The role of methyl-binding proteins in chromatin organization and epigenome maintenance

Alexandra Fournier, Nobuhiro Sasai, Mitsuyoshi Nakao and Pierre-Antoine Defossez

Advance Access publication date 19 December 2011

Abstract

Methylated DNA can be specifically recognized by a set of proteins called methyl-CpG-binding proteins (MBPs), which belong to three different structural families in mammals: the MBD family, the Kaiso and Kaiso-like proteins and the SRA domain proteins. A current view is that, once bound to methylated DNA, MBPs translate the DNA methylation signal into appropriate functional states, through interactions with diverse partners. However, if some of the biological functions of MBPs have been widely described—notably transcriptional repression—others are poorly understood, and more generally the extent of MBP activities remains unclear. Here we propose to discuss the role of MBPs in two crucial nuclear events: chromatin organization and epigenome maintenance. Finally, important challenges for future research as well as for biomedical applications in pathologies such as cancers—in which DNA methylation patterns are widely altered—will be mentioned.

Keywords: DNA methylation recognition; transcriptional repression; heterochromatin formation and maintenance; epigenetic cooperation; cancer

INTRODUCTION: WHERE THE METHYLATED DNA IS, AND HOW WE KNOW DNA METHYLATION IS IMPORTANT

In eukaryotes, only cytosine can be methylated—this differs from bacteria that can also display adenine methylation. In mammals, DNA methylation takes place mostly in the context of CpG dinucleotides, and 60–80% of all CpG dinucleotides are methylated, in a heterogeneous fashion. Most CpG islands are unmethylated, whereas the rest of the genome is globally methylated. In other words, the regions of the genome that are methylated include repeated elements, intergenic regions, gene bodies and certain CpG islands [1]. The regions that are unmethylated are mainly CpG islands. There are excellent recent reviews on CpG islands, to which the reader can refer for further information [2, 3], including a discussion of the possible mechanisms by which some, but not all, CpG islands may escape DNA methylation.

Three DNA methyltransferases exist in mammals: Dnmt1, Dnmt3a and Dnmt3b [4]. Their invalidation in mouse is lethal at different stages, showing that

Corresponding author. P.A. Defossez, CNRS UMR7216—Epigenetics and Cell Fate, Université Paris Diderot, Paris, France.
Tel: +33 1 57 278 926; Fax: +33 1 57 278 911; E-mail: pierre-antoine.defossez@univ-paris-diderot.fr

Alexandra Fournier is a postdoctoral fellow supported by Institut National du Cancer (INCa) in CNRS UMR7216/Université Paris Diderot. Her research project is about the role of UHRF proteins at the interface of DNA methylation, epigenetic repression and cancer.

Nobuhiro Sasai is a postdoctoral fellow supported by the Global COE Program (Cell Fate Regulation Research and Education Unit) in Kumamoto University. His research focuses on the molecular mechanism of transcriptional regulation mediated by DNA and histone modifications.

Mitsuyoshi Nakao is a professor and the director of the Institute of Molecular Embryology and Genetics (IMEG) at Kumamoto University. His research interest is the molecular basis of epigenetic cell regulation in development and human diseases.

Pierre-Antoine Defossez is Director of Research of the team ‘Functional domains of eukaryotic genomes’ within the CNRS UMR7216/Université Paris Diderot. His research work concerns the mechanisms of transcriptional silencing and the role of DNA methylation in mammals.

© The Author 2011. Published by Oxford University Press. All rights reserved. For permissions, please email: journals.permissions@oup.com
these genes are essential. Importantly, it has been shown for DNMT1 that the catalytic activity is essential: a catalytically inactive point mutant of the enzyme does not rescue the knockout [5, 6]. This contrasts with the situation in *Xenopus*, where certain key roles of DNMT1 do not depend on it methylating DNA [7, 8], as well as in *Zebrafish*, where mutations in *Dnmt1* are tolerated [9, 10].

The DNMT1 experiments tell us that it is essential for a mammalian organism that its DNA be methylated. Why is that? Many known effects of DNA methylation have to do with gene expression: DNA methylation has been shown to control the expression of parentally imprinted genes, inhibit the expression of repeated elements and lock in the repressed state of certain genes (including Oct4 in non-pluripotent cells). In general, the presence of DNA methylation correlates with the presence of heterochromatin and with an environment that is repressive for transcription [1].

