On how mammalian transcription factors recognize methylated DNA

Bethany A. Buck-Koehntop[1] and Pierre-Antoine Defossez[2]

[1]Department of Chemistry; University of Utah; Salt Lake City, UT USA; [2]CNRS UMR7216; Université Paris-Diderot; Paris, France

DNA methylation is an epigenetic mark that is essential for the development of mammals; it is frequently altered in diseases ranging from cancer to psychiatric disorders. The presence of DNA methylation attracts specialized methyl-DNA binding factors that can then recruit chromatin modifiers. These methyl-CpG binding proteins (MBPs) have key biological roles and can be classified into three structural families: methyl-CpG binding domain (MBD), zinc finger, and SET and RING finger-associated (SRA) domain. The structures of MBD and SRA proteins bound to methylated DNA have been previously determined and shown to exhibit two very different modes of methylated DNA recognition. The last piece of the puzzle has been recently revealed by the structural resolution of two different zinc finger proteins, Kaiso and ZFP57, in complex with methylated DNA. These structures show that the two methyl-CpG binding zinc finger proteins adopt differential methyl-CpG binding modes. Nonetheless, there are similarities with the MBD proteins suggesting some commonalities in methyl-CpG recognition across the various MBP domains. These fresh insights have consequences for the analysis of the many other zinc finger proteins present in the genome, and for the biology of methyl-CpG binding zinc finger proteins.

Introduction

DNA methylation is a paradigmatic epigenetic mark. It affects the function of DNA without changing its sequence, and is heritable through replication.[1] DNA methylation has a complex evolutionary history: it is an ancient phenomenon, which exists in archaebacteria, bacteria and eukaryotes, yet it has been lost in many organisms, such as the yeast *S. cerevisiae* or the worm *C. elegans*. In eukaryotes that do possess DNA methylation, its genomic distribution is highly variable between species, likely reflecting different physiological roles for this mark.[2] In the context of this review, we will focus on mammalian organisms, where epigenetic DNA methylation targets only cytosines, mostly in the context of CpG dinucleotides.[3] Mice lacking the DNA methylating enzymes DNMT1, DNMT3a or DNMT3b die during development or shortly after birth, establishing that DNA methylation is essential in mammals.[4]

Why then is DNA methylation so important in mammals? To answer this question, it is useful to know where DNA methylation is found in the genome. The tremendous advances in DNA sequencing technologies have made it possible to obtain full-genome DNA methylation maps in several important cell types such as mouse[5] and human embryonic stem (ES) cells.[6] Less systematic and less expensive alternatives, such as reduced representation bisulfite sequencing (RRBS)[7] or comprehensive high-throughput arrays for relative methylation (CHARM),[8] have also been performed in many tissues. Together, these experiments have provided a picture of where DNA methylation is typically localized within the genome of a given cell, and how much it varies between cell types.[9]

Several independent phenomena seem to account for the essential nature of DNA methylation in mammals.[10] First, DNA methylation controls X-chromosome inactivation and the expression of imprinted

genes, which are key regulators of development. Second, DNA methylation represses the transcription of repeated sequences and prevents relocation of transposable elements, therefore maintaining genome stability. Third, DNA methylation marks the bodies of active genes, with a possible influence on splicing.[11] Fourth, DNA methylation regulates gene expression by repressing promoter activity.

The repressive effects of DNA methylation on the expression of genes and repeated elements seem to involve two major mechanisms.[12] DNA methylation can prevent transcriptional regulators from recognizing their cognate targets; alternatively, it can recruit specialized regulators that specifically bind methylated DNA. These proteins, termed methyl-CpG binding proteins (MBPs), are the focus of this review. They belong to three structural families that we will now discuss: the MBD family, the zinc finger family and the SRA family (**Fig. 1**).

