DNA Modification Readers and Writers and Their Interplay

Cathia Rausch, Florian D. Hastert and M. Cristina Cardoso

Cell Biology and Epigenetics, *Department of Biology, Technische Universität Darmstadt, 64287 Darmstadt, Germany*

Correspondence to M. Cristina Cardoso: *Department of Biology, TU Darmstadt, Schnittspahnstr. 10, 64287 Darmstadt, Germany. Fax: +49 6151 16 21882. cardoso@bio.tu-darmstadt.de*
https://doi.org/10.1016/j.jmb.2019.12.018
Edited by Pierre-Antoine Defossez

Abstracts

Genomic DNA is modified in a postreplicative manner and several modifications, the enzymes responsible for their deposition as well as proteins that read these modifications, have been described. Here, we focus on the impact of DNA modifications on the DNA helix and review the writers and readers of cytosine modifications and how they interplay to shape genome composition, stability, and function.

Introduction

Epigenetic modifications of mammalian DNA are crucial regulators of organismic development but also in pathogenesis. The most abundant and most studied mammalian DNA modification was described by Rollin Hotchkiss in 1948, when he identified, in addition to the four canonical DNA bases, a fifth peak on a paper chromatogram from thymus deoxyribonucleic acids. He termed this novel mammalian DNA base "epicytosine" [1], unknowingly anticipating the modern definition of epigenetics and the additional layer of information that epigenetic DNA/chromatin modifications add to the genetic information of every cell. In the same study, Hotchkiss suggested that "epicytosine" corresponded to methylated cytosine. In 1972, a sixth DNA base, named 5-hydroxymethylcytosine (5hmC), was identified in different vertebrate tissues, although this was likely the result of an experimental artifact as in the same study 5-methylcytosine (5mC) was not detected [2]. Hence, 5mC remained the only noncanonical mammalian DNA base that was studied to a larger extent for a long time and helped explaining how different cellular phenotypes can arise from the same nucleotide sequence. The postreplicative covalent addition of a methyl group to the fifth carbon atom of the cytosine pyrimidine ring by the DNA methyltransferase (DNMT) enzymes and oxidation thereof by the Ten-eleven-translocation (TET) enzymes diversify the genome and shape the chromatin landscape (Fig. 1A). DNA cytosine methylation is predominantly found in the context of symmetric CpG dinucleotides and, in the human genome, 70%–80% of all CpGs harbor methylated cytosines [3,4]. This modification is crucial in numerous cellular processes such as X-chromosome inactivation and imprinting, gene repression, control of cellular development and differentiation, silencing of repetitive elements, and genome stability maintenance [4–7]. Meanwhile, two more cytosine variants, 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC), were identified as oxidation products of 5hmC mediated by TET enzymes [8] (Fig. 1A), shedding light into a possible active DNA demethylation pathway [9]. However, 5hmC and 5fC were also identified as stable epigenetic marks [10,11] and all cytosine modifications are recognized by specific reader proteins [12]. The (hydroxy)methylome is specifically read, amongst others, by members of the methyl-CpG binding domain (MBD) protein family (reviewed in Ref. [13]).

Here, we review the effects of DNA base modifications on DNA double helix structure and the interplay between cytosine methylation readers and modifiers. Although methylation of the nitrogen-6 of deoxyadenosine (m6dA) has been recently described in DNA from mammals [14,15] and linked to nucleosomal positioning in *Tetrahymena* [16], it

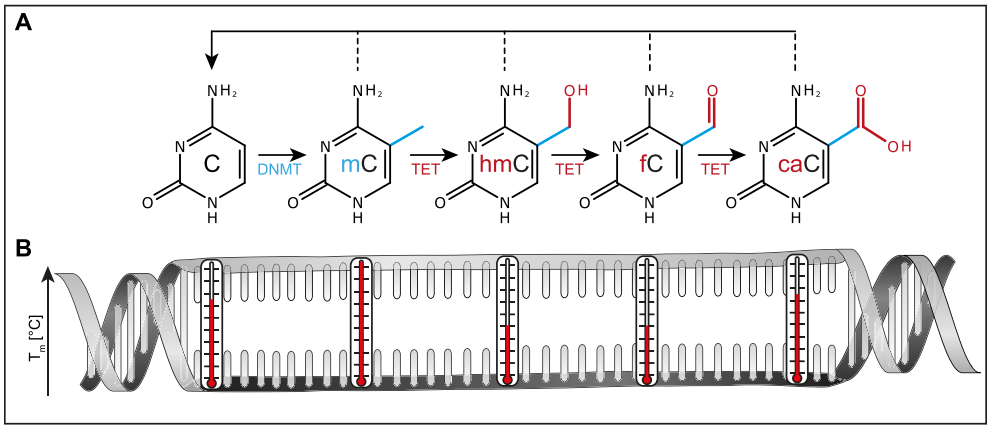

Fig. 1. Effects of epigenetic cytosine modifications. (A) DNA methyltransferases (DNMT) catalyze the addition of a methyl group to the fifth carbon atom of the cytosine pyrimidine ring creating 5-methylcytosine (5mC). The methyl group is oxidized by Ten-eleven-translocation (TET) proteins creating 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC), and 5-carboxylcytosine (5caC) in an iterative manner. **(B)** Schematic representation of the effect that different epigenetic cytosine modifications exert on the melting temperature (*T*m; graphically depicted as a thermometer) of the DNA double helix.

will be omitted from this review as its existence is still under debate [17].

Effect of cytosine modifications on DNA double helix structure and recognition by proteins

The flexibility of the dynamic DNA double helix system is a fundamental condition to allow tight protein-DNA contacts, is sequence dependent, and mainly determined by the stacking energy of base steps, hydrogen bondings between base pairs, and exocyclic functional groups pointing within the major or minor groove [18]. Since DNA replication, transcription, and repair all rely on unwinding of the DNA double helix, the influence of base modifications on helix parameters might directly affect DNA metabolic processes.

A variety of *in vitro* reports showed that, besides the mere DNA base composition that influences the melting temperature (T_m) of double-stranded DNA (dsDNA) [19], also the addition of chemical moieties, such as methyl or hydroxymethyl groups to DNA bases can influence the mechano-physical properties of the DNA double helix. C5 cytosine methylation was early on described to increase the melting temperature of an oligonucleotide carrying the modification in comparison to an unmethylated counterpart in a highly sequence-dependent manner [20,21]. In the following years, these findings were confirmed using UV spectroscopy, thermal melting analysis, and mechanical strand separation experiments and were also extended to the oxidative methylcytosine derivatives. Generally, cytosine methylation was reported to increase DNA melting temperature, indicative of a stabilization effect on the DNA double helix, whereas oxidation of the methyl group decreases the T_m, reversing the methylation-induced duplex stabilizing impact (Fig. 1B) [22–33]. 5hmC was in some studies found to not only revert the stabilizing influence of cytosine methylation, but to even further destabilize the double helix, resulting in a T_m lower than that of nonmodified cytosine (T_m: 5mC > C > 5hmC) [24,25,32].

The influence of cytosine modifications on DNA structure is still highly debated. The general view, however, is that the modified bases do not change the global B-DNA conformation, although local structural changes around the modification site were measured in a sequence-dependent manner [27,32,34–37]. Molecular dynamics and all-atom Monte-Carlo simulations showed that cytosine methylation changes the geometry of base pair steps resulting in an increase of the local curvature of the methylated DNA [34,37–40]. In addition, methylated DNA stretches can show slid structures that lead to local DNA distortions [22]; they are stiffer, more torsionally rigid, and less flexible, since the steric hindrance of the bulky methyl group within the major groove counteracts DNA bending [37,41–43]. Moreover, cytosine methylation might lead to a slight widening of the major groove and in turn a narrower minor groove [44]. On the contrary, 5hmC and 5fC containing DNA exhibit changes in base pair geometry that result in an increased flexibility [32,42], as well as alterations in the minor groove geometry [45]. Wanunu et al. showed computationally and experimentally that the increasing polarity of the cytosine modification leads to a lowering of the rigidity of the DNA stretch. In turn, the change from hydrophobicity (5mC) to hydrophilicity (oxidized

variants) of the DNA base modification can lead to a less stable DNA duplex [32], thereby linking DNA structural changes to DNA flexibility and rigidity, the solvent shell state, and eventually double helix stability.

The addition of a functional group to DNA bases can also influence DNA structure by changing base stacking and hydrogen bonding between base pairs. Cytosine methylation enhances base stacking interactions [22,28,30,34,46], due to a higher molecular polarizability upon methyl substitution [47]. Interestingly, base stacking abilities for 5hmC, 5fC, and 5caC were found to be similar [48]. Hydrogen bondings between base pairs containing 5mC or 5hmC were found not to be influenced by cytosine modifications [39,46,49]; however, there are reports that 5fC and 5caC substantially affect H-bonding interactions [34,50] and form intranucleobase hydrogen bonds between the oxygens of the ketone groups and the exocyclic N4 amino group of the cytosine [45,48].

These base modification–dependent changes in the local structure of DNA were proposed to represent a way to their selective recognition by proteins. High-resolution DNase I cleavage profiles and bisulfite sequencing data showed a significant increase in DNase I cleavage adjacent to methylated CpGs over unmethylated CpGs and this DNase hypersensitivity was attributed to methylation-induced minor groove narrowing [51]. The electron-withdrawing effect of the formyl and carboxyl groups and the concomitant weakening of the 5fC:G and 5caC:G base pairing was postulated to facilitate the flipping out of the double helix of these modified bases for recognition and excision by the thymine-DNA glycosylase (TDG) during the DNA demethylation pathway [50]. Accordingly, minor groove alterations in formylated and carboxylated DNA were proposed to be specifically recognized by TDG, resulting in its substrate specificity [26,45]. The contradictory results found by others, however, provide doubts to the 5fC-induced structural changes of DNA being a weighty DNA recognition platform for proteins, as the changes are rather subtle and very local [35,45]. In general, (modified) DNA readout can depend on base or shape readout. Both can be influenced by modification of cytosines; the addition of a bulky methyl group to the major groove could facilitate or interfere with protein side chain contacts [52], whereas the local structural change in DNA upon base modification may allow or inhibit protein recognition and binding [51]. Hence, changes in DNA shape were also demonstrated to play a role in enhancing DNA binding of the transcription factor p53 [53].

The changes in groove geometry that come along with cytosine methylation were related to alterations in solvation dynamics and water densities compared to unmodified DNA [32]. Conformingly and in view of DNA shape variations as a mechanism for specific protein recognition, MeCP2 was found to interact and recognize specific water molecules within the hydration shell of the major groove of methylated DNA, rather than the methylated cytosine *per se* [54]. Additionally, the above-mentioned narrowing of the minor grove observed in 5mC containing DNA enhances the negative electrostatic potential. The latter, in turn, can more efficiently attract basic side chains of DNA binding proteins [44,51,55]. Of note, the MBD domains of MeCP2, MBD1, and MBD2 were found to identify the mCpG dinucleotide via two conserved positively charged arginine rich fingers [54,56,57], which might exploit the described changes for binding site identification. Besides the recognition/binding of mCpG dinucleotides via its MBD domain, MeCP2 exhibits two conserved AT-hooks that specifically recognize adjacent AT-runs [58], which were also shown to have a tendency for binding narrow minor grooves [55]. It is therefore tempting to speculate about a possible DNA shape–based recognition mechanism of 5mC that relies on local changes of the DNA double helix structure arising upon base modification.

