, Schornack S et al (2009) Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* 326:1509–1512
- Bogdanovic O, Lister R (2017) DNA methylation and the preservation of cell identity. *Curr Opin Genet Dev* 46:9–14
- Bostick M, Kim JK, Esteve PO et al (2007) UHRF1 plays a role in maintaining DNA methylation in mammalian cells. *Science* 317:1760–1764
- Bruniquel D, Schwartz RH (2003) Selective, stable demethylation of the interleukin-2 gene enhances transcription by an active process. *Nat Immunol* 4:235–240
- Chedin F, Lieber MR, Hsieh CL (2002) The DNA methyltransferase-like protein DNMT3L stimulates de novo methylation by Dnmt3a. *Proc Natl Acad Sci U S A* 99:16916–16921
- Chen Q, Chen Y, Bian C et al (2013) TET2 promotes histone O-GlcNAcylation during gene transcription. *Nature* 493:561–564
- Cheng X, Blumenthal RM (2008) Mammalian DNA methyltransferases: a structural perspective. *Structure* 16:341–350
- Costa Y, Ding J, Theunissen TW et al (2013) NANOG-dependent function of TET1 and TET2 in establishment of pluripotency. *Nature* 495:370–374
- Crawford DJ, Liu MY, Nabel CS et al (2016) Tet2 catalyzes stepwise 5-methylcytosine oxidation by an iterative and de novo mechanism. *J Am Chem Soc* 138:730–733
- Delatte B, Wang F, Ngoc LV et al (2016) RNA biochemistry. Transcriptome-wide distribution and function of RNA hydroxymethylcytosine. *Science* 351:282–285
- DeNizio JE, Liu MY, Leddin E et al (2018) Selectivity and promiscuity in TET-mediated oxidation of 5-methylcytosine in DNA and RNA. *Biochemistry*. <https://doi.org/10.1021/acs.biochem.8b00912>
- Dhayalan A, Rajavelu A, Rathert P et al (2010) The Dnmt3a PWWP domain reads histone 3 lysine 36 trimethylation and guides DNA methylation. *J Biol Chem* 285:26114–26120
- Dunwell TL, McGuffin LJ, Dunwell JM et al (2013) The mysterious presence of a 5-methylcytosine oxidase in the *Drosophila* genome: possible explanations. *Cell Cycle* 12:3357–3365
- Egger G, Liang G, Aparicio A et al (2004) Epigenetics in human disease and prospects for epigenetic therapy. *Nature* 429:457–463
- Egger G, Jeong S, Escobar SG et al (2006) Identification of DNMT1 (DNA methyltransferase 1) hypomorphs in somatic knockouts suggests an essential role for DNMT1 in cell survival. *Proc Natl Acad Sci U S A* 103:14080–14085
- Emperle M, Rajavelu A, Reinhardt R et al (2014) Cooperative DNA binding and protein/DNA fiber formation increases the activity of the Dnmt3a DNA methyltransferase. *J Biol Chem* 289:29602–29613

- Ernst J, Kheradpour P, Mikkelsen TS et al (2011) Mapping and analysis of chromatin state dynamics in nine human cell types. *Nature* 473:43–49
- Fu L, Guerrero CR, Zhong N et al (2014) Tet-mediated formation of 5-hydroxymethylcytosine in RNA. *J Am Chem Soc* 136:11582–11585
- Galonska C, Charlton J, Mattei AL et al (2018) Genome-wide tracking of dCas9-methyltransferase footprints. *Nat Commun* 9:597
- Gowher H, Jeltsch A (2001) Enzymatic properties of recombinant Dnmt3a DNA methyltransferase from mouse: the enzyme modifies DNA in a non-processive manner and also methylates non-CpG [correction of non-CpA] sites. *J Mol Biol* 309:1201–1208
- Gowher H, Jeltsch A (2002) Molecular enzymology of the catalytic domains of the Dnmt3a and Dnmt3b DNA methyltransferases. *J Biol Chem* 277:20409–20414
- Gowher H, Jeltsch A (2018) Mammalian DNA methyltransferases: new discoveries and open questions. *Biochem Soc Trans* 46:1191–1202