Investigations pioneered in the laboratory of Adrian Bird led to the identification of the first MBP, MeCP2.[13] A search for MeCP2 homologs identified MBD1, MBD2 and MBD4, which all bind methylated DNA with higher affinity than unmethylated DNA, via their conserved MBD homology domain.[14] In contrast, MBD3, MBD5 and MBD6 do not bind methylated DNA, either due to amino acid alterations at critical positions or deletion of key DNA interacting residues in the MBD.[15] Structures for three members of the MBD family in complex with methylated DNA have been determined by either solution NMR spectroscopy (MBD1[16] and MBD2[17]) or X-ray crystallography (MeCP2,[18] **Fig. 2A**). The three proteins adopt a similar overall fold comprised of an α/β sandwich and a comparable mode of binding. The core of the β-sheet interacts with the major groove of the DNA making base specific contacts at the methyl-CpG (mCpG) site through conserved arginine and tyrosine residues. The binding interaction is further stabilized through extensive sugar and phosphate backbone interactions mediated by residues in the α-helix as well as a conserved loop within the β-sheet that forms a stabilized hairpin structure upon DNA recognition.

Interestingly, while the different MBD/methylated DNA complexes share common structural features and modes of mCpG recognition, they exhibit varying sequence preferences outside of the core mCpG site. In particular, MeCP2 selects for targets with an A/T stretch adjacent to the mCpG sites,[19] while MBD1 has highest recognition for mCpG sites in the context of T(mC)GCA or TG(mC)GCA sequences.[20] MBD2 has been shown to exhibit an orientation preference in the recognition of a target derived from the ρ-*globin* promoter [GGAT(mC)GGCTC] that is dependent on the identity of the base pairs flanking the mCpG site, suggesting that it too binds in a sequence-specific context.[17]

A second family of MBPs was discovered through the efforts of Yusuke Nakamura and his team, who established that UHRF1 and UHRF2, two related proteins, could bind methylated DNA via their SRA domains.[21] It was soon discovered that UHRF1 is an essential protein that binds hemimethylated DNA and recruits DNMT1 to facilitate maintenance DNA methylation; in the absence of UHRF1, there is a precipitous loss of DNA methylation.[22,23] The structure of UHRF1 in complex with hemimethylated DNA was promptly solved by X-ray crystallography, yielding important new insights (**Fig. 2B**). In this complex, the methylated cytosine is flipped out of the double helix and inserted into a deep hydrophobic pocket where it is stabilized by planar stacking contacts, hydrogen bonding, and van der Waals interactions that discriminate methylated from unmethylated cytosine.[24-26] This base flipping mechanism is reminiscent of DNA methylating enzymes,[27] but is totally unrelated to the mode of methylated DNA recognition exhibited by the MBD proteins. Further, positioning of an asparagine within proximity of the cross-strand unmethylated cytosine appears to provide selectivity for hemimethylated DNA as cytosine methylation would result in a steric clash with the asparagine, potentially leading to an overall structural instability for the complex. Finally, base specific interactions from UHRF1 are limited to the 5-mCpG/CpG pair, suggesting that in

contrast to the MBD family, sequence context outside of the mCpG site is not critical for recognition. This observation is consistent with involvement of UHRF1 in methylation maintenance of the entire genome, where specific recognition of sequence context outside of the mCpG site would be disadvantageous.

The third, and currently last, family of MBPs also came to light thanks to Egor Prokhortchouk, Adrian Bird and their coworkers who showed that a zinc finger protein, Kaiso, could discriminate methylated from unmethylated DNA.[28] Kaiso was independently shown to also bind a non-methylated consensus site, CTGCNA, called the Kaiso binding site (KBS).[29] Kaiso has two close paralogs in mammalian genomes: Zbtb4 and Zbtb38.[30] These proteins, like Kaiso, bind methylated DNA but can also bind a non-methylated consensus.[31,32] Very recently, another zinc finger protein, ZFP57, was also shown to bind methylated DNA and to act in DNA methylation-dependent maintenance of imprinted genes.[33]

After structural information for the MBD and SRA families of MBPs had been obtained, an important knowledge gap still remained: how do zinc finger proteins recognize methylated DNA? Do they adopt a canonical zinc finger structure? How does their recognition of methylated DNA compare with what we know about the MBD and SRA proteins? Does Kaiso use the same binding mode to engage methylated DNA and its non-methylated target? Do the structures illuminate the known biological differences between MBPs? Two recent papers, reporting the structures of the zinc finger proteins ZFP57[34] (**Fig. 2C**) and Kaiso[35] (**Fig. 2D**) in complex with methylated DNA have provided fresh insights into these questions. We will present these data and discuss their implications.