DNMT writers, MBD readers, and TET modifiers

DNMT writer enzymes and the 5mC mark

Methylation of cytosine in eukaryotic DNA relies on the action of the members of the S-adenosyl-L-methionine dependent DNA methyltransferase family (DNMT), which are capable of introducing new DNA methylation sites, as well as maintaining the methylation state upon DNA replication during the cell cycle. DNMT1, the first member of this family, can be considered an epigenetic janitor, since it methylates hemimethylated CpG sites that occur in the course of DNA replication and repair [59–61]. DNMT3a and DNMT3b, in contrast, are *de novo* methyltransferases; both show a strong affinity for unmethylated DNA and establish the initial DNA methylation patterns during embryogenesis and gametogenesis. Both proteins are essential for pluripotency repression during cell differentiation and development, are highly expressed in ES and germ cells, but are downregulated upon differentiation and in general in somatic tissues [62–64]. While DNMT3a was shown to play an important role in setting parental imprints [63], DNMT3b was linked to the correct methylation pattern of pericentric DNA regions and, hence, chromosome structure [62]. The physiological importance of DNMT1, DNMT3a, and DNMT3b is underlined by the negative effects their depletion has on organismic development and survival, as well as their link to human diseases.

The majority of cytosines in the human genome that reside within CpG dinucleotides are marked by a methyl group [3,4], and, due to the mutagenic potential of 5mC, a global CpG depletion is observed [65–67]. However, the majority of the CpGs are hypermethylated and found mostly in repetitive and parasitic DNA elements, such as retrotransposons, LINE1 (long interspersed nuclear elements), and Alu elements [68,69], whose silencing depends on CpG methylation. Recent studies showed that transcriptional regulation via DNA methylation is context-dependent, with methylation within promoter regions generally correlated with gene repression, and gene body methylation typically related to gene activation [70–73]. In addition to CpG methylation, recent sequencing-based studies demonstrated the presence of non-CpG methylation in pluripotent cells and neurons [74–78]. Interestingly, the methyl-CpG binding protein MeCP2 was recently shown to recognize and bind methylated cytosines in non-CpG contexts (mCAC) and to regulate the repression of long genes in the brain [56,79–82].

MBD reader proteins

DNA methylation was early-on linked to transcriptional repression amongst other mechanisms via the binding of members of the MBD protein family (Fig. 2A). The MBD (methyl-CpG binding domain) protein family is one of the best studied protein families capable of reading the genomic (hydroxy) methylome and thereby mediating between cytosine modifications, DNA and histone modifying enzyme complexes, non-coding RNAs and distinct chromatin states. The protein family comprises 11 members, namely MeCP2, MBD1-6, SETDB1, SETDB2, BAZ2A, and BAZ2B [83–88]. Besides the conserved methyl-CpG binding domain (MBD) that all

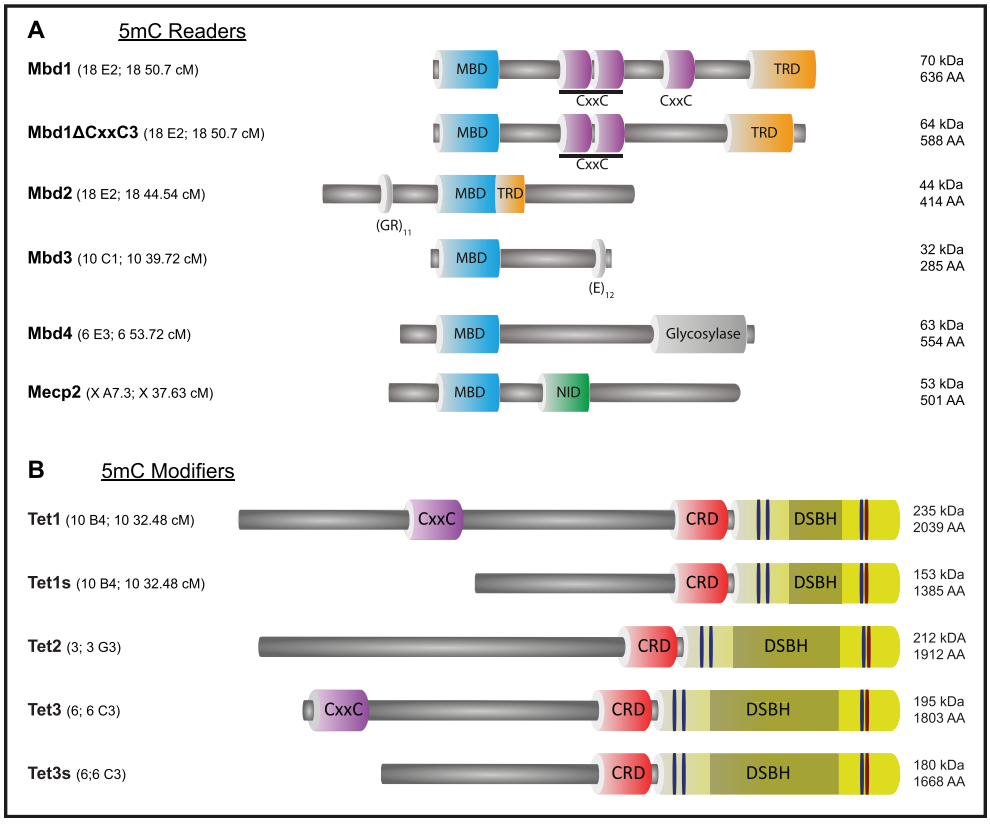

Fig. 2. Schematic representation of selected 5mC reader and modifier proteins. **(A)** Domain organization of selected members of the mouse methyl-CpG binding domain proteins including MBD1, MBD2, MBD3, MBD4, MeCP2, and a MBD1 isoform that lacks the third CXXC domain. The domains shown include the MBD (methyl-CpG binding domain), CXXC (CXXC zinc finger domain), TRD (transcriptional repression domain), and NID (NCoR/SMRT interaction domain, GR (glycine/arginine), E (glutamic acid). **(B)** Domain organizations of the long isoforms of mouse TET1, TET2, and TET3 and the short isoforms of TET1 and TET3. Shown are the N-terminal zinc finger domain (CXXC), the cysteine-rich domain (CRD), and the double-stranded beta helix (DSBH), which is interrupted by a low complexity insert and harbors the Fe(II)- and 2-oxoglutarate binding sites. The respective chromosomal locations are given in brackets next to the names. Protein length in amino acid (AA) and the predicted size in kilo Dalton (kDa) are given next to the structures. Protein domain graphs were generated using DOG 1.0 [213].

family members share, the domain composition of MBD proteins is highly specific for each member, conferring them distinct DNA and protein binding specificities and functions. Hereafter, we will focus on the best studied members of the family, MBD1-4 and MeCP2.

MeCP2

MeCP2 was the first protein described to selectively bind symmetrically methylated CpG dinucleotides [88], and its MBD domain was later used to identify further MBD domain containing proteins [86]. Alternative splicing leads to two MeCP2 isoforms (MeCP2 e1 and MeCP2 e2) that differ in their N-terminus [89] and are mainly considered to have similar functions although they exhibit different expression patterns in developing and postnatal mice brain [90,91]. Both variants contain an N-terminal MBD domain mediating specific binding of methylated CpGs [45] and a transcriptional repression domain (TRD) involved in transcriptional silencing [88,92]. MeCP2 DNA binding, however, is not merely mediated via its MBD domain. Regions outside this domain, including the TRD, the intervening domain (ID), and the C-terminal domain alpha, also contribute to the overall sequence unspecific binding of DNA [93,94]. Additional to the interactions with DNA, MeCP2 also harbors a chromatin interaction surface that was found to interact with histone H3 [95]. Besides the MeCP2-DNA interactions, various protein-protein interactions (reviewed in Ref. [13]), as for example, complexes with histone deacetylase 2 (HDAC2) and the transcriptional co-repressor SIN3A [96] as well as homo- and hetero-interactions of MeCP2 and MBD2 [97], were demonstrated and shown to establish and keep a repressive heterochromatin state in a dose-dependent manner [98]. MeCP2 is highly expressed in the central nervous system, especially in postmitotic neurons [99,100]. *MeCP2*$^{-/-}$ mice show a Rett syndrome like phenotype with reduced brain and neuronal cell size [101,102]. Various deletions and mutations within the *MeCP2* gene were found in Rett syndrome patients (reviewed in Ref. [13]).

MBD1

MBD1 is the largest MBD protein family member and has, due to its unique structure, particular binding abilities and specific functions in gene regulation. Besides its N-terminal MBD and C-terminal TRD domains, MBD1 contains, isoform-dependent, two or three CXXC zinc finger motifs [85,103,104]. Yeast-2-hybrid and co-immunoprecipitation assays revealed interactions with a variety of chromatin-associated proteins involved in histone modification and replication dependent assembly, DNA transcription and repair, polycomb repressive complex and sumoylation [105–115]. The isoform containing the first two CXXC motifs (CXXC1 and CXXC2) [103] can bind methylated CpGs via its MBD domain [116]. The presence of the third CXXC domain (CXXC3), which is homologous to, among others, the zinc finger of DNMT1, was shown to bind unmethylated CpGs [103,117], and enables MBD1 to also bind to unmodified DNA [103,118]. This dual-affinity to methylated and nonmodified DNA of the isoform containing three zinc fingers (MBD1a) enables MBD1 to regulate transcriptional repression regardless of the modification state of the DNA, whereas the isoform lacking the CXXC3 (MBD1b) can only repress expression of highly methylated genes [103,119,120]. The third zinc finger could thus facilitate MBD1 binding in genomes with low DNA methylation, as in preimplantation embryos or primordial germ cells (PGCs). Studies in fibroblast cells showed that the CXXC3 targets nonmodified CpGs, that MBD1 can repress the expression of non-methylated genes, that MBD1 harboring all three CXXC fingers can accumulate at hypomethylated constitutive heterochromatin in *Dnmt1*$^{-/-}$ fibroblasts, and that the deletion of CXXC3 drastically reduces localization at heterochromatin in these cells [103,121]. This is in contrast to a ChIP-seq study in mouse ES cells that showed no binding of MBD1 to unmethylated DNA, despite the presence of the CXXC3, and neither mutations of the MBD domain nor loss of DNA methylation could cause binding to nonmodified DNA [122]. This discrepancy may be related to the fact that, in the latter study, the long MBD1 isoform was over-expressed ectopically and constitutively as a fusion protein with BirA biotin ligase. Altogether these findings point towards MBD1 acting in a cell type specific mode and likely dependent on the level and/or ratio of its isoforms. Although MBD1 is expressed in a variety of tissues, high expression levels are especially found in neuronal stem cells (NSCs) and neurons. MBD1 plays an important role in the maintenance of NSC integrity and stemness and its deficiency in mice leads to accumulation of undifferentiated NSCs and impairs the transition into the neuronal lineage [123]. *Mbd1*$^{-/-}$ mice are viable and develop normally, but have decreased adult neurogenesis and knockout neural stem cells show less neuronal differentiation as well as increased genomic instability [124].

MBD2

MBD2 and MBD3 are the only MBD protein family members that share high sequence similarity also outside the MBD domain, suggesting an ancestral gene duplication [85,86]. MBD2 exists in three isoforms, MBD2a, MBD2b, and MBD2c (also referred to as MBD2t), which all contain the MBD

domain. MBD2a and MBD2b arise from different translational start sites and vary only in the inclusion of an N-terminal glycine and arginine (GR) repeat in MBD2a [85,125]. Both variants contain the C-terminal TRD domain, which is essential for interactions with corepressor complexes, such as the nucleosome remodeling and histone deacetylase (NuRD) or the SIN3A complex [126–129]. MBD2c includes an alternative exon, which results in a truncated protein lacking the TRD domain due to a premature stop codon [85]. This isoform is solely expressed in testes and ES cells and does not interact with the NuRD complex, due to the lack of the TRD domain [85,122,130]. Arginine methylation within the GR motif by protein arginine methyltransferase 1 and 5 (PRMT1 and PRMT5) [131,132] regulates MBD2-DNA and MBD2-HDAC repressor complex formation and reduces the MBD2 transcriptional repression action [132]. *MBD2*$^{-/-}$ mice are viable and fertile and, contrary to *Mecp2*$^{-/-}$ and *Mbd1*$^{-/-}$ mice, show only a mild phenotype, including pup nurturing deficits and hypoactivity, suggesting that the impact on brain function is smaller than that of MBD1 or MeCP2 [133,134]. Alternative splicing and the resulting switch between MBD2a and MBD2c is essential for self-renewal in human pluripotent stem cells (hPSCs). While overexpression of the long isoform MBD2a in hPSCs interferes with pluripotency and promotes differentiation, potentially through downregulating NANOG and OCT4 expression via the NuRD complex, overexpressing MBD2c, which shows no interaction with the NuRD complex, does not influence hPSCs. Of note, overexpressing MBD2c in combination with reprogramming factors in fibroblasts enhanced reprogramming and the formation of induced pluripotent stem cells [130].