Of note, DNA methylation is not always required at the cellular level. The clearest illustration is that ES cells can be engineered to lack all three DNA methyltransferases, therefore having no methylation of DNA [11]. This unmethylated state is tolerated in undifferentiated ES cells, and also in cells of the extra-embryonic tissues, but seems incompatible with differentiation into cells of the embryo proper [12].

Several mammalian proteins have the ability to distinguish methylated from unmethylated DNA. This occurs in two directions: some proteins bind DNA better or exclusively when it is unmethylated; others bind it better or exclusively when it is methylated. The latter group of proteins, called methyl (−CpG)-binding proteins (MBPs), is the object of this review. We will describe the three families of MBPs known in mammals, and discuss their role in chromatin organization and epigenome maintenance, in cooperation with other epigenetic factors.

Indeed, DNA methylation is mechanistically and functionally linked to other epigenetic mechanisms, notably histone modifications. Of particular importance is histone lysine methylation, which has been shown to regulate or act cooperatively with DNA methylation in a variety of biological processes [13–19]. Specifically, DNA methylation correlates strictly with the absence of H3K4 methylation—trimethylation of histone H3 lysine 4 (H3K4me3) is enriched within unmethylated intragenic CGIs—and correlates loosely with presence of H3K9 methylation [20]. DNA methylation is also strictly antagonistic with H2A.Z [21]. Dnmt3a and its cofactor Dnmt3L specifically recognize unmethylated H3K4 through the ATRX-Dnmt3-Dnmt3L (ADD) domain; the interaction is inhibited by methylation at H3K4 [22, 23].

The cross-talk between DNA methylation and histone modifications is ensured notably by interactions between MBPs and other epigenetic factors. The resulting modulation of chromatin organization is important for different nuclear regulation events that will be discussed in this review.

**THE THREE FAMILIES OF METHYL-CpG-Binding Proteins**

Methyl-CpG-binding proteins are represented in Figure 1.

**The MBD family**

In mammals, at least four MBD proteins, MeCP2, MBD1, MBD2 and MBD4, bind methylated DNA. Three other proteins, MBD3, MBD5 and MBD6, are members of this family but do not bind methylated DNA [24, 25]. This family of proteins has been well covered in other papers and reviews [24, 26], so we will say just a few words in this section.

While the MBD proteins share a common capacity for binding symmetrically methylated CpGs via the MBD, each protein shows distinct domain structure and interacts with different proteins. This suggests that the MBD proteins have primarily nonoverlapping functions and play roles in diverse biological processes. However, despite extensive studies to unravel the functions of the MBD proteins, the biological significance of the proteins is still unclear due to the absence of overt phenotypes in mice lacking each MBD gene. The only exception is MePC2, mutations/deletion of which are associated with a progressive neurological disorder, Rett syndrome, in humans (RTT) [27] and RTT-like phenotypes in mice. Given that both loss and excess of MeCP2 induce neurological phenotypes [28–30], appropriate regulation of MeCP2 activity and subsequent control of gene expression are considered to be critical for the normal function of the brain.

**Kaiso and Kaiso-like proteins**

Kaiso (also known as ZBTB33) was independently discovered first as a DNA-binding factor interacting with the adhesion molecule p120ctn [31], then as
a transcriptional repressor binding tandem methylated CpGs [32]. Contrary to MBD proteins, Kaiso also binds unmethylated consensus sequences that do not contain a CpG and are not methylatable [33, 34]. Kaiso is a member of the ZBTB family and contains a BTB/POZ domain at the N terminus and three tandem Krüppel-like C2H2 zinc fingers at the C terminus. The BTB/POZ domain permits protein–protein interactions [35], and zinc fingers allow to bind sequence-specific DNA (KBS) [33, 36]. Depletion of Kaiso in *Xenopus* embryos results in derepression of methylated genes before the midblastula transition [36–38]. Recently, Kaiso was showed to localize at centrosomes and spindle microtubules during mitosis, revealing a new functional role for Kaiso in centrosome separation during mitosis, microtubule nucleation and/or the G2-M checkpoint [39].