The Structure of ZFP57 in Complex with the Methylated DNA Sequence TGC(mC)GC

The two Cys_2His_2 zinc fingers (ZFs) in ZFP57, responsible for methylated DNA recognition, adopt the classical $\beta\beta\alpha$-motif, positioning the α-helices for making canonical major groove interactions
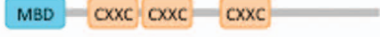
Figure 1. The three families of MBPs and their known targets. Left panels: schematic representation of the human proteins known to bind methylated DNA. Some of their structural domains are indicated: MBD, methyl-binding domain; CXXC, CXXC-type zinc finger; Ubl, ubiquitin-like domain; TTD, tandem tudor domain; PHD, plant homeo domain; SRA, SET and RING-finger associated domain; RING, really interesting new gene domain; BTB/POZ, broad complex, Tramtrack and bric-à-brac/pox virus and zinc finger Domain; ZF, Cys_2His_2 zinc finger domain; KRAB, Krüppel-associated box. The zinc fingers that are necessary for methylated DNA recognition are underlined. Right panel: some of the high-affinity DNA targets identified in vitro. The list is not exhaustive. M represents 5-methylcytosine, H represents 5-hydroxymethylcytosine and N is any nucleotide.

with three base pairs per zinc finger. ZF2 primarily makes base specific contacts with T^4GC^6, while ZF3 recognizes 5-mC^7GC^9. Conversion of the T^4GC^6 site to an ATG sequence abolished binding, indicating that similar to the MBD proteins, sequence context outside of the mCpG is critical for recognition and binding of cognate sequences by ZFP57. This is consistent with in vivo binding maps, showing that ZFP57 is not bound throughout the genome of ES cells, but only at some sites containing the TGC(mC)GC consensus.[33] Direct or water-mediated base specific contacts with the guanidines in the core TGC(mC)GC sequence are facilitated by several arginine residues. Mutation of these key arginines results in a complete loss of this binding interaction.

The methylated cytosines in the core mCpG site are recognized very differently. 5-mC8 is specifically recognized by Glu182, which makes a classical hydrogen bond with the N4-amino group as well as a CH···O hydrogen bond with the C5-methyl. Interestingly, a conservative mutation of Glu182 to glutamine has little effect on 5-mC recognition, presumably due to the N4-amino hydrogen bonding interaction being maintained. Additionally, van der Waals contributions from the side chain of Arg178, which is involved in hydrogen bonding