MBD3

MBD3 is the smallest member of the MBD protein family, yet three isoforms are described. Besides the N-terminal MBD domain, it contains a C-terminal poly-glutamate stretch [85]. The major difference to MBD2 is that MBD3 has two crucial amino acid exchanges in its MBD domain, which impair the binding of methylated DNA. Binding to 5hmC was proposed, is however still debated [12,122,135–138]. Similar to MBD2, MBD3 is also part of the NuRD complex; however, biochemical analyses showed that they are mutually exclusive within the complex [131,139,140]. MBD3 deletion leads to early embryonic lethality in mice [133] and stem cells lacking MBD3 fail to differentiate and show aberrant self-renewal, even in the absence of leukemia inhibitory factor (LIF) [141,142]. MBD3 depletion in somatic cells together with expression of the reprogramming factors OCT4, SOX2, KLF4, and c-MYC enhanced reprogramming efficiency [143,144], and accordingly MBD3 was proposed to play an important role in pluripotency and lineage commitment [137,145,146]. Nevertheless, opposing results were found when neural and epiblast-derived stem cells were used [147], indicating a context-dependent influence of MBD3 on pluripotency.

MBD4

MBD4 (also known as MED1) is associated with specific functions due to its unique domain composition. Besides its N-terminal MBD domain, MBD4 contains a C-terminal glycosylase motif, unique among MBD proteins [85]. The MBD enables MBD4 binding to methylated CpG as well as mCAC sites for the human variant [56]; however, it was found to also bind 5hmC in neural progenitor cells [12]. The glycosylase domain of MBD4 is similar to the one found in 3-methyladenine DNA glycosylase II (AlkA) [148], recognizes mCG/TG mismatches, and shows enzymatic activity towards T/G or U/G mismatches that occur via deamination of 5mC or cytosine, respectively [125,149]. Base removal via MBD4 relies on several steps, including a base flipping mechanism and, eventually, repair of the abasic site via BER proteins [148,150,151]. MBD4 null mice are viable and fertile, but they show increased number of CpG to TpG transitions [152]. Although MBD4 was shown to excise 5hmU *in vitro* and was therefore implicated in DNA demethylation [153], oxidation-dependent reactivation of methylated reporter genes was only possible in the presence of TDG but not MBD4 [154].

DNA methylation and binding of MBD readers

In line with *in vitro* binding assays [85,88], genome-wide mapping of MBD protein binding in mouse ES and differentiated cells showed a general linear dependence of binding and methylation density for all MBDs except MBD3, as well as the requirement of a functional MBD domain and methylated CGs, CAs, and CACs [56,155]. For some MBDs, additional sequence preference has been reported, with high affinity binding of MeCP2 relying upon an A/T run adjacent to the mCpG [156] and the MBD1 MBD shown to preferentially bind to methylated CpGs within T^{m}CGCA and TG^{m}CGCA sequences [157]. These prerequisites target MBD1/2/4 and MeCP2 to methylated, regulatory, and inactive regions of the genome and, conformingly, *de novo* methylated CpG-rich regions gain MBD binding during differentiation [122]. Still, substantial binding to unmethylated CpG islands was found for MBD2 and MBD4 [122] and DNA binding of MeCP2 was previously shown to not only rely on DNA

methylation [93,155,158,159]. Sequence unspecific and unmethylated DNA binding is described to be mediated by the respective additional domains or protein-protein interactions. An example for an additional domain that confers unmethylated DNA binding is the isoform-specific CXXC3 domain of MBD1 as mentioned before [103]. MBD3 distribution is highly controversial, as binding downstream of CpG-rich and hydroxymethylated promoters, as well as CpG density and (hydroxy)methylation independent binding of regulatory regions were reported [122,137]. Some studies also report binding of MeCP2 as well as MBD4 to 5hmC [12,137,160], but this conclusion has not been validated by other studies [94,99,153].

Microscopic analysis in tissues and in cultured wild type and DNMT deficient cells of the subnuclear distribution of MBDs relative to DNA distribution, compaction, and methylation unanimously reported their enrichment at highly methylated and compacted DNA, with the exception of MBD3 [85,92,98,103,121,158]. These types of analyses by their nature do not exclude tandem or other repeat DNA sequences, which are though mostly non- or under-represented in genomics studies.

TET modifiers and oxidized cytosine variants

Ten-eleven-translocation proteins, oxidative cytosine variants, and DNA demethylation

The 5-methylcytosine mark was for long regarded to be a relatively stable epigenetic mark, due to the chemical nature of the C-C bond between the pyrimidine C5 and the methyl group carbon. Therefore, DNA demethylation was believed to occur via replication-dependent dilution either upon inhibiting DNA methyltransferases or in their complete absence. This hypothesis, however, had to be reviewed when a replication-independent demethylation wave was found in mouse zygotes, where the paternal genome was shown to experience an almost complete loss of the methyl marks before the first round of DNA replication [161,162]. Several attempts to identify the methyl cytosine demethylase were followed by intense debate and not confirmed in other laboratories. A major breakthrough in the field of active loss of DNA methylation was achieved in 2009, when an additional modified cytosine base, 5hmC, and the corresponding writer enzymes were discovered [163,164]. Besides the sixth base 5hmC, three human homologs of trypanosomal 2-oxoglutarate- and iron(II)-dependent dioxygenases JBP1 and JBP2 were identified, named TET1, TET2, and TET3 (Fig. 2B). The acronym TET stands for "Ten-Eleven Translocation" and derives from a t(10;11)(q22;q23) translocation that generates a MLL-TET1 fusion protein, which is found in acute myeloid leukemia [165,166]. TETs were shown to catalyze the oxidation reaction from 5mC to 5hmC in cultured cells as well as *in vivo* [164]. Subsequently, TET proteins were shown to further oxidize 5hmC to 5fC and 5caC [8,167,168]. After their discovery, TET proteins were positioned in the context of active DNA demethylation, assuming that the demethylation reaction relies on the stepwise oxidation of 5mC to 5caC by TETs, followed by thymine DNA glycosylase (TDG) or Nei-like 1 (NEIL1) glycosylase mediated excision of 5fC or 5caC [9,154,169–171]. The resulting abasic site is subsequently targeted by base-excision repair (BER) proteins, leading to the insertion of a nonmodified cytosine [167,172]. TET enzymes and oxidative cytosine variants are suggested to play an essential role in major demethylation cascades during embryonic development [162,173–179] and accordingly, the loss of functional TET and TDG enzymes impairs mouse ES cell differentiation and is embryonically lethal in mice [101,102]. Recent reports, however, suggest that 5hmC and 5fC can also act as stable epigenetic marks, as opposed to being mere DNA demethylation intermediates [9–11,171]. Quantification of genomic cytosine modifications revealed drastic differences in the abundance of the different derivatives. Whereas methylated cytosines account for ~4.5% of all cytosines, 5hmC levels are ten-fold lower, with 0.45% of all cytosines in average in brain tissue and 0.39% in mouse embryonic stem (ES) cells [8,168,180]. In mouse ES cells, 5fC levels are 10 to 100-fold lower than 5hmC levels, only 0.02% to 0.002% of all cytosines carry the formyl mark [8,168] and 5caC is even less abundant, accounting for only 0.0002% of all cytosines [8].

Genomic distribution of oxidative cytosine variants

Genome-wide studies in ES cells and brain tissues found 5hmC to be enriched in euchromatic chromatin regions and especially at/around transcriptional start sites, promoters with moderate and low CpG content and gene bodies, where the amount is positively correlated to gene expression [181–185]. In ES cells, the oxidative mark is widespread but predominantly found at developmentally regulated bivalent genes that are polycomb repressive complex 2 (PRC2) targets, near transcription factor binding sites, at gene enhancers and CTCF binding sites as well as DNAse I hypersensitive chromatin [186–188]. In the brain, hydroxymethylated cytosines are more enriched at poised enhancers than at active enhancers [189]. Similar to 5hmC, 5fC is found at promoters of actively transcribed genes, H3K4me3 and RNA polymerase II marked transcriptional start sites, exons, and poised enhancers in ES cells [190–192].

TET isoforms

A structural feature that separates the three TET proteins from one another and finetunes their respective biological function in different developmental stages and tissues is their N-terminal CXXC domain. Although the zinc finger domains of TET1 and TET3 as well as the TET2 associated IDAX/CXXC4 belong to the same subgroup of CXXC-domains [193], they highly differ in their respective binding affinities and substrates. TET2 lost its zinc finger during evolution in a chromosomal inversion, and it is now encoded by the genomically adjacent IDAX/CXXC4, which negatively regulates TET2 expression and modulates its catalytic activity [194]. TET1 and TET3, in contrast, kept their respective zinc finger domains and the CXXC domain of TET3 was shown to preferentially bind 5caC [195]. However, TET1 CXXC domain was found to bind mostly nonmodified DNA [117], which was implicated in the prevention of DNA methylation spreading into unmethylated, euchromatic regions [196]. While three different isoforms of TET2 and also TET3 have been characterized to date [195,197], an N-terminally truncated TET1 isoform, which lacks the CXXC domain, was described recently and attributed a role in mouse development, but also in cancerogenesis [198,199].

Interplay of DNA methylation readers and modifiers

Cytosine methylation writers, readers, and modifiers need a tight spatio-temporal regulation to avoid aberrant DNA methylation, a hallmark of numerous human disorders. Deregulation of members of the DNMT, MBD, or TET protein families and their interplay can impair the genomic methylome, as well as chromatin structure and organization, as observed in patients suffering from the neurodegenerative HSANIE disorder (hereditary sensory and autonomic neuropathy, mutations in the TS domain of DNMT1), the Rett syndrome (MeCP2 mutations), or AML (acute myeloid leukemia, TET2 mutations/deletions) [98,200–202]. Shaping of DNA modification patterns can be regulated by the localization, activity, stability, interactome, substrate preference, and cofactor availability of the writer and modifier enzymes [168,203–209]. As previously discussed, the distribution of DNA modifications throughout the genome is not random but tightly associated with highly specific genomic regions and concomitantly presenting distinct protein recognition platforms. Consequently, the binding profile of cytosine modification reader proteins is greatly influenced and defined by the genomic landscape of the modifications and the latter is proposed to shape the genomic location of the readers. Indeed, the catalytic activity of TET proteins was found to affect the binding of different members of the MBD protein family [12]. However, since MBD readers and TET modifiers share (at least) one common substrate, the opposite role can also be envisaged: recognition and binding of 5mC (and/or 5hmC) by MBDs modulate the activity of TET proteins and in turn the epigenetic makeup of DNA, its physical properties, and hence DNA metabolism. According to this model, the reader proteins shape the DNA modification distribution via regulation of the accessibility of the modified DNA bases to modifier enzymes.

Hashimoto and coworkers found by *in vitro* inhibition experiments that binding of methylated DNA by MeCP2 inhibits TET1 activity, thereby suggesting a potential mechanism for 5hmC level regulation *in vivo* [153]. Another study suggested that MeCP2 binding of 5mC itself was correlated with lack of 5mC oxidation [210]. Recently, a more indepth analysis on the interplay of 5mC, MBD, and TET proteins interrogated whether and how binding of MBD proteins to a methylated substrate influences TET-mediated DNA demethylation. Interestingly and in line with the respective substrate preference, especially the binding of the MBD protein family members MeCP2 and MBD2 was found to protect 5mC from TET1 mediated oxidation. Importantly, quantification of the TET1 level demonstrated that it was similar to the physiological levels and the same was the case for MeCP2 levels that mimicked its concentration in neuronal cells. Hence, the conditions used reflected the physiological situation. This blocking of TET-mediated 5mC oxidation was MBD protein concentration dependent and relied both on MeCP2 and MBD2 binding 5mC directly via their MBD or, in the case of MeCP2, additionally by its binding to nonmodified DNA (Fig. 3). Thus, protection of TET1-dependent 5mC oxidation arose from direct blocking of access to the DNA as well as 5mC.