A homology search based on Kaiso-like zinc fingers led to the identification of two proteins closely related to Kaiso, ZBTB4 and ZBTB38 [also known as CtBP-interacting BTB zinc finger protein (CIBZ)], which share many of its properties [40]. ZBTB4 and ZBTB38 repress transcription of methylated promoters in a reporter assay, and localize to densely methylated pericentromeric heterochromatin regions in mouse cells in a DNA methylation-dependent manner [41]. They bind methylated DNA [35, 42] in a sequence-specific manner: the nucleotides surrounding the methylated CpG directly contribute to the binding affinity [34]. They also bind certain unmethylated sequences; ZBTB4, but not ZBTB38, is able to bind a KBS-related sequence [34].

Although the expression patterns of Kaiso, ZBTB4 and ZBTB38 are overlapping in several adult mouse tissues [31, 40, 43], the degree of redundancy between them is currently unknown. The human genome encodes about 700 proteins containing zinc fingers. Kaiso and its paralogs apparently recognize methylated DNA by a canonical mechanism [34], raising the possibility that additional zinc finger proteins will be found to bind methylated DNA [135]. Thus, Bartke *et al.* [44] identified a number of proteins enriched in CpG-methylated DNA, and among them some ZBTBs that would be interesting candidates.

**The SRA domain proteins**

UHRF1, also known as Nuclear Protein of 95 kDa (Np95) in mouse, was initially identified as a protein that binds to a CCAAT box in the promoter of the topoisomerase IIα gene [45]. It was then showed to bind methylated CpGs [46]. UHRF1 contains at least five functional domains [47]: an ubiquitin-like domain (Ubl) at the N-terminus; a tandem Tudor

---

**Figure I:** The three families of methyl-CpG-binding proteins in mammals. (A) The MBD family: MeCP2, MBD1, MBD2 and MBD4; of note, MBD3, MBD5 and MBD6, also part of this family but not binding meDNA, are not represented here. MBD, methyl-CpG-binding domain; TRD, transcriptional repression domain; CTD, chromatin compaction; CXXC ZF, unmethylated-CpG-binding zinc finger; glycosylase, DNA glycosylase activity. (B) Kaiso and Kaiso-like proteins: Kaiso, ZBTB4 (Zinc finger and BTB domain containing 4) and ZBTB38. BTB, BTB/POZ (Broad complex, Tramtrack, Bric a brac/Pox virus and Zinc finger) domain; mZF, methyl-CpG-binding zinc fingers. (C) SRA domain proteins: UHRF1 (Ubiquitin-like with PHD and RING finger domains I) and UHRF2. Ubl, ubiquitin-like domain; TTD, tandem tudor domain; PHD, plant homeodomain finger domain; SRA, SET and RING associated domain; RING, Really Interesting New Gene finger domain.
domain (TTD) that binds histone H3 tails trimethylated on residue K9 [48–50]; a plant homeodomain (PHD), which can bind modified histones; a SET and RING Associated domain (SRA) that binds methylated DNA; and finally a Really Interesting New Gene (RING) domain at its C-terminus, which is endowed with catalytic activity as an E3 ubiquitin ligase.

UHRF1 is unique among all methylcytosine-binding proteins in being essential in mouse; its deletion leads to rapid developmental arrest [51]. UHRF1 is involved in the regulation of the cell cycle and DNA damage control; it regulates cell proliferation by activating cell cycle inducer genes such as topoisomerase IIa, and blocking cell cycle inhibitory factors like Rb [46, 52–55]. UHRF1 is involved in heterochromatin function [48, 49, 56] and maintenance [51, 57], as we will discuss later.

In mammalian genomes, UHRF1 has only one close paralog: UHRF2 [58]. It can also bind methylated DNA through its SRA domain, and H3K9me3 through its TTD domain [59]. Contrary to Uhrf1, localization of Uhrf2 to pericentric heterochromatin does not depend on DNA methylation but on H3K9me3 [59]. Finally, Uhrf2 has been implicated in cell cycle progression and tumorigenesis [60].

Uhrf1 is expressed in ESCs and down regulated during differentiation, whereas Uhrf2 is upregulated and highly expressed in differentiated mouse tissues. Moreover, Uhrf2 does not restore DNA methylation in Uhrf1-deficient cells [59]. Taken together, these data argue against a functional redundancy of Uhrf1 and Uhrf2.

Of note, related SRA proteins in plants, such as VIM1, bind methylated cytosines in non-CpG contexts [61]. As non-CpG methylation has been identified in mammalian ES cells [62–65], one could speculate that UHRF proteins might recognize this mark [66], even if some in vitro experiments were shown to be negative [57].