- Gowher H, Liebert K, Hermann A et al (2005) Mechanism of stimulation of catalytic activity of Dnmt3A and Dnmt3B DNA-(cytosine-C5)-methyltransferases by Dnmt3L. *J Biol Chem* 280:13341–13348
- Gowher H, Loutchanwoot P, Vorobjeva O et al (2006) Mutational analysis of the catalytic domain of the murine Dnmt3a DNA-(cytosine C5)-methyltransferase. *J Mol Biol* 357:928–941
- Goyal R, Reinhardt R, Jeltsch A (2006) Accuracy of DNA methylation pattern preservation by the Dnmt1 methyltransferase. *Nucleic Acids Res* 34:1182–1188
- Guo F, Li X, Liang D et al (2014) Active and passive demethylation of male and female pronuclear DNA in the mammalian zygote. *Cell Stem Cell* 15:447–459
- Guo X, Wang L, Li J et al (2015) Structural insight into autoinhibition and histone H3-induced activation of DNMT3A. *Nature* 517:640–644
- Hajkova P, Erhardt S, Lane N et al (2002) Epigenetic reprogramming in mouse primordial germ cells. *Mech Dev* 117:15–23
- Handa V, Jeltsch A (2005) Profound flanking sequence preference of Dnmt3a and Dnmt3b mammalian DNA methyltransferases shape the human epigenome. *J Mol Biol* 348:1103–1112
- Harrison JS, Cornett EM, Goldfarb D et al (2016) Hemi-methylated DNA regulates DNA methylation inheritance through allosteric activation of H3 ubiquitylation by UHRF1. *elife* 5:e17101
- Hashimoto H, Horton JR, Zhang X et al (2008) The SRA domain of UHRF1 flips 5-methylcytosine out of the DNA helix. *Nature* 455:826–829
- Hausinger RP, Schofield CJ (2015) 2-oxoglutarate-dependent oxygenases. *RSC metallobiology*, vol 3. Royal Society of Chemistry, Cambridge
- He Y, Ecker JR (2015) Non-CG methylation in the human genome. *Annu Rev Genomics Hum Genet* 16:55–77
- He YF, Li BZ, Li Z et al (2011) Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science* 333:1303–1307
- 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* 279:48350–48359
- Hodges E, Smith AD, Kendall J et al (2009) High definition profiling of mammalian DNA methylation by array capture and single molecule bisulfite sequencing. *Genome Res* 19:1593–1605
- Hsu PD, Lander ES, Zhang F (2014) Development and applications of CRISPR-Cas9 for genome engineering. *Cell* 157:1262–1278
- Hu L, Li Z, Cheng J et al (2013) Crystal structure of TET2-DNA complex: insight into TET-mediated 5mC oxidation. *Cell* 155:1545–1555
- Hu L, Lu J, Cheng J et al (2015) Structural insight into substrate preference for TET-mediated oxidation. *Nature* 527:118–122
- Huang YH, Su J, Lei Y et al (2017) DNA epigenome editing using CRISPR-Cas SunTag-directed DNMT3A. *Genome Biol* 18:176

- Iida T, Suetake I, Tajima S et al (2002) PCNA clamp facilitates action of DNA cytosine methyltransferase 1 on hemimethylated DNA. *Genes Cells* 7:997–1007
- Ito S, Shen L, Dai Q et al (2011) Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science* 333:1300–1303
- Iyer LM, Tahiliani M, Rao A et al (2009) Prediction of novel families of enzymes involved in oxidative and other complex modifications of bases in nucleic acids. *Cell Cycle* 8:1698–1710
- Jang HS, Shin WJ, Lee JE et al (2017) CpG and non-CpG methylation in epigenetic gene regulation and brain function. *Genes (Basel)* 8(6). pii: E148
- Jeltsch A, Jurkowska RZ (2013) Multimerization of the dnmt3a DNA methyltransferase and its functional implications. *Prog Mol Biol Transl Sci* 117:445–464