interactions with the neighboring G7, further contributes to 5-mC8 recognition. This so called "5-mC-Arg-G"[34] interaction is also observed in the MBD structures, indicating a conserved mechanism for 5-mC recognition that is utilized within the context of very different structural motifs. 5-mC7 on the other hand has no base specific contacts from ZFP57 and is surrounded by a network of ordered water molecules. In the MBD/ methylated DNA structures, an increased solvation at one of the 5-mC bases is also observed and mediates interactions between the 5-mC and a highly conserved tyrosine residue, enhancing the overall binding.[18]
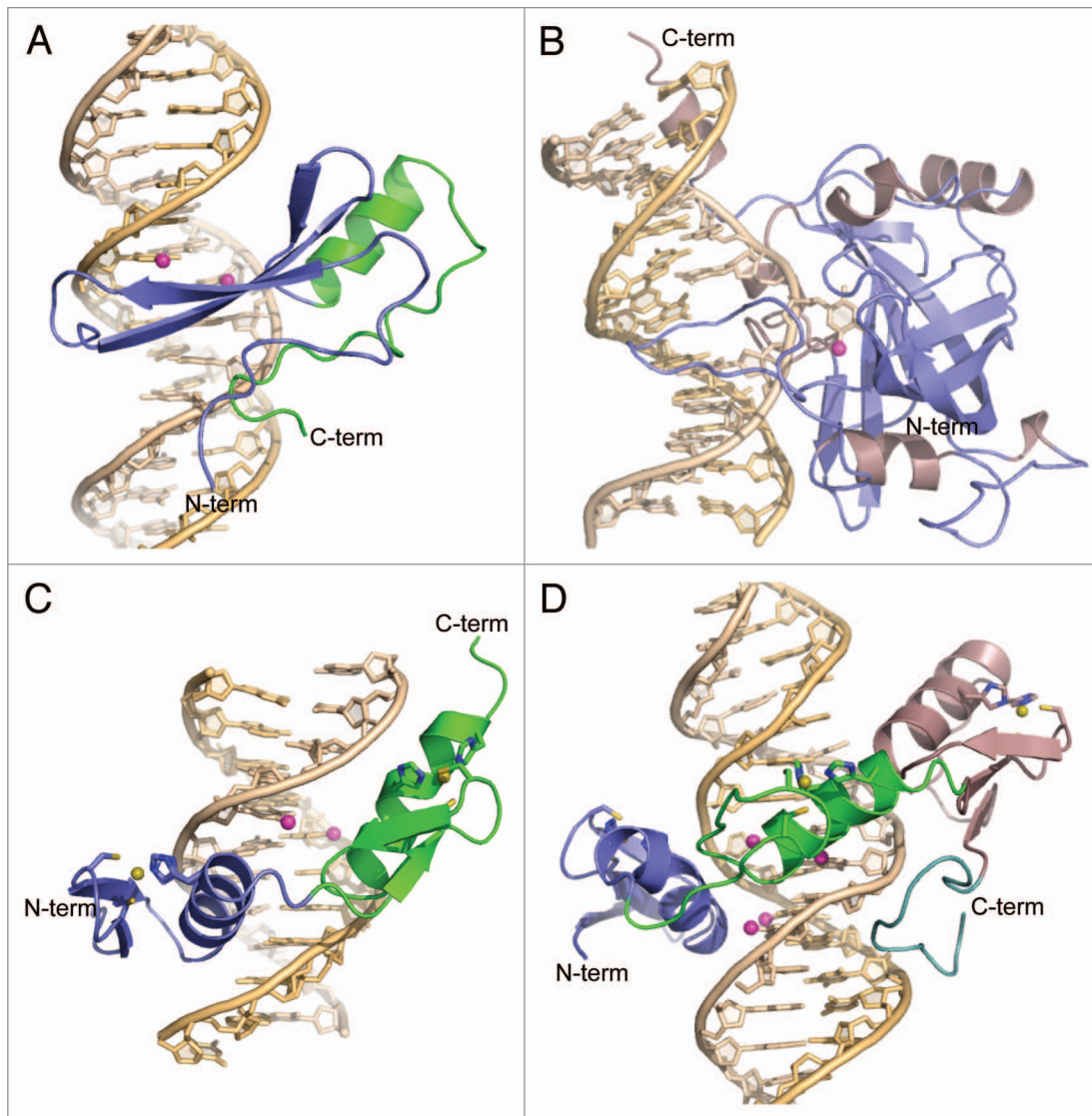
Figure 2. Representative structures for each MBP family. (**A**) Crystal structure of the MBD of human MeCP2 in complex with a methylated DNA consensus site derived from the *brain-derived neurotrophic factor* (BDNF) promoter (PDB 3C2I).[18] (**B**) Crystal structure of the SRA domain of human UHRF1 in complex with hemimethylated DNA (PDB 3CLZ).[25] (**C**) Crystal structure of the mouse Cys_2His_2 zinc finger protein ZFP57 in complex with a methylated DNA target known to be localized in imprinting control regions (PDB 4GZN).[34] (**D**) Crystal structure of the human Cys_2His_2 zinc finger protein Kaiso in complex with a methylated DNA consensus site derived from the *E-cadherin* (CDH1) promoter (PDB 4F6N).[35] Pink spheres represent methyl groups in the methylated cytosines.