ROLE OF METHYL-CpG-BINDING PROTEINS IN CHROMATIN ORGANIZATION

Site-specific chromatin organization

Kaiso and Kaiso–like proteins interact with different sets of corepressors

Kaiso interacts with the NCoR corepressor complex, which includes HDAC3, TBL1/TBLR1, GPS2 and IR10 [67]. This recruitment of NCoR is required for histone hypoacetylation and H3K9 methylation over the promoter of a target gene, at least in cancer cells [67]. ZBTB38 interacts with the co-repressors CtBP1s (C-terminal Binding Proteins) [43]; the CtBP complex includes histone deacetylases, methyltransferases and demethylases [68]. ZBTB4 can repress transcription in a methylation-independent process; it heterodimerizes with MIZ1, represses P21CIP1 expression and inhibits cell cycle arrest in response to p53 activation, through the recruitment of a Sin3/HDAC complex [69].

ZBTB4 was identified as a new prognostic factor in breast cancer [70]. ZBTB4 is downregulated in breast tumors, and its expression is significantly correlated with relapse-free survival. Survival differences were not observed for other ZBTBs (notably Kaiso and ZBTB38). Similar results were noticed for lymphoma and lung cancer tumors, confirming the prognostic significance of ZBTB4. The authors showed that ZBTB4 repressed a unique set of genes that was not regulated by Kaiso or ZBTB38. The specificity of ZBTB4 for these genes probably does not come from intrinsic DNA-binding capacities as the three of them have similar targets, at least in vitro [34], and so remains to be explained. In breast cancer cells, ZBTB4 negatively regulates the expression of the transcriptional activators Sp1, Sp3 and Sp4, and some of their important target genes such as VEGF, Survivin and VEGFR1. This involves a competition process between ZBTB4 and Sp proteins for binding to GC-rich promoter sequences, which may be influenced by the methylation status of the DNA [70]. Finally, the authors showed that ZBTB4 was repressed by miRNAs derived from the oncogenic miR-17-92 cluster (Figure 2). Restoration of ZBTB4 expression by inhibiting these oncogenic miRNAs suppressed expression of Sp1/3/4 and Sp regulated pro-oncogenic genes, resulting in inhibition of cancer cell proliferation and invasion. So drugs targeting the oncogenic miR-17-92 cluster may be interesting for clinical applications.

UHRF1 and regulation of gene expression

Recently, several pieces of work showed that the PHD domain of UHRF1 could recognize and bind unmodified H3R2, but not methylated H3R2 [71–73]. The authors [72] identify ADAM19 and SUSD2 as targets for repression by UHRF1 in human cell lines. These genes are bound by UHRF1 near promoters and in gene bodies. Their promoters
lack H3R2 methylation and also H3K9 trimethylation, which may explain the dependency on the PHD domain for UHRF1 regulation of these genes. Finally, these two genes present DNA methylation on their promoter, and methylation is reduced upon knockdown of UHRF1, meaning that UHRF1-mediated repression of these two genes would involve DNA methylation. Taken together, these data suggest that the primary function of PHDUHRF1 is to localize UHRF1 to euchromatic targets lacking H3K9me3, through binding the unmodified R2 residue. Interestingly, they found that the presence of H3K9me3 counteracted the inhibitory effect of H3R2 dimethylation on UHRF1 binding to histone H3 in vitro [72].

Global chromatin organization
MeCP2 as a global chromatin regulator
Several studies suggested that MeCP2 is associated with transcriptional activation [74, 75]. These evidences argued against a classical model of MeCP2 function in which MeCP2 alters local chromatin environment to repress transcription through recruitment of the Sin3A/HDAC complex to target promoters [76]. In addition, there is still a debate about whether MeCP2 is a transcriptional repressor of specific target genes or a global chromatin regulator.