To ensure genomic stability, minimize transcriptional noise of noncoding genomic repeats (e.g., major satellite repeats, MaSat), and insertions of mobile retrotransposons (e.g., long interspersed nuclear elements, LINEs), cells developed a variety of mechanisms to silence the expression of these elements, the most important being cytosine methylation. MBD protein–mediated blocking of 5mC oxidation was shown to prevent the reactivation of silenced repetitive elements such as major satellite repeats or LINE elements, thereby maintaining chromatin composition, as well as genome stability [158,211]. Accordingly, chromocentric 5hmC levels in the pons of a mouse model for Rett syndrome ($MeCP2^{y/-}$) were significantly increased and *in situ* RNA-FISH as well as quantitative PCR analyses revealed higher levels of major satellite expression compared to wildtype. This suggested that

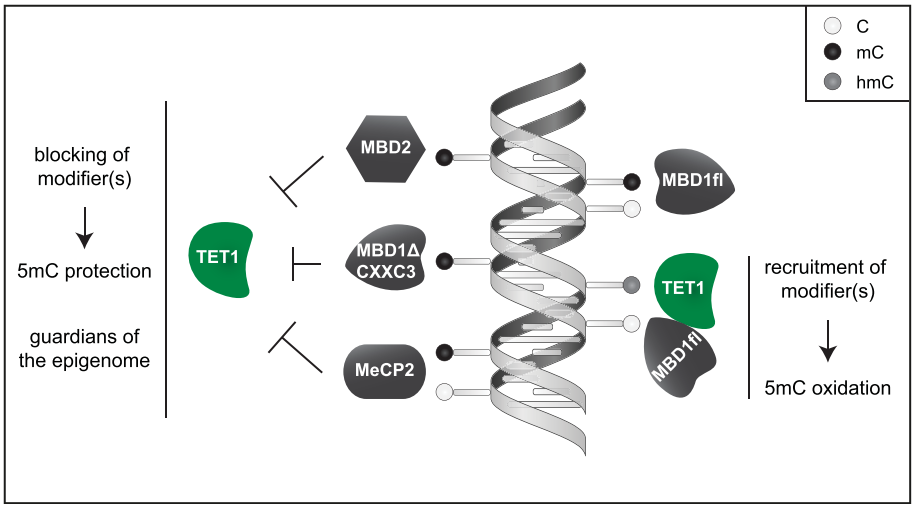

Fig. 3. Role of MBD proteins in modifying and protecting cytosine modifications. Graphical summary of the specific DNA binding abilities of MBD proteins and their influence on the DNA base modifier TET1. Whereas MBD1 full-length (fl) recruits TET1 to methylated cytosine thereby enhancing 5-hydroxymethylcytosine generation, MeCP2, MBD2, and MBD1ΔCXXC3 bind 5-methylcytosine, blocking TET1 access and protecting 5-methylcytosine from oxidation. C = cytosine, mC = 5-methylcytosine, hmC = 5-hydroxymethylcytosine.

unrestricted TET1 activity might increase repeat element transcription in animal models of Rett syndrome. Thus, these studies disclose the potential of a pathophysiological pathway of TET enzymes in Rett syndrome, as well as a role of MeCP2 and MBD2 as guardians of the epigenome and in genome stability.

In contrast to MeCP2 and MBD2, MBD3 and MBD4 were not found to influence TET1-mediated 5hmC formation. On the other hand, MBD1 was shown to enhance TET1, but not TET2 and TET3, mediated 5mC oxidation. The underlying mechanism relies on an MBD1-dependent increase of TET1 localization at highly methylated heterochromatin, and this recruitment is favored by protein-protein interactions between MBD1 and TET1 [121]. Interestingly, the recruitment is cell-cycle stage independent. Akin to the blocking effect observed upon MeCP2 or MBD2 binding to 5mC, the MBD1-dependent TET1 recruitment effect observed in low TET1 overexpressing cells reflects physiological situation comparable to mouse ES cells. Recruitment and catalytic activity of TET1 at heterochromatin ultimately leads to the displacement of MBD1, as 5hmC is a less favored binding substrate relative to 5mC. Interestingly, MBD1-mediated enhancement of TET1 localization at heterochromatin as well as 5mC oxidation is MBD1 isoform dependent (Fig. 2A). In contrast to the MBD1 isoform containing all three CXXC domains (MBD1 fl or MBD1a), the short isoform lacking the third zinc finger (MBD1ΔCXXC3 or MBD1b) blocks TET1-mediated 5hmC formation, suggesting different biological functions of the MBD1 isoforms (Fig. 3). This CXXC3 dependency might indicate that the latter targets MBD1 to CpG-dense regions via binding to unmethylated CpGs, recruiting TET1 without simultaneously blocking its ability to oxidize methylated DNA substrates [121].

Concluding remarks/Perspectives

Altogether these studies unveil new mechanisms of a time and dose-dependent cross-regulation of cytosine modifications and their readers and modifiers, relying on inhibiting each other's binding. Until now, DNA modifications are generally seen as shapers of the reader's binding profiles, and, hence, their genomic localization [160,212]. Conversely, cytosine modification readers do not only play important roles in translating the modifications into distinct chromatin states, thereby regulating transcription and conserving genome integrity, but also in shaping the genomic modification landscape by protecting from or enhancing oxidation.

Acknowledgments

We thank all the past and present members of our laboratory for their many contributions along the years. Last but not least, we thank our collaborators over the years, who have made our work so much more enjoyable.

Author contributions

Florian D. Hastert, Cathia Rausch: Conceptualization, Writing—Original Draft Preparation, Visualization.

M. Cristina Cardoso: Conceptualization, Writing—Reviewing and Editing, Funding Acquisition.

Declarations of interest

None.

Funding

Our research has been supported by grants of the German Research Foundation, Germany (DFG CA 198/12-1 to M.C.C.).

Received 14 September 2019;
Received in revised form 28 November 2019;
Accepted 5 December 2019
Available online 20 December 2019

Keywords:
chromatin;
DNA helix stability;
DNA methylation readers;
DNA modifications;
DNA modification writers

Abbreviations used:

5caC, 5-carboxylcytosine; 5fC, 5-formylcytosine; 5hmC, 5-hydroxymethylcytosine; 5mC, 5-methylcytosine; C, Cytosine; CGI, CpG island; CpG, Cytosine-phosphate-guanine dinucleotide; CXXC, Cysteine/X/X/cysteine zinc finger motif; DNMT, DNA methyltransferase; ES cells, Embryonic stem cells; ID, Intervening domain; MBD, Methyl-CpG-binding domain; TET, Ten-eleven translocation; T_m, Melting temperature; TRD, Transcriptional repression domain.

References

[1] R.D. Hotchkiss, The quantitative separation of purines, pyrimidines, and nucleosides by paper chromatography, J. Biol. Chem. 175 (1948) 315–332.
[2] N.W. Penn, R. Suwalski, C. O'Riley, K. Bojanowski, R. Yura, The presence of 5-hydroxymethylcytosine in animal deoxyribonucleic acid, Biochem. J. 126 (1972) 781–790.
[3] M. Ehrlich, M.A. Gama-Sosa, L.H. Huang, R.M. Midgett, K.C. Kuo, R.A. McCune, et al., Amount and distribution of 5-methylcytosine in human DNA from different types of tissues of cells, Nucleic Acids Res. 10 (1982) 2709–2721.
[4] A. Bird, DNA methylation patterns and epigenetic memory, Genes Dev. 16 (2002) 6–21.
[5] E. Li, C. Beard, R. Jaenisch, Role for DNA methylation in genomic imprinting, Nature 366 (1993) 362–365.
[6] M.M. Suzuki, A. Bird, DNA methylation landscapes: provocative insights from epigenomics, Nat. Rev. Genet. 9 (2008) 465–476.
[7] C.P. Walsh, J.R. Chaillet, T.H. Bestor, Transcription of IAP endogenous retroviruses is constrained by cytosine methylation, Nat. Genet. 20 (1998) 116–117.
[8] S. Ito, L. Shen, Q. Dai, S.C. Wu, L.B. Collins, J.A. Swenberg, et al., Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine, Science 333 (2011) 1300–1303.
[9] R.M. Kohli, Y. Zhang, TET enzymes, TDG and the dynamics of DNA demethylation, Nature 502 (2013) 472–479.
[10] M. Bachman, S. Uribe-Lewis, X. Yang, H.E. Burgess, M. Iurlaro, W. Reik, et al., 5-Formylcytosine can be a stable DNA modification in mammals, Nat. Chem. Biol. 11 (2015) 555–557.
[11] M. Bachman, S. Uribe-Lewis, X. Yang, M. Williams, A. Murrell, S. Balasubramanian, 5-Hydroxymethylcytosine is a predominantly stable DNA modification, Nat. Chem. 6 (2014) 1049–1055.
[12] C.G. Spruijt, F. Gnerlich, A.H. Smits, T. Pfaffeneder, P.W. Jansen, C. Bauer, et al., Dynamic readers for 5-(hydroxy)methylcytosine and its oxidized derivatives, Cell 152 (2013) 1146–1159.
[13] A.K. Ludwig, P. Zhang, M.C. Cardoso, Modifiers and readers of DNA modifications and their impact on genome structure, expression, and stability in disease, Front. Genet. 7 (2016) 115.
[14] M.J. Koziol, C.R. Bradshaw, G.E. Allen, A.S.H. Costa, C. Frezza, J.B. Gurdon, Identification of methylated deoxyadenosines in vertebrates reveals diversity in DNA modifications, Nat. Struct. Mol. Biol. 23 (2016) 24–30.
[15] T.P. Wu, T. Wang, M.G. Seetin, Y. Lai, S. Zhu, K. Lin, et al., DNA methylation on N(6)-adenine in mammalian embryonic stem cells, Nature 532 (2016) 329–333.
[16] G.Z. Luo, Z. Hao, L. Luo, M. Shen, D. Sparvoli, Y. Zheng, et al., N(6)-methyldeoxyadenosine directs nucleosome positioning in Tetrahymena DNA, Genome Biol. 19 (2018) 200.
[17] S. Schiffers, C. Ebert, R. Rahimoff, O. Kosmatchev, J. Steinbacher, A.V. Bohne, et al., Quantitative LC-MS provides No evidence for m6 dA or m4 dC in the genome of mouse embryonic stem cells and tissues, Angew. Chem. Int. Ed. 56 (2017) 11268–11271.
[18] A.A. Travers, The structural basis of DNA flexibility, Philos. Trans. Ser. Math. Phys. Eng. Sci. 362 (2004) 1423–1438.
[19] J. Marmur, P. Doty, Heterogeneity in deoxyribonucleic acids. I. Dependence on composition of the configurational stability of deoxyribonucleic acids, Nature 183 (1959) 1427–1429.
[20] J.D. Hoheisel, A.G. Craig, H. Lehrach, Effect of 5-bromo- and 5-methyldeoxycytosine on duplex stability and discrimination of the NotI octadeoxynucleotide. Quantitative measurements using thin-layer chromatography, J. Biol. Chem. 265 (1990) 16656–16660.
[21] A. Lefebvre, O. Mauffret, S. el Antri, M. Monnot, E. Lescot, S. Fermandjian, Sequence dependent effects of CpG cytosine methylation. A joint 1H-NMR and 31P-NMR study, Eur. J. Biochem. 229 (1995) 445–454.
[22] C. Acosta-Silva, V. Branchadell, J. Bertran, A. Oliva, Mutual relationship between stacking and hydrogen bonding in DNA. Theoretical study of guanine-cytosine, guanine-5-