- Jeltsch A, Jurkowska RZ (2014) New concepts in DNA methylation. *Trends Biochem Sci* 39:310–318
- Jeltsch A, Jurkowska RZ (2016) Allosteric control of mammalian DNA methyltransferases – a new regulatory paradigm. *Nucleic Acids Res* 44:8556–8575
- Jia D, Jurkowska RZ, Zhang X et al (2007) Structure of Dnmt3a bound to Dnmt3L suggests a model for de novo DNA methylation. *Nature* 449:248–251
- Jin SG, Zhang ZM, Dunwell TL et al (2016) Tet3 reads 5-carboxylcytosine through its CXXX domain and is a potential guardian against neurodegeneration. *Cell Rep* 14:493–505
- Jinek M, Chylinski K, Fonfara I et al (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337:816–821
- Jurkowska RZ, Jeltsch A (2016) Enzymology of mammalian DNA methyltransferases. *Adv Exp Med Biol* 945:87–122
- Jurkowska RZ, Anspach N, Urbanke C et al (2008) Formation of nucleoprotein filaments by mammalian DNA methyltransferase Dnmt3a in complex with regulator Dnmt3L. *Nucleic Acids Res* 36:6656–6663
- Jurkowska RZ, Jurkowski TP, Jeltsch A (2011a) Structure and function of mammalian DNA methyltransferases. *Chembiochem* 12:206–222
- Jurkowska RZ, Siddique AN, Jurkowski TP et al (2011b) Approaches to enzyme and substrate design of the murine Dnmt3a DNA methyltransferase. *Chembiochem* 12:1589–1594
- Jurkowski TP, Anspach N, Kulishova L et al (2007) The M.EcoRV DNA-(adenine N6)-methyltransferase uses DNA bending for recognition of an expanded EcoDam recognition site. *J Biol Chem* 282:36942–36952
- Jurkowski TP, Ravichandran M, Stepper P (2015) Synthetic epigenetics-towards intelligent control of epigenetic states and cell identity. *Clin Epigenetics* 7:18
- Kinde B, Gabel HW, Gilbert CS et al (2015) Reading the unique DNA methylation landscape of the brain: non-CpG methylation, hydroxymethylation, and MeCP2. *Proc Natl Acad Sci U S A* 112:6800–6806
- Klimasauskas S, Kumar S, Roberts RJ et al (1994) HhaI methyltransferase flips its target base out of the DNA helix. *Cell* 76:357–369
- Koch A, Joosten SC, Feng Z et al (2018) Analysis of DNA methylation in cancer: location revisited. *Nat Rev Clin Oncol* 15:459–466
- Kolasinska-Zwierz P, Down T, Latorre I et al (2009) Differential chromatin marking of introns and expressed exons by H3K36me3. *Nat Genet* 41:376–381
- Kungulovski G, Nunna S, Thomas M et al (2015) Targeted epigenome editing of an endogenous locus with chromatin modifiers is not stably maintained. *Epigenetics Chromatin* 8:12
- Lau CH, Suh Y (2018) In vivo epigenome editing and transcriptional modulation using CRISPR technology. *Transgenic Res* 27:489–509
- Lei Y, Zhang X, Su J et al (2017) Targeted DNA methylation in vivo using an engineered dCas9-MQ1 fusion protein. *Nat Commun* 8:16026
- Lei Y, Huang YH, Goodell MA (2018) DNA methylation and de-methylation using hybrid site-targeting proteins. *Genome Biol* 19:187
- Li BZ, Huang Z, Cui QY et al (2011) Histone tails regulate DNA methylation by allosterically activating de novo methyltransferase. *Cell Res* 21:1172–1181

- Lin L, Liu Y, Xu F et al (2018) Genome-wide determination of on-target and off-target characteristics for RNA-guided DNA methylation by dCas9 methyltransferases. *GigaScience* 7:1–19
- Little EJ, Babic AC, Horton NC (2008) Early interrogation and recognition of DNA sequence by indirect readout. *Structure* 16:1828–1837