MeCP2 is a basic nuclear protein abundantly expressed in neurons. Although MeCP2 can bind naked DNA duplexes in a methylation-dependent fashion [77], it is currently widely accepted that MeCP2 is a chromatin architectural protein. MeCP2 prefers to bind the DNA entry and exit sites of the nucleosomes in a similar mode to linker histones [78–80]. Supporting these findings is that, by a proteomic screen for chromatin interacting proteins, MeCP2 was enriched on methylated DNA only in the context of nucleosomes [44]. Furthermore, in vivo studies showed that, in neuronal nuclei, MeCP2 molecules are nearly as abundant as the histone octamer, and MeCP2 chromatin binding tracks the methylated CpG density of the genome [81]. In MeCP2 knockout neurons, increased histone acetylation and doubling of the linker histone H1 levels were observed [81]. These findings suggest a model in which MeCP2 replaces H1 and deacetylates histones throughout the genome to establish a neuron-specific chromatin structure. This specialized chromatin would be required to suppress transcriptional noise that might be problematic for the brain function. These observations argue that, at least in neurons, MeCP2 serves as a global chromatin regulator rather than a gene-specific transcriptional repressor. If this is the case, why does MeCP2 deficiency in the brain have subtle effects on gene expression profiles [82]? Getting to an answer is challenging. For one thing, the brain is a heterogeneous tissue, with different cell types showing heterogeneity in MeCP2 expression. In contrast, culture cell lines are homogeneous but do not reflect physiological situations. In addition, it cannot be excluded the possibility that MeCP2 may act as a gene-specific transcriptional repressor in particular cells, as MeCP2 is expressed in nonneuronal cells/tissues albeit with comparatively lower levels.

What is the molecular mechanism of MeCP2-mediated chromatin regulation? Biochemical analysis of MeCP2 protein has clarified that neither Sin3A nor other proteins form a stable complex with MeCP2 [83]. Nevertheless, a broad range of proteins has been identified as MeCP2 partners, including
corepressor proteins, histone modification/binding proteins and chromatin remodeling factors [84]. This raised the possibility that the interaction occurs transiently or requires specific environment in vivo. Of particular interest among the MeCP2 interactors is an ATP-dependent SWI/SNF DNA helicase, ATRX [85]. MeCP2 mutations found in RTT patients disrupt the interaction with ATRX and displace ATRX from heterochromatin [86]. MeCP2, in cooperation with ATRX, Cohesin and an insulator protein CTCF, regulates the expression of a subset of imprinting genes in an allele-specific manner in the postnatal mouse brain [87]. Given that Cohesin and CTCF regulate gene expression through the formation of chromatin loops [88], MeCP2 may be involved in higher order chromatin regulation as well [89]. MeCP2 has been proposed to regulate the expression of the imprinted gene DLX5 through chromatin looping [90]. Likewise, another MBD protein MBD2 is suggested to be involved in the chromatin loop-mediated gene regulation at the GATA-4 gene locus when the gene is fully silenced by PcG protein occupation and hypermethylation of multiple CpG islands in HCT116 adult cancer cells [91]. As these are observations at specific genomic loci and/or in cancer cells, further studies are required to test whether the chromatin loop-mediated gene regulation by MBD proteins is a generalized mechanism or not. Recently developed approaches, such as chromosome conformation capture (3C) or its derivatives [92], would be quite helpful to understand the roles of MeCP2 or other MBD proteins in the regulation of higher order chromatin conformation.

**UHRF1 is involved in large-scale organization of pericentric heterochromatin**

During mid-to-late S phase, Uhrf1 specifically relocalizes to chromocenters where it is involved in reformation of pericentric heterochromatin [56]. Uhrf1 determines large-scale modifications of chromocenters and its PHD domain would be required for this process. Depletion of Uhrf1 leads to a reduced number and an increased size of chromocenters; at the opposite, overexpression of Uhrf1 induces decondensation of chromocenters (including of pericentric satellite DNA) [49]. This phenomenon would occur in a H3K9me3- and DNA methylation-independent manner, whereas other works show that localization of UHRF1 to pericentric heterochromatin depends on interaction with both H3K9me3 and methylated DNA [48, 93].

Two other points are under debate. The first one concerns the domain of UHRF1 involved in large-scale chromatin organization. Contrary to the previous authors [49], for Karagiani et al. [48], it is the RING domain that is important; they showed that overexpression of a RING-mutant form of UHRF1 disrupted higher order organization of heterochromatin in interphase mouse cells. The RING domain of UHRF1 is endowed with an E3 ligase activity toward H3 [48, 94], and here, UHRF1 binds to H3K9me3 and promotes local chromatin ubiquitination. Interestingly, in Schizosaccharomyces pombe, protein ubiquitination seems to play a critical role in heterochromatin formation and H3K9 methylation [95, 96]. So in mammals, this mark could signal the recruitment of downstream molecules, such as methyltransferases and deacetylases, which would induce spreading of heterochromatin [48].