The Structure of Kaiso in Complex with Methylated DNA or the KBS

Kaiso has been crystallized in complex with two different nucleotides: the first, MeECad, contains two consecutive mCpG sites and is derived from the *E-cadherin* promoter [C(mC)G(mC)GT];[28] the second, KBS, contains no CpGs in the sequence and is unmethylated (TCCTGCCA).[29] The overall structures of Kaiso in complex with the KBS or methylated DNA are nearly identical and manifest similar modes of DNA

recognition. Unlike ZFP57, Kaiso utilizes three Cys_2His_2 zinc fingers and additional N- and C-terminal extensions to provide structural stability and enhance the overall binding affinity. ZF2 adopts the standard $\beta\beta\alpha$ fold, while ZF1 and ZF3 each contain three-stranded β-sheets with $\beta\beta\beta\alpha$ and $\beta\beta\alpha\beta$ motifs, respectively. While all three zinc fingers orient with their helices in the DNA major groove, in contrast to ZFP57 and other canonical Cys_2His_2 zinc fingers,[35] Kaiso only contacts a core of 5–6 base pairs between the three zinc fingers. Base specific interactions are

mediated through residues located within ZF1 and ZF2. ZF3 contributes additional electrostatic interactions with the phosphate backbone, but primarily functions as a scaffold upon which the C-terminal extension (CTE) can form a stabilizing loop through interactions with the helix of ZF2, positioning the remainder of the extension for making minor groove DNA contacts. These minor groove contacts are essential for high-affinity DNA recognition by Kaiso as truncation of the CTE beyond ZF3 dramatically reduces the binding interaction.[46] Of all the MBP

structures solved to date, Kaiso appears to be the only one that utilizes both major and minor groove interactions to confer high-affinity binding interactions. The biological necessity and implications of this in terms of distinguishing binding targets is not entirely clear.

Further, Kaiso is the only MBP that has been identified to have preference for recognizing two consecutive mCpG sites.[28] From the structures, it can be seen that this is partially dictated by the base specific interactions from Arg511 (ZF1) which forms cross-strand hydrogen bonds with guanines of the central two 5-mCpG pairs. While this interaction is insensitive to the cytosine methylation status, it does dictate that a guanine must occupy both positions (as is the case in KBS). Similar to ZFP57, E535 (ZF2) in Kaiso makes classical and CH···O hydrogen bonds with both 5-mCs in the first palindromic 5-mCpG pair, which is further stabilized by a "5-mC-Arg-G" interaction between Arg511 and the coding strand 5-mC. In KBS recognition, Glu535 makes similar contacts with cross-strand thymine and cytosine bases, presumably through a protonation of the carboxylic side chain that allows for the formation of a hydrogen bond between the Glu535 $O^{\varepsilon2}$ and the thymine O4. Interestingly, mutation of Glu182 to an alanine in ZFP57 had a minimal effect on methylated DNA recognition, while mutation of the equivalent to Glu535 in the Kaiso paralog ZBTB4 (Glu350) resulted in a substantial reduction in mCpG recognition.[32] The 5-mCs in the second mCpG site are primarily localized within a hydrophobic pocket formed by residues in ZF1.

Biological Consequences

The possible existence of other methyl-CpG binding transcription factors. There are 700 zinc finger proteins in the human proteome,[36] most of which are uncharacterized. The KRAB-zinc finger family (which contains ZFP57), in particular, has witnessed a rapid expansion in mammals.[37] It is then possible that ZFP57 paralogs may exist that can also bind methylated DNA. At first evaluation, Kaiso has no close paralogs besides ZBTB4 and ZBTB38, but identification

of the critical residues for interaction with methylated cytosines now makes it possible to look for other zinc finger proteins in the genome that share residues with Kaiso at these key locations, and to test whether they also bind methylated DNA.