methylcytosine, and their dimers, J. Phys. Chem. B 114 (2010) 10217–10227.
[23] M. Banyay, A. Graslund, Structural effects of cytosine methylation on DNA sugar pucker studied by FTIR, J. Mol. Biol. 324 (2002) 667–676.
[24] L. Lercher, M.A. McDonough, A.H. El-Sagheer, A. Thalhammer, S. Kriaucionis, T. Brown, et al., Structural insights into how 5-hydroxymethylation influences transcription factor binding, Chem. Commun. 50 (2014) 1794–1796.
[25] C.M. Lopez, A.J. Lloyd, K. Leonard, M.J. Wilkinson, Differential effect of three base modifications on DNA thermostability revealed by high resolution melting, Anal. Chem. 84 (2012) 7336–7342.
[26] E.A. Raiber, P. Murat, D.Y. Chirgadze, D. Beraldi, B.F. Luisi, S. Balasubramanian, 5-Formylcytosine alters the structure of the DNA double helix, Nat. Struct. Mol. Biol. 22 (2015) 44–49.
[27] D. Renciuk, O. Blacque, M. Vorlickova, B. Spingler, Crystal structures of B-DNA dodecamer containing the epigenetic modifications 5-hydroxymethylcytosine or 5-methylcytosine, Nucleic Acids Res. 41 (2013) 9891–9900.
[28] P.M. Severin, X. Zou, H.E. Gaub, K. Schulten, Cytosine methylation alters DNA mechanical properties, Nucleic Acids Res. 39 (2011) 8740–8751.
[29] P.M. Severin, X. Zou, K. Schulten, H.E. Gaub, Effects of cytosine hydroxymethylation on DNA strand separation, Biophys. J. 104 (2013) 208–215.
[30] Q. Song, Z. Qiu, H. Wang, Y. Xia, J. Shen, Y. Zhang, The effect of methylation on the hydrogen-bonding and stacking interaction of nucleic acid bases, Struct. Chem. 24 (2013) 55–65.
[31] A. Thalhammer, A.S. Hansen, A.H. El-Sagheer, T. Brown, C.J. Schofield, Hydroxylation of methylated CpG dinucleotides reverses stabilisation of DNA duplexes by cytosine 5-methylation, Chem. Commun. 47 (2011) 5325–5327.
[32] M. Wanunu, D. Cohen-Karni, R.R. Johnson, L. Fields, J. Benner, N. Peterman, et al., Discrimination of methylcytosine from hydroxymethylcytosine in DNA molecules, J. Am. Chem. Soc. 133 (2011) 486–492.
[33] P. Zhang, F.D. Hastert, A.K. Ludwig, K. Breitwieser, M. Hofstätter, M.C. Cardoso, DNA base flipping analytical pipeline, Biol. Methods Protoc. 2 (2017) 1–12.
[34] A.T.P. Carvalho, M.L. Gouveia, C. Raju Kanna, S. Warmlander, J. Platts, S.C.L. Kamerlin, Theoretical modelling of epigenetically modified DNA sequences, F1000Research 4 (2015) 52.
[35] J.S. Hardwick, D. Ptchelkine, A.H. El-Sagheer, I. Tear, D. Singleton, S.E.V. Phillips, et al., 5-Formylcytosine does not change the global structure of DNA, Nat. Struct. Mol. Biol. 24 (2017) 544–552.
[36] A. Moser, R. Guza, N. Tretyakova, D.M. York, Density functional study of the influence of C5 cytosine substitution in base pairs with guanine, Theor. Chem. Acc. 122 (2009) 179–188.
[37] A. Perez, C.L. Castellazzi, F. Battistini, K. Collinet, O. Flores, O. Deniz, et al., Impact of methylation on the physical properties of DNA, Biophys. J. 102 (2012) 2140–2148.
[38] A.T. Carvalho, L. Gouveia, C.R. Kanna, S.K. Warmlander, J.A. Platts, S.C. Kamerlin, Understanding the structural and dynamic consequences of DNA epigenetic modifications: computational insights into cytosine methylation and hydroxymethylation, Epigenetics 9 (2014) 1604–1612.
[39] J. Horberg, A. Reymer, A sequence environment modulates the impact of methylation on the torsional rigidity of DNA, Chem. Commun. 54 (2018) 11885–11888.
[40] S. Rao, T.P. Chiu, J.F. Kribelbauer, R.S. Mann, H.J. Bussemaker, R. Rohs, Systematic prediction of DNA shape changes due to CpG methylation explains epigenetic effects on protein-DNA binding, Epigenet. Chromatin 11 (2018) 6.
[41] A. Mendonca, E.H. Chang, W. Liu, C. Yuan, Hydroxymethylation of DNA influences nucleosomal conformation and stability in vitro, Biochim. Biophys. Acta 1839 (2014) 1323–1329.
[42] T.T. Ngo, J. Yoo, Q. Dai, Q. Zhang, C. He, A. Aksimentiev, et al., Effects of cytosine modifications on DNA flexibility and nucleosome mechanical stability, Nat. Commun. 7 (2016) 10813.
[43] A. Jeltsch, Beyond Watson and Crick: DNA methylation and molecular enzymology of DNA methyltransferases, Chembiochem : Euro. J. Chem. Biol. 3 (2002) 274–293.
[44] A.C. Dantas Machado, T. Zhou, S. Rao, P. Goel, C. Rastogi, A. Lazarovici, et al., Evolving insights on how cytosine methylation affects protein-DNA binding, Briefings Funct. Genom. 14 (2015) 61–73.
[45] T. Fu, L. Liu, Q.L. Yang, Y. Wang, P. Xu, L. Zhang, et al., Thymine DNA glycosylase recognizes the geometry alteration of minor grooves induced by 5-formylcytosine and 5-carboxylcytosine, Chem. Sci. 10 (2019) 7407–7417.
[46] J. Norberg, M. Vihinen, Molecular dynamics simulation of the effects of cytosine methylation on structure of oligonucleotides, J. Mol. Struct. THEOCHEM 546 (2001) 51–62.
[47] L.C. Sowers, B.R. Shaw, W.D. Sedwick, Base stacking and molecular polarizability: effect of a methyl group in the 5-position of pyrimidines, Biochem. Biophys. Res. Commun. 148 (1987) 790–794.
[48] M.W. Szulik, P.S. Pallan, B. Nocek, M. Voehler, S. Banerjee, S. Brooks, et al., Differential stabilities and sequence-dependent base pair opening dynamics of Watson-Crick base pairs with 5-hydroxymethylcytosine, 5-formylcytosine, or 5-carboxylcytosine, Biochemistry 54 (2015) 1294–1305.
[49] M.B. Helabad, N. Kanaan, P. Imhof, Base flip in DNA studied by molecular dynamics simulationsof differently-oxidized forms of methyl-Cytosine, Int. J. Mol. Sci. 15 (2014) 11799–11816.
[50] Q. Dai, P.J. Sanstead, C.S. Peng, D. Han, C. He, A. Tokmakoff, Weakened N3 hydrogen bonding by 5-formylcytosine and 5-carboxylcytosine reduces their base-pairing stability, ACS Chem. Biol. 11 (2016) 470–477.
[51] A. Lazarovici, T. Zhou, A. Shafer, A.C. Dantas Machado, T.R. Riley, R. Sandstrom, et al., Probing DNA shape and methylation state on a genomic scale with DNase I, Proc. Natl. Acad. Sci. U. S. A. 110 (2013) 6376–6381.
[52] Y. Liu, X. Zhang, R.M. Blumenthal, X. Cheng, A common mode of recognition for methylated CpG, Trends Biochem. Sci. 38 (2013) 177–183.
[53] M. Kitayner, H. Rozenberg, R. Rohs, O. Suad, D. Rabinovich, B. Honig, et al., Diversity in DNA recognition by p53 revealed by crystal structures with Hoogsteen base pairs, Nat. Struct. Mol. Biol. 17 (2010) 423–429.
[54] K.L. Ho, I.W. McNae, L. Schmiedeberg, R.J. Klose, A.P. Bird, M.D. Walkinshaw, MeCP2 binding to DNA depends upon hydration at methyl-CpG, Mol. Cell 29 (2008) 525–531.