The second point of debate is about the domain of UHRF1 that binds to H3K9me3. Some propose a cooperation between the SRA and the PHD domains; the SRA domain would increase the binding of UHRF1 to H3K9me3 through its PHD domain [48, 94]. But others found that H3K9me3 binding was mediated by the TTD domain [50, 93]. More precisely, the TTD domain of UHRF1 would simultaneously recognize H3K9me3 and unmodified H3K4 by a single domain that undergoes a conformational change to accommodate both lysines [93]. Despite a possible difference in techniques or model used explaining this discrepancy, it could also reflect the dynamic interactions between protein domains that may vary upon conformational changes.

**ROLE OF METHYL-CpG-BINDING PROTEINS IN MAINTENANCE OF EPIGENOME AND GENOME INTEGRITY**

**Maintenance of heterochromatin**

Methyl-binding proteins associate with chromatin modifying activities that establish and/or maintain heterochromatin.

**Maintenance of heterochromatin structure by the MBD proteins**

Epigenetic information is faithfully inherited concomitant with DNA replication to maintain
a given chromatin state and a cell type [97]. An intriguing finding concerning MBPs on this process is that MBD1 mediates the maintenance of H3K9 methylation at DNA methylated loci. MBD1 specifically interacts with the p150 subunit of chromatin assembly factor 1 (CAF-1) in S-phase [98, 99]. Moreover, MBD1 recruits a H3K9-specific methyltransferase SETDB1 to CAF-1 on progressing replication forks to inherit H3K9 methylation status after replication of chromatin in HeLa cells [99]. The extent to which MBD1 contributes to epigenome maintenance remains unclear given the fact that MBD1 is not essential for normal development, which is in sharp contrast to the critical roles of SETDB1 and Dnmts for development. Indeed, SETDB1 is reported to bind and repress endogenous retroviruses independent of DNA methylation in mouse ES cells [100]. There is also a possibility of functional redundancy or compensation among MBD proteins or H3K9 methyltransferases during chromatin replication. In addition to SETDB1, another H3K9 methyltransferase SUV39H1 and the polycomb group proteins PC2 and Ring1b are known to interact with MBD1, although mechanistic action of the complexes during replication of chromatin remains to be elucidated [101, 102]. Moreover, it has been reported that the MBD2 complex localizes to replication foci and associates with DNMT1 in human cells during late S-phase [103], and that MeCP2 associates with DNMT1 [104]. However, there is no report about the interaction of these proteins with CAF-1 or SETDB1, and it remains unknown whether MBD2 and MeCP2 have a similar or compensatory function with MBD1 in epigenome maintenance. It would also be interesting to see how UHRF1 and MBD proteins function interconnectively during chromatin replication.

Another remarkable feature of the MBD1 complex is that its activity is modulated by a SUMOylation system. MBD1, CAF-1, SETDB1 and its cofactor MCAF1 all contain a SUMO interacting motif (SIM) and/or are a substrate of SUMOylation enzymes [105–109]. Generally, SUMOylation of proteins regulates its activity, stability, localization or interaction with partner proteins [110]. In the case of MBD1, the SUMOylated form is recognized by the SIM of MCAF1 and contributes to the formation of heterochromatin [105]. Understanding how and when the SUMOylation system modifies the activity of the MBD1 complex would be of great interest in clarifying the role of SUMOylation in heterochromatin formation and epigenome maintenance.

Maintenance of heterochromatin by UHRF1

DNA methylation patterns must be faithfully copied from parent to daughter cells. This inheritance process is mediated by DNMT1, but additional factor(s) are needed to recruit DNMT1 to the replication foci, and possibly regulate it to increase its preference for hemimethylated CpGs and/or prevent aberrant methylation of previously unmethylated sites [111, 112].