The structural work has revealed a small degree of similarities in the modes of methylated DNA recognition between MBD and zinc fingers, but it has also shown that three completely distinct protein folds can be used to discriminate methylated from non-methylated DNA. There is no reason to exclude the possibility that transcription factors in yet other protein families may have the same capacity. Along these lines, the transcription factor RBP-J, which belongs to the CSL family, was recently shown to have higher affinity for methylated than for unmethylated DNA in vitro.[38]

Attempts to engineer zinc finger proteins that can specifically discriminate methylated cytosines from unmethylayed cytosine and thymine have also yielded interesting results. Using a phage display approach, ZF-2 and -3 of the three Zif268 canonical Cys_2His_2 zinc fingers were randomized to search for a highly selective binder of 5-mC.[39] Out of these studies, two clones were determined to have preferential binding to methylated DNA sites. Consistently, the two clones utilize an arginine to make base specific contacts with the guanines in the mCpG site, reminiscent of what has been observed for both the MBD and zinc finger protein structures. Additionally, one of these engineered constructs utilizes a glutamine for making direct contact with the 5-mC in a comparable manner to that observed for the Glu182Gln mutation in ZFP57. The other construct selected for a tyrosine in proximity of the 5-mC, suggesting a possible mode of water-mediated interactions similar to 5-mC recognition exhibited by the MBD family. These findings, in the context of the structural knowledge for both the MBD and zinc finger families of MBPs, suggest that some conserved mechanisms for reading mCpG sites exist across MBPs.

Binding to hemimethylated DNA: molecular directionality and DNA methylation maintenance. An unexpected finding from the work on Kaiso and

ZFP57 is that both proteins bind methylated DNA asymmetrically, with one strand contributing most of the binding affinity. Therefore, they are "hemimethyl-DNA" as much as "methyl-DNA" binding proteins. It is currently unclear whether the in vivo targets of ZFP57 and Kaiso are symmetrically methylated or hemi-methylated, and a combination of ChIP and bisulfite sequencing[40] will be necessary to ascertain this point. Nonetheless, if the in vitro findings are reflective of the situation in cells, there may be three biological consequences for this observed hemimethylated DNA recognition. First, it is possible that this asymmetry provides directionality for the MBPs in identifying their cognate binding sites. The second consequence of asymmetrical methylated DNA recognition by ZFP57 and Kaiso is that these proteins might facilitate the orientation of protein complexes, such as chromatin remodeling co-repressors, at the target site. This idea was suggested for MBD2;[17] even though it binds DNA methylated on both strands, it shows a clear directionality that has not been demonstrated for MeCP2 or MBD1. The third consequence is that Kaiso and ZFP57 might bind hemimethylated sites to mediate proper DNA methylation during development. Kaiso mutant mice have no overt phenotype and no indication of gross methylation abnormalities,[41] but this does not rule out more subtle defects. In contrast, ZFP57 mutant mice do show a loss of maintenance at certain imprinted loci that is consistent with this possibility.[42] Indeed, ZFP57 has been shown to recognize methylated DNA in the context of imprinting control regions, which then interacts with the protein TRIM28, recruiting DNMT1 and UHRF1 to the target site.[33]

DNA hydroxymethylation. A breakthrough in the field of epigenetics has occurred within the past few years, with the discovery that methylated cytosines can be remodeled through an active demethylation process.[43] One pathway involves the TET enzymes, which mediate the demethylation of 5-mC through a series of oxidative intermediates including 5-hydromethylcytosine (5hmC), 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC). It is currently unclear whether

these species are merely reaction intermediates along the demethylation pathway, or whether they participate in specific regulatory functions in the cell, possibly by recruiting dedicated binders. Out of all the MBPs identified to date, only UHRF1 has been shown to bind 5hmC in addition to 5-mC.[44] In contrast, MBD proteins have been shown to be incapable of binding 5hmC containing sites,[45] and the presence of 5hmC also decreases binding by ZFP57.[34] Given that there are three structurally divergent domains that can specifically distinguish methylated cytosines from unmethylated cytosines, it will be interesting to see what diversity in structural domains have evolved to recognize any or all of these oxidative intermediates.