[55] R. Rohs, S.M. West, A. Sosinsky, P. Liu, R.S. Mann, B. Honig, The role of DNA shape in protein-DNA recognition, Nature 461 (2009) 1248–1253.
[56] K. Liu, C. Xu, M. Lei, A. Yang, P. Loppnau, T.R. Hughes, et al., Structural basis for the ability of MBD domains to bind methyl-CG and TG sites in DNA, J. Biol. Chem. 293 (2018) 7344–7354.
[57] I. Ohki, N. Shimotake, N. Fujita, J. Jee, T. Ikegami, M. Nakao, et al., Solution structure of the methyl-CpG binding domain of human MBD1 in complex with methylated DNA, Cell 105 (2001) 487–497.
[58] M.J. Lyst, J. Connelly, C. Merusi, A. Bird, Sequence-specific DNA binding by AT-hook motifs in MeCP2, FEBS Lett. 590 (2016) 2927–2933.
[59] L.S. Chuang, H.I. Ian, T.W. Koh, H.H. Ng, G. Xu, B.F. Li, Human DNA-(cytosine-5) methyltransferase-PCNA complex as a target for p21WAF1, Science 277 (1997) 1996–2000.
[60] H. Leonhardt, A.W. Page, H.U. Weier, T.H. Bestor, A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei, Cell 71 (1992) 865–873.
[61] O. Mortusewicz, L. Schermelleh, J. Walter, M.C. Cardoso, H. Leonhardt, Recruitment of DNA methyltransferase I to DNA repair sites, Proc. Natl. Acad. Sci. U. S. A. 102 (2005) 8905–8909.
[62] M. Okano, D.W. Bell, D.A. Haber, E. Li, DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development, Cell 99 (1999) 247–257.
[63] M. Kaneda, M. Okano, K. Hata, T. Sado, N. Tsujimoto, E. Li, et al., Essential role for de novo DNA methyltransferase Dnmt3a in paternal and maternal imprinting, Nature 429 (2004) 900–903.
[64] Z.D. Smith, A. Meissner, DNA methylation: roles in mammalian development, Nat. Rev. Genet. 14 (2013) 204–220.
[65] A.P. Bird, DNA methylation and the frequency of CpG in animal DNA, Nucleic Acids Res. 8 (1980) 1499–1504.
[66] D.N. Cooper, M. Krawczak, Cytosine methylation and the fate of CpG dinucleotides in vertebrate genomes, Hum. Genet. 83 (1989) 181–188.
[67] C. Coulondre, J.H. Miller, P.J. Farabaugh, W. Gilbert, Molecular basis of base substitution hotspots in Escherichia coli, Nature 274 (1978) 775–780.
[68] K.D. Robertson, A.P. Wolffe, DNA methylation in health and disease, Nat. Rev. Genet. 1 (2000) 11–19.
[69] Y. Zheng, B.T. Joyce, L. Liu, Z. Zhang, W.A. Kibbe, W. Zhang, et al., Prediction of genome-wide DNA methylation in repetitive elements, Nucleic Acids Res. 45 (2017) 8697–8711.
[70] M.P. Ball, J.B. Li, Y. Gao, J.H. Lee, E.M. LeProust, I.H. Park, et al., Targeted and genome-scale strategies reveal gene-body methylation signatures in human cells, Nat. Biotechnol. 27 (2009) 361–368.
[71] P.A. Jones, Functions of DNA methylation: islands, start sites, gene bodies and beyond, Nat. Rev. Genet. 13 (2012) 484–492.
[72] D. Schubeler, Function and information content of DNA methylation, Nature 517 (2015) 321–326.
[73] X. Yang, H. Han, D.D. De Carvalho, F.D. Lay, P.A. Jones, G. Liang, Gene body methylation can alter gene expression and is a therapeutic target in cancer, Cancer Cell 26 (2014) 577–590.
[74] J.H. Lee, S.J. Park, K. Nakai, Differential landscape of non-CpG methylation in embryonic stem cells and neurons caused by DNMT3s, Sci. Rep. 7 (2017) 11295.
[75] R. Lister, E.A. Mukamel, J.R. Nery, M. Urich, C.A. Puddifoot, N.D. Johnson, et al., Global epigenomic reconfiguration during mammalian brain development, Science 341 (2013) 1237905.
[76] R. Lister, M. Pelizzola, R.H. Dowen, R.D. Hawkins, G. Hon, J. Tonti-Filippini, et al., Human DNA methylomes at base resolution show widespread epigenomic differences, Nature 462 (2009) 315–322.
[77] M.D. Schultz, Y. He, J.W. Whitaker, M. Hariharan, E.A. Mukamel, D. Leung, et al., Human body epigenome maps reveal noncanonical DNA methylation variation, Nature 523 (2015) 212–216.
[78] M.J. Ziller, F. Muller, J. Liao, Y. Zhang, H. Gu, C. Bock, et al., Genomic distribution and inter-sample variation of non-CpG methylation across human cell types, PLoS Genet. 7 (2011) e1002389.
[79] H.W. Gabel, B. Kinde, H. Stroud, C.S. Gilbert, D.A. Harmin, N.R. Kastan, et al., Disruption of DNA-methylation-dependent long gene repression in Rett syndrome, Nature 522 (2015) 89–93.
[80] J.U. Guo, Y. Su, J.H. Shin, J. Shin, H. Li, B. Xie, et al., Distribution, recognition and regulation of non-CpG methylation in the adult mammalian brain, Nat. Neurosci. 17 (2014) 215–222.
[81] S. Lagger, J.C. Connelly, G. Schweikert, S. Webb, J. Selfridge, B.H. Ramsahoye, et al., MeCP2 recognizes cytosine methylated tri-nucleotide and di-nucleotide sequences to tune transcription in the mammalian brain, PLoS Genet. 13 (2017), e1006793.
[82] M.J. Sperlazza, S.M. Bilinovich, L.M. Sinanan, F.R. Javier, D.C. Williams Jr., Structural basis of MeCP2 distribution on non-CpG methylated and hydroxymethylated DNA, J. Mol. Biol. 429 (2017) 1581–1594.
[83] H.I. Baymaz, A. Fournier, S. Laget, Z. Ji, P.W. Jansen, A.H. Smits, et al., MBD5 and MBD6 interact with the human PR-DUB complex through their methyl-CpG-binding domain, Proteomics 14 (2014) 2179–2189.
[84] S.H. Cross, R.R. Meehan, X. Nan, A. Bird, A component of the transcriptional repressor MeCP1 shares a motif with DNA methyltransferase and HRX proteins, Nat. Genet. 16 (1997) 256–259.
[85] B. Hendrich, A. Bird, Identification and characterization of a family of mammalian methyl-CpG binding proteins, Mol. Cell. Biol. 18 (1998) 6538–6547.
[86] B. Hendrich, S. Tweedie, The methyl-CpG binding domain and the evolving role of DNA methylation in animals, Trends Genet. : TIG (Trends Genet.) 19 (2003) 269–277.
[87] S. Laget, M. Joulie, F. Le Masson, N. Sasai, E. Christians, S. Pradhan, et al., The human proteins MBD5 and MBD6 associate with heterochromatin but they do not bind methylated DNA, PLoS One 5 (2010) e11982.
[88] J.D. Lewis, R.R. Meehan, W.J. Henzel, I. Maurer-Fogy, P. Jeppesen, F. Klein, et al., Purification, sequence, and cellular localization of a novel chromosomal protein that binds to methylated DNA, Cell 69 (1992) 905–914.
[89] S. Kriaucionis, A. Bird, The major form of MeCP2 has a novel N-terminus generated by alternative splicing, Nucleic Acids Res. 32 (2004) 1818–1823.
[90] A. Krithivas, M. Fujimuro, M. Weidner, D.B. Young, S.D. Hayward, Protein interactions targeting the latency-associated nuclear antigen of Kaposi's sarcoma-associated

herpesvirus to cell chromosomes, J. Virol. 76 (2002) 11596–11604.
[91] Y. Li, H. Wang, J. Muffat, A.W. Cheng, D.A. Orlando, J. Lovén, et al., Global transcriptional and translational repression in human-embryonic-stem-cell-derived Rett syndrome neurons, Cell Stem Cell 13 (2013) 446–458.
[92] X. Nan, F.J. Campoy, A. Bird, MeCP2 is a transcriptional repressor with abundant binding sites in genomic chromatin, Cell 88 (1997) 471–481.
[93] R.P. Ghosh, T. Nikitina, R.A. Horowitz-Scherer, L.M. Gierasch, V.N. Uversky, K. Hite, et al., Unique physical properties and interactions of the domains of methylated DNA binding protein 2, Biochemistry 49 (2010) 4395–4410.
[94] V. Valinluck, H.H. Tsai, D.K. Rogstad, A. Burdzy, A. Bird, L.C. Sowers, Oxidative damage to methyl-CpG sequences inhibits the binding of the methyl-CpG binding domain (MBD) of methyl-CpG binding protein 2 (MeCP2), Nucleic Acids Res. 32 (2004) 4100–4108.
[95] T. Nikitina, X. Shi, R.P. Ghosh, R.A. Horowitz-Scherer, J.C. Hansen, C.L. Woodcock, Multiple modes of interaction between the methylated DNA binding protein MeCP2 and chromatin, Mol. Cell. Biol. 27 (2007) 864–877.
[96] P.L. Jones, G.J. Veenstra, P.A. Wade, D. Vermaak, S.U. Kass, N. Landsberger, et al., Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription, Nat. Genet. 19 (1998) 187–191.
[97] A. Becker, L. Allmann, M. Hofstatter, V. Casa, P. Weber, A. Lehmkuhl, et al., Direct homo- and hetero-interactions of MeCP2 and MBD2, PLoS One 8 (2013) e53730.
[98] A. Brero, H.P. Easwaran, D. Nowak, I. Grunewald, T. Cremer, H. Leonhardt, et al., Methyl CpG–binding proteins induce large-scale chromatin reorganization during terminal differentiation, J. Cell Biol. 169 (2005) 733–743.
[99] M. Iurlaro, G. Ficz, D. Oxley, E.A. Raiber, M. Bachman, M.J. Booth, et al., A screen for hydroxymethylcytosine and formylcytosine binding proteins suggests functions in transcription and chromatin regulation, Genome Biol. 14 (2013) R119.
[100] J. Barau, A. Teissandier, N. Zamudio, S. Roy, V. Nalesso, Y. Herault, et al., The DNA methyltransferase DNMT3C protects male germ cells from transposon activity, Science 354 (2016) 909–912.
[101] D. Cortazar, C. Kunz, J. Selfridge, T. Lettieri, Y. Saito, E. MacDougall, et al., Embryonic lethal phenotype reveals a function of TDG in maintaining epigenetic stability, Nature 470 (2011) 419–423.
[102] M.M. Dawlaty, A. Breiling, T. Le, M.I. Barrasa, G. Raddatz, Q. Gao, et al., Loss of Tet enzymes compromises proper differentiation of embryonic stem cells, Dev. Cell 29 (2014) 102–111.
[103] H.F. Jorgensen, I. Ben-Porath, A.P. Bird, Mbd1 is recruited to both methylated and nonmethylated CpGs via distinct DNA binding domains, Mol. Cell. Biol. 24 (2004) 3387–3395.
[104] H.H. Ng, P. Jeppesen, A. Bird, Active repression of methylated genes by the chromosomal protein MBD1, Mol. Cell. Biol. 20 (2000) 1394–1406.
[105] N. Fujita, S. Watanabe, T. Ichimura, Y. Ohkuma, T. Chiba, H. Saya, et al., MCAF mediates MBD1-dependent transcriptional repression, Mol. Cell. Biol. 23 (2003) 2834–2843.
[106] N. Fujita, S. Watanabe, T. Ichimura, S. Tsuruzoe, Y. Shinkai, M. Tachibana, et al., Methyl-CpG binding domain 1 (MBD1) interacts with the Suv39h1-HP1 heterochromatic complex for DNA methylation-based transcriptional repression, J. Biol. Chem. 278 (2003) 24132–24138.
[107] U.F. Hameed, J. Lim, Q. Zhang, M.A. Wasik, D. Yang, K. Swaminathan, Transcriptional repressor domain of MBD1 is intrinsically disordered and interacts with its binding partners in a selective manner, Sci. Rep. 4 (2014) 4896.
[108] T. Ichimura, S. Watanabe, Y. Sakamoto, T. Aoto, N. Fujita, M. Nakao, Transcriptional repression and heterochromatin formation by MBD1 and MCAF/AM family proteins, J. Biol. Chem. 280 (2005) 13928–13935.
[109] M.J. Lyst, X. Nan, I. Stancheva, Regulation of MBD1-mediated transcriptional repression by SUMO and PIAS proteins, EMBO J. 25 (2006) 5317–5328.
[110] B.E. Reese, K.E. Bachman, S.B. Baylin, M.R. Rountree, The methyl-CpG binding protein MBD1 interacts with the p150 subunit of chromatin assembly factor 1, Mol. Cell. Biol. 23 (2003) 3226–3236.
[111] Y. Sakamoto, S. Watanabe, T. Ichimura, M. Kawasuji, H. Koseki, H. Baba, et al., Overlapping roles of the methylated DNA-binding protein MBD1 and polycomb group proteins in transcriptional repression of HOXA genes and heterochromatin foci formation, J. Biol. Chem. 282 (2007) 16391–16400.
[112] Y. Uchimura, T. Ichimura, J. Uwada, T. Tachibana, S. Sugahara, M. Nakao, et al., Involvement of SUMO modification in MBD1- and MCAF1-mediated heterochromatin formation, J. Biol. Chem. 281 (2006) 23180–23190.
[113] R. Villa, L. Morey, V.A. Raker, M. Buschbeck, A. Gutierrez, F. De Santis, et al., The methyl-CpG binding protein MBD1 is required for PML-RARalpha function, Proc. Natl. Acad. Sci. U. S. A. 103 (2006) 1400–1405.
[114] S. Watanabe, T. Ichimura, N. Fujita, S. Tsuruzoe, I. Ohki, M. Shirakawa, et al., Methylated DNA-binding domain 1 and methylpurine-DNA glycosylase link transcriptional repression and DNA repair in chromatin, Proc. Natl. Acad. Sci. U. S. A. 100 (2003) 12859–12864.
[115] J. Xu, W. Zhu, W. Xu, X. Cui, L. Chen, S. Ji, et al., Silencing of MBD1 reverses pancreatic cancer therapy resistance through inhibition of DNA damage repair, Int. J. Oncol. 42 (2013) 2046–2052.
[116] C. Bianchi, R. Zangi, Molecular dynamics study of the recognition of dimethylated CpG sites by MBD1 protein, J. Chem. Inf. Model. 55 (2015) 636–644.
[117] C. Frauer, A. Rottach, D. Meilinger, S. Bultmann, K. Fellinger, S. Hasenoder, et al., Different binding properties and function of CXXC zinc finger domains in Dnmt1 and Tet1, PLoS One 6 (2011) e16627.
[118] M. Nakao, S. Matsui, S. Yamamoto, K. Okumura, M. Shirakawa, N. Fujita, Regulation of transcription and chromatin by methyl-CpG binding protein MBD1, Brain Dev. 23 (Suppl 1) (2001) S174–S176.
[119] N. Fujita, N. Shimotake, I. Ohki, T. Chiba, H. Saya, M. Shirakawa, et al., Mechanism of transcriptional regulation by methyl-CpG binding protein MBD1, Mol. Cell. Biol. 20 (2000) 5107–5118.
[120] N. Fujita, S. Takebayashi, K. Okumura, S. Kudo, T. Chiba, H. Saya, et al., Methylation-mediated transcriptional silencing in euchromatin by methyl-CpG binding protein MBD1 isoforms, Mol. Cell. Biol. 19 (1999) 6415–6426.
[121] P. Zhang, C. Rausch, F.D. Hastert, B. Boneva, A. Filatova, S.J. Patil, et al., Methyl-CpG binding domain protein 1 regulates localization and activity of Tet1 in a CXXC3