UHRF1 robustly accumulates into the DAPI-dense heterochromatin loci during mid-to-late S-phase [57, 113, 114], in a manner dependent on the presence of (hemi)methylated DNA [51]. During S-phase, UHRF1 forms complexes with DNMT1 and PCNA at the replicating heterochromatin [51, 115]. These observations suggested a potential direct role for UHRF1 in recruiting DNMT1 at replication foci. Experiments performed in Uhrf1−/− murine ES cells showed that Dnmt1 presented a diffused pattern and no significant enrichment in the replicating heterochromatin during S-phase. Expression of a Myc-tagged Uhrf1 protein in Uhrf1−/− cells rescued the Dnmt1 localization pattern [51], meaning that Uhrf1 is required for the recruitment of Dnmt1 into replication foci during S-phase.

UHRF1 binds hemimethylated DNA through its SRA domain and would interact with DNMT1 through its PHD domain [57] and/or its SRA domain [116] (Figure 3). The hemimethylated sites are small, so it seems impossible for large proteins like DNMT1 and UHRF1 to bind such sites simultaneously. This raises questions about the sequence of events leading to correct methylation of daughter DNA strand. The X-ray crystal structure of the UHRF1 SRA domain bound to hemimethylated oligonucleotides revealed that the SRA domain contacts DNA by two loops [113, 117–121]; one recognizes the hemimethylated DNA, while the other flips out the hemimethylated cytosine molecule from the DNA helix, presumably to help DNMT1 to methylate DNA at the right place on the daughter DNA strand. So UHRF1 would act as a guide or anchor, rather than a recruiter, to present the cytosine to be methylated to DNMT1 [122]. DNMT1 methyltransferase activity subsequently restores symmetrical CpG methylation, thereby maintaining the parental methylation pattern (Figure 3).
UHRF1 plays a crucial role in DNA methylation maintenance following DNA replication. Moreover, since UHRF1 binds to H3K9me3 and G9a [48, 51, 114], it could be important for the maintenance of histone marks as well.

Maintenance of (epi)genome integrity
Role of UHRF1 in DNMT1 stability: a subtle molecular balance
MBPs associate with multiple protein partners to modify surrounding chromatin, providing the appropriate context for crucial nuclear events to occur. This complex and dynamic interplay requires a fine regulation of all actors. Recently, a subtle mechanism was revealed to ensure regulation of DNMT1 abundance [123, 124]. Du et al. showed that DNMT1 was destabilized by acetylation by the acetyltransferase Tip60, which allows ubiquitination by UHRF1, thereby targeting DNMT1 for proteasomal degradation. In contrast, DNMT1 was stabilized by HDAC1 and the deubiquitinase USP7, also known as HAUSP. Of course, these events are tightly regulated during the cell cycle, and the initiation of DNMT1 degradation was coordinated with the end of DNA replication, and so of the need for DNMT activity [123].

Notably, the importance of these regulation mechanisms is highlighted by a recent work [125], which shows that the disruption of DNMT1/PCNA/UHRF1 interactions promotes a global DNA hypo-methylation in human gliomas, which is a source of chromosomal instability, notably.

Roles of MBD proteins in DNA repair and DNA demethylation
Formation of specific heterochromatic domains is important in maintaining genome integrity [126]. Several lines of evidence suggest that MBD proteins contribute to maintain genome stability. An engineered L1-EGFP transgene system showed an increase of retrotransposition in MeCP2 null brain [127]. MeCP2 may contribute directly or indirectly to brain’s genetic heterogeneity by suppressing L1 retrotransposition.

MBD1-deficient adult neural stem cells (ANCs) display elevated expression of the endogenous retrovirus IAP and increased genomic instability, such as aneuploidy [128]. There are also interesting correlations between the MBD proteins and a DNA repair system. MBD1 interacts with methylpurine–DNA glycosylase (MPG), which catalyzes the excision of damaged bases [129].
MBD1 and MPG coexist in heterochromatin to repress transcription. Upon base damage, MBD1 dissociates from chromatin to facilitate MPG-mediated damaged base repair. In line with the observation, MBD1-depleted cells are more sensitive to a DNA-damaging stimulus [129].