Concluding Remarks

With the addition of structural knowledge for the final zinc finger family of MBPs, we are now able to discern commonalities as well as differences in the mode of methyl-CpG recognition across the three MBP families. Nonetheless, several fundamental questions still exist. There is increasing evidence to suggest that the various MBPs have differential DNA targets, and for the MBD and zinc finger families it appears that this could be dictated by sequence context flanking the mCpG site. However, a comprehensive analysis of the in vivo DNA targets for these proteins has yet to be established. This analysis will facilitate our understanding for why it was necessary to evolve three distinctly different structural domains for methyl-CpG recognition, and for determining a more detailed understanding of the complex biological functions these proteins adopt in both normal cellular function as well as disease.

References

1. Reik W. Stability and flexibility of epigenetic gene regulation in mammalian development. Nature 2007; 447:425-32; PMID:17522676; http://dx.doi.org/10.1038/nature05918.

2. Suzuki MM, Bird A. DNA methylation landscapes: provocative insights from epigenomics. Nat Rev Genet 2008; 9:465-76; PMID:18463664; http://dx.doi.org/10.1038/nrg2341.

3. Lister R, Ecker JR. Finding the fifth base: genome-wide sequencing of cytosine methylation. Genome Res 2009; 19:959-66; PMID:19273618; http://dx.doi.org/10.1101/gr.083451.108.

4. Li E. Chromatin modification and epigenetic reprogramming in mammalian development. Nat Rev Genet 2002; 3:662-73; PMID:12209141; http://dx.doi.org/10.1038/nrg887.

5. Stadler MB, Murr R, Burger L, Ivanek R, Lienert F, Schöler A, et al. DNA-binding factors shape the mouse methylome at distal regulatory regions. Nature 2011; 480:490-5; PMID:22170606.

6. Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J, et al. Human DNA methylomes at base resolution show widespread epigenomic differences. Nature 2009; 462:315-22; PMID:19829295; http://dx.doi.org/10.1038/nature08514.

7. Boyle P, Clement K, Gu H, Smith ZD, Ziller M, Fostel JL, et al. Gel-free multiplexed reduced representation bisulfite sequencing for large-scale DNA methylation profiling. Genome Biol 2012; 13:R92; PMID:23034176; http://dx.doi.org/10.1186/gb-2012-13-10-r92.

8. Irizarry RA, Ladd-Acosta C, Carvalho B, Wu H, Brandenburg SA, Jeddeloh JA, et al. Comprehensive high-throughput arrays for relative methylation (CHARM). Genome Res 2008; 18:780-90; PMID:18316654; http://dx.doi.org/10.1101/gr.7301508.

9. Landan G, Cohen NM, Mukamel Z, Bar A, Molchadsky A, Brosh R, et al. Epigenetic polymorphism and the stochastic formation of differentially methylated regions in normal and cancerous tissues. Nat Genet 2012; 44:1207-14; PMID:23064413; http://dx.doi.org/10.1038/ng.2442.

10. Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nat Rev Genet 2012; 13:484-92; PMID:22641018; http://dx.doi.org/10.1038/nrg3230.

11. Shukla S, Kavak E, Gregory M, Imashimizu M, Shutinoski B, Kashlev M, et al. CTCF-promoted RNA polymerase II pausing links DNA methylation to splicing. Nature 2011; 479:74-9; PMID:21964334; http://dx.doi.org/10.1038/nature10442.

12. Klose RJ, Bird AP. Genomic DNA methylation: the mark and its mediators. Trends Biochem Sci 2006; 31:89-97; PMID:16403636; http://dx.doi.org/10.1016/j.tibs.2005.12.008.

13. Lewis JD, Meehan RR, Henzel WJ, Maurer-Fogy I, Jeppesen P, Klein F, et al. Purification, sequence, and cellular localization of a novel chromosomal protein that binds to methylated DNA. Cell 1992; 69:905-14; PMID:1606614; http://dx.doi.org/10.1016/0092-8674(92)90610-O.