domain-dependent manner, Nucleic Acids Res. 45 (2017) 7118–7136.
[122] T. Baubec, R. Ivanek, F. Lienert, D. Schubeler, Methylation-dependent and -independent genomic targeting principles of the MBD protein family, Cell 153 (2013) 480–492.
[123] E.M. Jobe, Y. Gao, B.E. Eisinger, J.K. Mladucky, C.C. Giuliani, L.E. Kelnhofer, et al., Methyl-CpG-binding protein MBD1 regulates neuronal lineage commitment through maintaining adult neural stem cell identity, J. Neurosci. 37 (2017) 523–536.
[124] X. Zhao, T. Ueba, B.R. Christie, B. Barkho, M.J. McConnell, K. Nakashima, et al., Mice lacking methyl-CpG binding protein 1 have deficits in adult neurogenesis and hippocampal function, Proc. Natl. Acad. Sci. U. S. A. 100 (2003) 6777–6782.
[125] B. Hendrich, C. Abbott, H. McQueen, D. Chambers, S. Cross, A. Bird, Genomic structure and chromosomal mapping of the murine and human Mbd1, Mbd2, Mbd3, and Mbd4 genes, Mamm. Genome : Off. J. Int. Mamm. Genome Soc. 10 (1999) 906–912.
[126] J. Boeke, O. Ammerpohl, S. Kegel, U. Moehren, R. Renkawitz, The minimal repression domain of MBD2b overlaps with the methyl-CpG-binding domain and binds directly to Sin3A, J. Biol. Chem. 275 (2000) 34963–34967.
[127] M.A. Desai, H.D. Webb, L.M. Sinanan, J.N. Scarsdale, N.M. Walavalkar, G.D. Ginder, et al., An intrinsically disordered region of methyl-CpG binding domain protein 2 (MBD2) recruits the histone deacetylase core of the NuRD complex, Nucleic Acids Res. 43 (2015) 3100–3113.
[128] Q. Feng, Y. Zhang, The MeCP1 complex represses transcription through preferential binding, remodeling, and deacetylating methylated nucleosomes, Genes Dev. 15 (2001) 827–832.
[129] M.N. Gnanapragasam, J.N. Scarsdale, M.L. Amaya, H.D. Webb, M.A. Desai, N.M. Walavalkar, et al., p66Alpha-MBD2 coiled-coil interaction and recruitment of Mi-2 are critical for globin gene silencing by the MBD2-NuRD complex, Proc. Natl. Acad. Sci. U. S. A. 108 (2011) 7487–7492.
[130] Y. Lu, Y.H. Loh, H. Li, M. Cesana, S.B. Ficarro, J.R. Parikh, et al., Alternative splicing of MBD2 supports self-renewal in human pluripotent stem cells, Cell Stem Cell 15 (2014) 92–101.
[131] X. Le Guezennec, M. Vermeulen, A.B. Brinkman, W.A. Hoeijmakers, A. Cohen, E. Lasonder, et al., MBD2/NuRD and MBD3/NuRD, two distinct complexes with different biochemical and functional properties, Mol. Cell. Biol. 26 (2006) 843–851.
[132] C.P. Tan, S. Nakielny, Control of the DNA methylation system component MBD2 by protein arginine methylation, Mol. Cell. Biol. 26 (2006) 7224–7235.
[133] B. Hendrich, J. Guy, B. Ramsahoye, V.A. Wilson, A. Bird, Closely related proteins MBD2 and MBD3 play distinctive but interacting roles in mouse development, Genes Dev. 15 (2001) 710–723.
[134] K.H. Wood, B.S. Johnson, S.A. Welsh, J.Y. Lee, Y. Cui, E. Krizman, et al., Tagging methyl-CpG-binding domain proteins reveals different spatiotemporal expression and supports distinct functions, Epigenomics 8 (2016) 455–473.
[135] M.F. Fraga, E. Ballestar, G. Montoya, P. Taysavang, P.A. Wade, M. Esteller, The affinity of different MBD proteins for a specific methylated locus depends on their intrinsic binding properties, Nucleic Acids Res. 31 (2003) 1765–1774.
[136] M. Saito, F. Ishikawa, The mCpG-binding domain of human MBD3 does not bind to mCpG but interacts with NuRD/Mi2 components HDAC1 and MTA2, J. Biol. Chem. 277 (2002) 35434–35439.
[137] O. Yildirim, R. Li, J.H. Hung, P.B. Chen, X. Dong, L.S. Ee, et al., Mbd3/NURD complex regulates expression of 5-hydroxymethylcytosine marked genes in embryonic stem cells, Cell 147 (2011) 1498–1510.
[138] J.M. Cramer, J.N. Scarsdale, N.M. Walavalkar, W.A. Buchwald, G.D. Ginder, D.C. Williams Jr., Probing the dynamic distribution of bound states for methylcytosine-binding domains on DNA, J. Biol. Chem. 289 (2014) 1294–1302.
[139] K. Gunther, M. Rust, J. Leers, T. Boettger, M. Scharfe, M. Jarek, et al., Differential roles for MBD2 and MBD3 at methylated CpG islands, active promoters and binding to exon sequences, Nucleic Acids Res. 41 (2013) 3010–3021.
[140] Y. Zhang, H.H. Ng, H. Erdjument-Bromage, P. Tempst, A. Bird, D. Reinberg, Analysis of the NuRD subunits reveals a histone deacetylase core complex and a connection with DNA methylation, Genes Dev. 13 (1999) 1924–1935.
[141] K. Kaji, I.M. Caballero, R. MacLeod, J. Nichols, V.A. Wilson, B. Hendrich, The NuRD component Mbd3 is required for pluripotency of embryonic stem cells, Nat. Cell Biol. 8 (2006) 285–292.
[142] K. Kaji, J. Nichols, B. Hendrich, Mbd3, a component of the NuRD co-repressor complex, is required for development of pluripotent cells, Development 134 (2007) 1123–1132.
[143] M. Luo, T. Ling, W. Xie, H. Sun, Y. Zhou, Q. Zhu, et al., NuRD blocks reprogramming of mouse somatic cells into pluripotent stem cells, Stem Cells 31 (2013) 1278–1286.
[144] Y. Rais, A. Zviran, S. Geula, O. Gafni, E. Chomsky, S. Viukov, et al., Deterministic direct reprogramming of somatic cells to pluripotency, Nature 502 (2013) 65–70.
[145] M. Hirasaki, A. Ueda, M.N. Asaka, K. Uranishi, A. Suzuki, M. Kohda, et al., Identification of the coiled-coil domain as an essential methyl-CpG-binding domain protein 3 element for preserving lineage commitment potential of embryonic stem cells, Stem Cells 36 (2018) 1355–1367.
[146] N. Reynolds, P. Latos, A. Hynes-Allen, R. Loos, D. Leaford, A. O'Shaughnessy, et al., NuRD suppresses pluripotency gene expression to promote transcriptional heterogeneity and lineage commitment, Cell Stem Cell 10 (2012) 583–594.
[147] R.L. Dos Santos, L. Tosti, A. Radzisheuskaya, I.M. Caballero, K. Kaji, B. Hendrich, et al., MBD3/NuRD facilitates induction of pluripotency in a context-dependent manner, Cell Stem Cell 15 (2014) 392.
[148] P. Wu, C. Qiu, A. Sohail, X. Zhang, A.S. Bhagwat, X. Cheng, Mismatch repair in methylated DNA. Structure and activity of the mismatch-specific thymine glycosylase domain of methyl-CpG-binding protein MBD4, J. Biol. Chem. 278 (2003) 5285–5291.
[149] J. Otani, K. Arita, T. Kato, M. Kinoshita, H. Kimura, I. Suetake, et al., Structural basis of the versatile DNA recognition ability of the methyl-CpG binding domain of methyl-CpG binding domain protein 4, J. Biol. Chem. 288 (2013) 6351–6362.