Involvement of MBD4 in DNA repair is also reported. Unlike MBD1, MBD4 itself contains a thymine DNA glycosylase domain (TGD). The MBD of MBD4 has a preference for the TG:meCG mismatch DNA, which is a deamination product of the methylated cytosine and is repaired through glycosylation by the TGD [130]. Interestingly, recent reports suggested the involvement of MBD4 in an active DNA demethylation reaction [131, 132]. The removal of DNA methylation can occur by passive dilution if the mark is not replaced after each round of replication, but can also be an active process involving enzymatic activities. Currently, several mechanisms are proposed to mediate DNA demethylation [133]. MBD4 seems to function in a pathway that involves DNA glycosylation followed by base excision repair. Although MBD4 has a preference for mismatch DNA for glycosylation, a hormone-induced phosphorylation of MBD4 by protein kinase C confers glycosylation activity for symmetrical meCpGs. This action is followed by the DNA repair mechanism and results in demethylation and activation of the target gene CYP27B1 [132]. As severe DNA methylation defects were not observed in MBD4 knockout mice, the MBD4-mediated demethylation is not likely to occur genome-wide and seems to act as a loci-specific demethylation mechanism.

**FINAL REMARKS**

The role of DNA methylation in mammals is not yet fully elucidated, although this mark is strongly correlated with transcription states. The role of the DNMT1 protein itself is not completely clear either, as its ability to methylate DNA is only one of its functions. This challenging context has to be kept in mind when thinking about potential functions of methyl-binding proteins.

In this review, we wanted to discuss the role of mammalian methyl-binding proteins in chromatin organization and epigenome maintenance. The occurrence and regulation of these crucial nuclear events involve several MBP functions. MBPs differ from each other first by the way they recognize methyl-CpG; this includes the typical protein domain used, and/or the sequence context required for DNA methylation recognition. Second, the variety among MBPs is highlighted by the distinct set of partners they interact with. These differences allow them to act on different targets, and to ensure multiple particular functions, depending notably on genome and chromatin context, and on the cell cycle stage. Even if only UHRF1 has been shown to be essential, all nine MBPs show to have important and different roles, and do not seem to be redundant [134]. In addition, other MBPs might exist [134] and remain to be fully identified [44, 135].

**Hydroxymethylation and its recognition**

Importantly, 5-methylcytosine (5mC) can be further modified in 5-hydroxymethylcytosine (5hmC) by TET enzymes [136–138]. The role of 5hmC is still under debate. As 5hmC is also resistant to bisulfite treatment, the techniques used so far do not properly discriminate between 5mC and 5hmC [139–141], which sheds doubt on some of the methylation maps generated up to this point. But new methods are being developed to resolve this issue [142, 143]. An interesting question is how this mark is recognized and whether the MBPs could bind 5hmC. Among proteins from MBD family, MeCP2 does not seem to bind 5hmC since replacing 5mC by 5hmC on one DNA strand strongly decreases the affinity of MeCP2-MBD for DNA [144]. Very recently, Frauer et al. [145] showed that UHRF1 was able to bind 5hmC through its SRA domain. It will be interesting to know whether UHRF2 shares this property. The binding of Kaiso and Kaiso-like proteins to 5hmC also remains to be tested.

**DNA demethylation and new epigenetic therapies**

Given the abundance of aberrant methylation events in cancers, the removal of DNA methylation represents an important area of investigation. In addition to a role for MBD4 in this event, it has been proposed that 5hmC could represent an intermediate step in the demethylation process [146]. Further investigations should be promising for future therapeutic applications. Currently, 5-aza-(2-deoxy)cytidine, a well-known drug that inhibits DNA methylation, is widely used in treatment of several hematological malignancies, either alone or in combination with HDAC inhibitors, such as SAHA. In view of recent results, new epigenetic drug
Key Points
- MBPs recruit different sets of epigenetic regulators to manage site-specific chromatin organization and transcriptional repression.
- MeCP2 and UHRF1 are involved in large-scale chromatin organization, notably pericentromeric heterochromatin formation.
- MBD1 and UHRF1 play an important role in maintenance of heterochromatin structures.
- MBPs are involved in processes essential for genome integrity, such as DNA repair.
- Several MBPs are found to be deregulated in cancer.

FUNDING
Work in the Nakao lab is supported in part by Global Centers of Excellence (COE) program, Ministry of Education, Culture, Sports, Science & Technology (MEXT), Japan. Work in the Defossez lab is supported by Centre National de la Recherche Scientifique, Institut National du Cancer, Association pour la Recherche contre le Cancer, Ligue contre le Cancer (Comité de Paris), and Groupement des Entreprises Françaises pour la Lutte contre le Cancer. Alexandra Fournier is the recipient of a post-doctoral fellowship from Institut National du Cancer.

References