- Liu XS, Wu H, Ji X et al (2016) Editing DNA methylation in the mammalian genome. *Cell* 167:233–247 e217
- Liu MY, Torabifard H, Crawford DJ et al (2017) Mutations along a TET2 active site scaffold stall oxidation at 5-hydroxymethylcytosine. *Nat Chem Biol* 13:181–187
- Liu XS, Wu H, Krzisch M et al (2018) Rescue of fragile X syndrome neurons by DNA methylation editing of the FMR1 gene. *Cell* 172:979–992 e976
- 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* 286:35334–35338
- Martinowich K, Hattori D, Wu H et al (2003) DNA methylation-related chromatin remodeling in activity-dependent BDNF gene regulation. *Science* 302:890–893
- Mayer W, Niveleau A, Walter J et al (2000) Demethylation of the zygotic paternal genome. *Nature* 403:501–502
- McDonald JI, Celik H, Rois LE et al (2016) Reprogrammable CRISPR/Cas9-based system for inducing site-specific DNA methylation. *Biol Open* 5:866–874
- Meissner A, Mikkelsen TS, Gu H et al (2008) Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature* 454:766–770
- Mojica FJ, Diez-Villasenor C, Garcia-Martinez J et al (2005) Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J Mol Evol* 60:174–182
- Morselli M, Pastor WA, Montanini B et al (2015) In vivo targeting of de novo DNA methylation by histone modifications in yeast and mouse. *elife* 4:e06205
- Muller U, Bauer C, Siegl M et al (2014) TET-mediated oxidation of methylcytosine causes TDG or NEIL glycosylase dependent gene reactivation. *Nucleic Acids Res* 42:8592–8604
- Nady N, Lemak A, Walker JR et al (2011) Recognition of multivalent histone states associated with heterochromatin by UHRF1 protein. *J Biol Chem* 286:24300–24311
- Neri F, Rapelli S, Krepelova A et al (2017) Intragenic DNA methylation prevents spurious transcription initiation. *Nature* 543:72–77
- Nishiyama A, Yamaguchi L, Sharif J et al (2013) Uhrf1-dependent H3K23 ubiquitylation couples maintenance DNA methylation and replication. *Nature* 502:249–253
- Noh KM, Wang H, Kim HR et al (2015) Engineering of a histone-recognition domain in Dnmt3a alters the epigenetic landscape and phenotypic features of mouse ESCs. *Mol Cell* 59:89–103
- Norvil AB, Petell CJ, Alabdi L et al (2018) Dnmt3b methylates DNA by a noncooperative mechanism, and its activity is unaffected by manipulations at the predicted dimer interface. *Biochemistry* 57:4312–4324
- Nunna S, Reinhardt R, Ragozin S et al (2014) Targeted methylation of the epithelial cell adhesion molecule (EpCAM) promoter to silence its expression in ovarian cancer cells. *PLoS One* 9: e87703
- Oswald J, Engemann S, Lane N et al (2000) Active demethylation of the paternal genome in the mouse zygote. *Curr Biol* 10:475–478
- Otani J, Nankumo T, Arita K et al (2009) Structural basis for recognition of H3K4 methylation status by the DNA methyltransferase 3A ATRX-DNMT3-DNMT3L domain. *EMBO Rep* 10:1235–1241
- Pabo CO, Peisach E, Grant RA (2001) Design and selection of novel Cys2His2 zinc finger proteins. *Annu Rev Biochem* 70:313–340
- Parrilla-Doblas JT, Ariza RR, Roldan-Arjona T (2017) Targeted DNA demethylation in human cells by fusion of a plant 5-methylcytosine DNA glycosylase to a sequence-specific DNA binding domain. *Epigenetics* 12:296–303
- Perera A, Eisen D, Wagner M et al (2015) TET3 is recruited by REST for context-specific hydroxymethylation and induction of gene expression. *Cell Rep* 11:283–294