[150] H.E. Krokan, H. Nilsen, F. Skorpen, M. Otterlei, G. Slupphaug, Base excision repair of DNA in mammalian cells, FEBS Lett. 476 (2000) 73–77.
[151] B.A. Manvilla, A. Maiti, M.C. Begley, E.A. Toth, A.C. Drohat, Crystal structure of human methyl-binding domain IV glycosylase bound to abasic DNA, J. Mol. Biol. 420 (2012) 164–175.
[152] C.B. Millar, J. Guy, O.J. Sansom, J. Selfridge, E. MacDougall, B. Hendrich, et al., Enhanced CpG mutability and tumorigenesis in MBD4-deficient mice, Science 297 (2002) 403–405.
[153] H. Hashimoto, Y. Liu, A.K. Upadhyay, Y. Chang, S.B. Howerton, P.M. Vertino, et al., Recognition and potential mechanisms for replication and erasure of cytosine hydroxymethylation, Nucleic Acids Res. 40 (2012) 4841–4849.
[154] U. Muller, C. Bauer, M. Siegl, A. Rottach, H. Leonhardt, TET-mediated oxidation of methylcytosine causes TDG or NEIL glycosylase dependent gene reactivation, Nucleic Acids Res. 42 (2014) 8592–8604.
[155] H. Stroud, S.C. Su, S. Hrvatin, A.W. Greben, W. Renthal, L.D. Boxer, et al., Early-life gene expression in neurons modulates lasting epigenetic states, Cell 171 (2017) 1151–1164 e16.
[156] R.J. Klose, S.A. Sarraf, L. Schmiedeberg, S.M. McDermott, I. Stancheva, A.P. Bird, DNA binding selectivity of MeCP2 due to a requirement for A/T sequences adjacent to methyl-CpG, Mol. Cell 19 (2005) 667–678.
[157] T. Clouaire, J.I. de Las Heras, C. Merusi, I. Stancheva, Recruitment of MBD1 to target genes requires sequence-specific interaction of the MBD domain with methylated DNA, Nucleic Acids Res. 38 (2010) 4620–4634.
[158] A.K. Ludwig, P. Zhang, F.D. Hastert, S. Meyer, C. Rausch, H.D. Herce, et al., Binding of MBD proteins to DNA blocks Tet1 function thereby modulating transcriptional noise, Nucleic Acids Res. 45 (2017) 2438–2457.
[159] P.J. Skene, R.S. Illingworth, S. Webb, A.R. Kerr, K.D. James, D.J. Turner, et al., Neuronal MeCP2 is expressed at near histone-octamer levels and globally alters the chromatin state, Mol. Cell 37 (2010) 457–468.
[160] M. Mellén, P. Ayata, S. Dewell, S. Kriaucionis, N. Heintz, MeCP2 binds to 5hmC enriched within active genes and accessible chromatin in the nervous system, Cell 151 (2012) 1417–1430.
[161] W. Mayer, A. Niveleau, J. Walter, R. Fundele, T. Haaf, Demethylation of the zygotic paternal genome, Nature 403 (2000) 501–502.
[162] J. Oswald, S. Engemann, N. Lane, W. Mayer, A. Olek, R. Fundele, et al., Active demethylation of the paternal genome in the mouse zygote, Curr. Biol. : CB (Curr. Biol.) 10 (2000) 475–478.
[163] S. Kriaucionis, N. Heintz, The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain, Science 324 (2009) 929–930.
[164] M. Tahiliani, K.P. Koh, Y. Shen, W.A. Pastor, H. Bandukwala, Y. Brudno, et al., Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1, Science 324 (2009) 930–935.
[165] R.B. Lorsbach, J. Moore, S. Mathew, S.C. Raimondi, S.T. Mukatira, J.R. Downing, TET1, a member of a novel protein family, is fused to MLL in acute myeloid leukemia containing the t(10;11)(q22;q23), Leukemia 17 (2003) 637–641.
[166] R. Ono, T. Taki, T. Taketani, M. Taniwaki, H. Kobayashi, Y. Hayashi, LCX, leukemia-associated protein with a CXXC domain, is fused to MLL in acute myeloid leukemia with trilineage dysplasia having t(10;11)(q22;q23), Cancer Res. 62 (2002) 4075–4080.
[167] Y.F. He, B.Z. Li, Z. Li, P. Liu, Y. Wang, Q. Tang, et al., Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA, Science 333 (2011) 1303–1307.
[168] T. Pfaffeneder, B. Hackner, M. Truss, M. Munzel, M. Muller, C.A. Deiml, et al., The discovery of 5-formylcytosine in embryonic stem cell DNA, Angew. Chem. 50 (2011) 7008–7012.
[169] S.R. Dalton, A. Bellacosa, DNA demethylation by TDG, Epigenomics 4 (2012) 459–467.
[170] A. Slyvka, K. Mierzejewska, M. Bochtler, Nei-like 1 (NEIL1) excises 5-carboxylcytosine directly and stimulates TDG-mediated 5-formyl and 5-carboxylcytosine excision, Sci. Rep. 7 (2017) 9001.
[171] H. Wu, Y. Zhang, Mechanisms and functions of Tet protein-mediated 5-methylcytosine oxidation, Genes Dev. 25 (2011) 2436–2452.
[172] A. Maiti, A.C. Drohat, Thymine DNA glycosylase can rapidly excise 5-formylcytosine and 5-carboxylcytosine: potential implications for active demethylation of CpG sites, J. Biol. Chem. 286 (2011) 35334–35338.
[173] J.A. Hackett, J.J. Zylicz, M.A. Surani, Parallel mechanisms of epigenetic reprogramming in the germline, Trends Genet. : TIG (Trends Genet.) 28 (2012) 164–174.
[174] P. Hajkova, S. Erhardt, N. Lane, T. Haaf, O. El-Maarri, W. Reik, et al., Epigenetic reprogramming in mouse primordial germ cells, Mech. Dev. 117 (2002) 15–23.
[175] P.W.S. Hill, H.G. Leitch, C.E. Requena, Z. Sun, R. Amouroux, M. Roman-Trufero, et al., Epigenetic reprogramming enables the transition from primordial germ cell to gonocyte, Nature 555 (2018) 392–396.
[176] K. Iqbal, S.G. Jin, G.P. Pfeifer, P.E. Szabo, Reprogramming of the paternal genome upon fertilization involves genome-wide oxidation of 5-methylcytosine, Proc. Natl. Acad. Sci. U. S. A. 108 (2011) 3642–3647.
[177] D.M. Messerschmidt, B.B. Knowles, D. Solter, DNA methylation dynamics during epigenetic reprogramming in the germline and preimplantation embryos, Genes Dev. 28 (2014) 812–828.
[178] Y. Okamoto, N. Yoshida, T. Suzuki, N. Shimozawa, M. Asami, T. Matsuda, et al., DNA methylation dynamics in mouse preimplantation embryos revealed by mass spectrometry, Sci. Rep. 6 (2016) 19134.
[179] H. Sasaki, Y. Matsui, Epigenetic events in mammalian germ-cell development: reprogramming and beyond, Nat. Rev. Genet. 9 (2008) 129–140.
[180] M. Munzel, D. Globisch, T. Bruckl, M. Wagner, V. Welzmiller, S. Michalakis, et al., Quantification of the sixth DNA base hydroxymethylcytosine in the brain, Angew. Chem. 49 (2010) 5375–5377.
[181] M.J. Booth, M.R. Branco, G. Ficz, D. Oxley, F. Krueger, W. Reik, et al., Quantitative sequencing of 5-methylcytosine and 5-hydroxymethylcytosine at single-base resolution, Science 336 (2012) 934–937.
[182] G. Ficz, M.R. Branco, S. Seisenberger, F. Santos, F. Krueger, T.A. Hore, et al., Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation, Nature 473 (2011) 398–402.

[183] S.G. Jin, X. Wu, A.X. Li, G.P. Pfeifer, Genomic mapping of 5-hydroxymethylcytosine in the human brain, Nucleic Acids Res. 39 (2011) 5015–5024.
[184] W.A. Pastor, U.J. Pape, Y. Huang, H.R. Henderson, R. Lister, M. Ko, et al., Genome-wide mapping of 5-hydroxymethylcytosine in embryonic stem cells, Nature 473 (2011) 394–397.
[185] Y. Xu, F. Wu, L. Tan, L. Kong, L. Xiong, J. Deng, et al., Genome-wide regulation of 5hmC, 5mC, and gene expression by Tet1 hydroxylase in mouse embryonic stem cells, Mol. Cell 42 (2011) 451–464.
[186] F. Neri, D. Incarnato, A. Krepelova, S. Rapelli, A. Pagnani, R. Zecchina, et al., Genome-wide analysis identifies a functional association of Tet1 and Polycomb repressive complex 2 in mouse embryonic stem cells, Genome Biol. 14 (2013) R91.
[187] A.B. Robertson, J.A. Dahl, C.B. Vagbo, P. Tripathi, H.E. Krokan, A. Klungland, A novel method for the efficient and selective identification of 5-hydroxymethylcytosine in genomic DNA, Nucleic Acids Res. 39 (2011) e55.
[188] M. Yu, G.C. Hon, K.E. Szulwach, C.X. Song, L. Zhang, A. Kim, et al., Base-resolution analysis of 5-hydroxymethylcytosine in the mammalian genome, Cell 149 (2012) 1368–1380.
[189] L. Wen, X. Li, L. Yan, Y. Tan, R. Li, Y. Zhao, et al., Whole-genome analysis of 5-hydroxymethylcytosine and 5-methylcytosine at base resolution in the human brain, Genome Biol. 15 (2014) R49.
[190] F. Neri, D. Incarnato, A. Krepelova, S. Rapelli, F. Anselmi, C. Parlato, et al., Single-base resolution analysis of 5-formyl and 5-carboxyl cytosine reveals promoter DNA methylation dynamics, Cell Rep. 10 (2015) 674–683.
[191] E.A. Raiber, D. Beraldi, G. Ficz, H.E. Burgess, M.R. Branco, P. Murat, et al., Genome-wide distribution of 5-formylcytosine in embryonic stem cells is associated with transcription and depends on thymine DNA glycosylase, Genome Biol. 13 (2012) R69.
[192] C.X. Song, K.E. Szulwach, Q. Dai, Y. Fu, S.Q. Mao, L. Lin, et al., Genome-wide profiling of 5-formylcytosine reveals its roles in epigenetic priming, Cell 153 (2013) 678–691.
[193] C. Xu, K. Liu, M. Lei, A. Yang, Y. Li, T.R. Hughes, et al., DNA sequence recognition of human CXXC domains and their structural determinants, Structure 26 (2018) 85–95 e3.
[194] M. Ko, J. An, H.S. Bandukwala, L. Chavez, T. Aijo, W.A. Pastor, et al., Modulation of TET2 expression and 5-methylcytosine oxidation by the CXXC domain protein IDAX, Nature 497 (2013) 122–126.
[195] S.G. Jin, Z.M. Zhang, T.L. Dunwell, M.R. Harter, X. Wu, J. Johnson, et al., Tet3 reads 5-carboxylcytosine through its CXXC domain and is a potential guardian against neurodegeneration, Cell Rep. 14 (2016) 493–505.
[196] C. Jin, Y. Lu, J. Jelinek, S. Liang, M.R. Estecio, M.C. Barton, et al., TET1 is a maintenance DNA demethylase that prevents methylation spreading in differentiated cells, Nucleic Acids Res. 42 (2014) 6956–6971.
[197] S.M. Langemeijer, R.P. Kuiper, M. Berends, R. Knops, M.G. Aslanyan, M. Massop, et al., Acquired mutations in TET2 are common in myelodysplastic syndromes, Nat. Genet. 41 (2009) 838–842.
[198] W. Zhang, W. Xia, Q. Wang, A.J. Towers, J. Chen, R. Gao, et al., Isoform switch of TET1 regulates DNA demethylation and mouse development, Mol. Cell 64 (2016) 1062–1073.
[199] C.R. Good, J. Madzo, B. Patel, S. Maegawa, N. Engel, J. Jelinek, et al., A novel isoform of TET1 that lacks a CXXC domain is overexpressed in cancer, Nucleic Acids Res. 45 (2017) 8269–8281.
[200] N. Agarwal, A. Becker, K.L. Jost, S. Haase, B.K. Thakur, A. Brero, et al., MeCP2 Rett mutations affect large scale chromatin organization, Hum. Mol. Genet. 20 (2011) 4187–4195.
[201] M. Ko, Y. Huang, A.M. Jankowska, U.J. Pape, M. Tahiliani, H.S. Bandukwala, et al., Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2, Nature 468 (2010) 839–843.
[202] M. Smets, S. Link, P. Wolf, K. Schneider, V. Solis, J. Ryan, et al., DNMT1 mutations found in HSANIE patients affect interaction with UHRF1 and neuronal differentiation, Hum. Mol. Genet. 26 (2017) 1522–1534.
[203] C. Bauer, K. Gobel, N. Nagaraj, C. Colantuoni, M. Wang, U. Muller, et al., Phosphorylation of TET proteins is regulated via O-GlcNAcylation by the O-linked N-acetylglucosamine transferase (OGT), J. Biol. Chem. 290 (2015) 4801–4812.
[204] K. Blaschke, K.T. Ebata, M.M. Karimi, J.A. Zepeda-Martinez, P. Goyal, S. Mahapatra, et al., Vitamin C induces Tet-dependent DNA demethylation and a blastocyst-like state in ES cells, Nature 500 (2013) 222–226.
[205] M.C. Cardoso, H. Leonhardt, DNA methyltransferase is actively retained in the cytoplasm during early development, J. Cell Biol. 147 (1999) 25–32.
[206] S.R. Kinney, S. Pradhan, Regulation of expression and activity of DNA (cytosine-5) methyltransferases in mammalian cells, Prog. Mol. Biol. Transl. Sci. 101 (2011) 311–333.
[207] W. Qin, H. Leonhardt, G. Pichler, Regulation of DNA methyltransferase 1 by interactions and modifications, Nucleus 2 (2011) 392–402.
[208] F. Spada, A. Haemmer, D. Kuch, U. Rothbauer, L. Schermelleh, E. Kremmer, et al., DNMT1 but not its interaction with the replication machinery is required for maintenance of DNA methylation in human cells, J. Cell Biol. 176 (2007) 565–571.
[209] P. Vella, A. Scelfo, S. Jammula, F. Chiacchiera, K. Williams, A. Cuomo, et al., Tet proteins connect the O-linked N-acetylglucosamine transferase Ogt to chromatin in embryonic stem cells, Mol. Cell 49 (2013) 645–656.
[210] K.E. Szulwach, X. Li, Y. Li, C.X. Song, H. Wu, Q. Dai, et al., 5-hmC-mediated epigenetic dynamics during postnatal neurodevelopment and aging, Nat. Neurosci. 14 (2011) 1607–1616.
[211] P. Zhang, A.K. Ludwig, F.D. Hastert, C. Rausch, A. Lehmkuhl, I. Hellmann, et al., L1 retrotransposition is activated by Ten-eleven-translocation protein 1 and repressed by methyl-CpG binding proteins, Nucleus 8 (2017) 548–562.
[212] M. Mellen, P. Ayata, N. Heintz, 5-hydroxymethylcytosine accumulation in postmitotic neurons results in functional demethylation of expressed genes, Proc. Natl. Acad. Sci. U. S. A. 114 (2017) E7812–E7821.
[213] J. Ren, L. Wen, X. Gao, C. Jin, Y. Xue, X. Yao, DOG 1.0: illustrator of protein domain structures, Cell Res. 19 (2009) 271–273.