- Petell CJ, Alabdi L, He M et al (2016) An epigenetic switch regulates de novo DNA methylation at a subset of pluripotency gene enhancers during embryonic stem cell differentiation. *Nucleic Acids Res* 44:7605–7617
- Pfaffeneder T, Spada F, Wagner M et al (2014) Tet oxidizes thymine to 5-hydroxymethyluracil in mouse embryonic stem cell DNA. *Nat Chem Biol* 10:574–581
- Qi LS, Larson MH, Gilbert LA et al (2013) Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* 152:1173–1183
- Qin W, Wolf P, Liu N et al (2015) DNA methylation requires a DNMT1 ubiquitin interacting motif (UIM) and histone ubiquitination. *Cell Res* 25:911–929
- Qiu C, Sawada K, Zhang X et al (2002) The PWWP domain of mammalian DNA methyltransferase Dnmt3b defines a new family of DNA-binding folds. *Nat Struct Biol* 9:217–224
- Rajavelu A, Lungu C, Emperle M et al (2018) Chromatin-dependent allosteric regulation of DNMT3A activity by MeCP2. *Nucleic Acids Res* 46:9044–9056
- Ramsahoye BH, Binizskiewicz D, Lyko F et al (2000) Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. *Proc Natl Acad Sci U S A* 97:5237–5242
- Ravichandran M, Jurkowska RZ, Jurkowski TP (2018) Target specificity of mammalian DNA methylation and demethylation machinery. *Org Biomol Chem* 16:1419–1435
- Rondelet G, Dal Maso T, Willems L et al (2016) Structural basis for recognition of histone H3K36me3 nucleosome by human de novo DNA methyltransferases 3A and 3B. *J Struct Biol* 194:357–367
- Rothbart SB, Krajewski K, Nady N et al (2012) Association of UHRF1 with methylated H3K9 directs the maintenance of DNA methylation. *Nat Struct Mol Biol* 19:1155–1160
- Saunderson EA, Stepper P, Gomm JJ et al (2017) Hit-and-run epigenetic editing prevents senescence entry in primary breast cells from healthy donors. *Nat Commun* 8:1450
- Scholze H, Boch J (2010) TAL effector-DNA specificity. *Virulence* 1:428–432
- Schubeler D (2015) Function and information content of DNA methylation. *Nature* 517:321–326
- Serandour AA, Avner S, Mahe EA et al (2016) Single-CpG resolution mapping of 5-hydroxymethylcytosine by chemical labeling and exonuclease digestion identifies evolutionarily unconserved CpGs as TET targets. *Genome Biol* 17:56
- Sharif J, Muto M, Takebayashi S et al (2007) The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. *Nature* 450:908–912
- Siddique AN, Nunna S, Rajavelu A et al (2013) Targeted methylation and gene silencing of VEGF-A in human cells by using a designed Dnmt3a-Dnmt3L single-chain fusion protein with increased DNA methylation activity. *J Mol Biol* 425:479–491
- Smith ZD, Meissner A (2013) DNA methylation: roles in mammalian development. *Nat Rev Genet* 14:204–220
- Song J, Teplova M, Ishibe-Murakami S et al (2012) Structure-based mechanistic insights into DNMT1-mediated maintenance DNA methylation. *Science* 335:709–712
- Stepper P, Kungulovski G, Jurkowska RZ et al (2017) Efficient targeted DNA methylation with chimeric dCas9-Dnmt3a-Dnmt3L methyltransferase. *Nucleic Acids Res* 45:1703–1713
- Stolzenburg S, Rots MG, Beltran AS et al (2012) Targeted silencing of the oncogenic transcription factor SOX2 in breast cancer. *Nucleic Acids Res* 40:6725–6740
- Stroynowska-Czerwinska A, Piasecka A, Bochtler M (2018) Specificity of MLL1 and TET3 CXXC domains towards naturally occurring cytosine modifications. *Biochim Biophys Acta Gene Regul Mech* 1861:1093–1101
- Sutcliffe JS, Nelson DL, Zhang F et al (1992) DNA methylation represses FMR-1 transcription in fragile X syndrome. *Hum Mol Genet* 1:397–400
- Suzuki T, Shimizu Y, Furuhashi E et al (2017) RUNX1 regulates site specificity of DNA demethylation by recruitment of DNA demethylation machineries in hematopoietic cells. *Blood Adv* 1:1699–1711
- Tahiliani M, Koh KP, Shen Y et al (2009) Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* 324:930–935

- Tamanaha E, Guan S, Marks K et al (2016) Distributive processing by the iron(II)/alpha-ketoglutarate-dependent catalytic domains of the TET enzymes is consistent with epigenetic roles for oxidized 5-methylcytosine bases. *J Am Chem Soc* 138:9345–9348
- Vakoc CR, Sachdeva MM, Wang H et al (2006) Profile of histone lysine methylation across transcribed mammalian chromatin. *Mol Cell Biol* 26:9185–9195
- Varley KE, Gertz J, Bowling KM et al (2013) Dynamic DNA methylation across diverse human cell lines and tissues. *Genome Res* 23:555–567
- Vella P, Scelfo A, Jammula S et al (2013) Tet proteins connect the O-linked N-acetylglucosamine transferase Ogt to chromatin in embryonic stem cells. *Mol Cell* 49:645–656
- Vilkaitis G, Suetake I, Klimasauskas S et al (2005) Processive methylation of hemimethylated CpG sites by mouse Dnmt1 DNA methyltransferase. *J Biol Chem* 280:64–72
- Vojta A, Dobrinic P, Tadic V et al (2016) Repurposing the CRISPR-Cas9 system for targeted DNA methylation. *Nucleic Acids Res* 44:5615–5628
- von Meyenn F, Iurlaro M, Habibi E et al (2016) Impairment of DNA methylation maintenance is the main cause of global demethylation in naive embryonic stem cells. *Mol Cell* 62:983
- Wolfe SA, Nekludova L, Pabo CO (2000) DNA recognition by Cys2His2 zinc finger proteins. *Annu Rev Biophys Biomol Struct* 29:183–212
- Xiong T, Meister GE, Workman RE et al (2017) Targeted DNA methylation in human cells using engineered dCas9-methyltransferases. *Sci Rep* 7:6732
- Xu GL, Bestor TH (1997) Cytosine methylation targeted to pre-determined sequences. *Nat Genet* 17:376–378
- Xu Y, Xu C, Kato A et al (2012) Tet3 CXXC domain and dioxygenase activity cooperatively regulate key genes for *Xenopus* eye and neural development. *Cell* 151:1200–1213
- Xu C, Liu K, Lei M et al (2018) DNA sequence recognition of human CXXC domains and their structural determinants. *Structure* 26:85–95 e83
- Yamazaki Y, Mann MR, Lee SS et al (2003) Reprogramming of primordial germ cells begins before migration into the genital ridge, making these cells inadequate donors for reproductive cloning. *Proc Natl Acad Sci U S A* 100:12207–12212
- Yang L, Rau R, Goodell MA (2015) DNMT3A in haematological malignancies. *Nat Rev Cancer* 15:152–165
- Zhang H, Zhang X, Clark E et al (2010a) TET1 is a DNA-binding protein that modulates DNA methylation and gene transcription via hydroxylation of 5-methylcytosine. *Cell Res* 20:1390–1393
- Zhang Y, Jurkowska R, Soeroes S et al (2010b) Chromatin methylation activity of Dnmt3a and Dnmt3a/3L is guided by interaction of the ADD domain with the histone H3 tail. *Nucleic Acids Res* 38:4246–4253
- Zhang ZM, Lu R, Wang P et al (2018) Structural basis for DNMT3A-mediated de novo DNA methylation. *Nature* 554:387–391
- Ziller MJ, Muller F, Liao J et al (2011) Genomic distribution and inter-sample variation of non-CpG methylation across human cell types. *PLoS Genet* 7:e1002389