14. Hendrich B, Bird A. Identification and characterization of a family of mammalian methyl-CpG binding proteins. Mol Cell Biol 1998; 18:6538-47; PMID:9774669.

15. Laget S, Joulie M, Le Masson F, Sasai N, Christians E, Pradhan S, et al. The human proteins MBD5 and MBD6 associate with heterochromatin but they do not bind methylated DNA. PLoS One 2010; 5:e11982; PMID:20700456; http://dx.doi.org/10.1371/journal.pone.0011982.

16. Ohki I, Shimotake N, Fujita N, Jee J, Ikegami T, Nakao M, et al. Solution structure of the methyl-CpG binding domain of human MBD1 in complex with methylated DNA. Cell 2001; 105:487-97; PMID:11371345; http://dx.doi.org/10.1016/S0092-8674(01)00324-5.

17. Scarsdale JN, Webb HD, Ginder GD, Williams DC Jr. Solution structure and dynamic analysis of chicken MBD2 methyl binding domain bound to a target-methylated DNA sequence. Nucleic Acids Res 2011; 39:6741-52; PMID:21531701; http://dx.doi.org/10.1093/nar/gkr262.

18. Ho KL, McNae IW, Schmiedeberg L, Klose RJ, Bird AP, Walkinshaw MD. MeCP2 binding to DNA depends upon hydration at methyl-CpG. Mol Cell 2008; 29:525-31; PMID:18313390; http://dx.doi.org/10.1016/j.molcel.2007.12.028.

19. Klose RJ, Sarraf SA, Schmiedeberg L, McDermott SM, Stancheva I, Bird AP. DNA binding selectivity of MeCP2 due to a requirement for A/T sequences adjacent to methyl-CpG. Mol Cell 2005; 19:667-78; PMID:16137622; http://dx.doi.org/10.1016/j.molcel.2005.07.021.

20. Clouaire T, de Las Heras JI, Merusi C, Stancheva I. Recruitment of MBD1 to target genes requires sequence-specific interaction of the MBD domain with methylated DNA. Nucleic Acids Res 2010; 38:4620-34; PMID:20378711; http://dx.doi.org/10.1093/nar/gkq228.

21. Unoki M, Nishidate T, Nakamura Y. ICBP90, an E2F-1 target, recruits HDAC1 and binds to methyl-CpG through its SRA domain. Oncogene 2004; 23:7601-10; PMID:15361834; http://dx.doi.org/10.1038/sj.onc.1208053.

22. Bostick M, Kim JK, Estève PO, Clark A, Pradhan S, Jacobsen SE. UHRF1 plays a role in maintaining DNA methylation in mammalian cells. Science 2007; 317:1760-4; PMID:17673620; http://dx.doi.org/10.1126/science.1147939.

23. Sharif J, Muto M, Takebayashi S, Suetake I, Iwamatsu A, Endo TA, et al. The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. Nature 2007; 450:908-12; PMID:17994007; http://dx.doi.org/10.1038/nature06397.

24. Arita K, Ariyoshi M, Tochio H, Nakamura Y, Shirakawa M. Recognition of hemi-methylated DNA by the SRA protein UHRF1 by a base-flipping mechanism. Nature 2008; 455:818-21; PMID:18772891; http://dx.doi.org/10.1038/nature07249.

25. Avvakumov GV, Walker JR, Xue S, Li Y, Duan S, Bronner C, et al. Structural basis for recognition of hemi-methylated DNA by the SRA domain of human UHRF1. Nature 2008; 455:822-5; PMID:18772889; http://dx.doi.org/10.1038/nature07273.

26. Hashimoto H, Horton JR, Zhang X, Bostick M, Jacobsen SE, Cheng X. The SRA domain of UHRF1 flips 5-methylcytosine out of the DNA helix. Nature 2008; 455:826-9; PMID:18772888; http://dx.doi.org/10.1038/nature07280.

27. Song J, Teplova M, Ishibe-Murakami S, Patel DJ. Structure-based mechanistic insights into DNMT1-mediated maintenance DNA methylation. Science 2012; 335:709-12; PMID:22323818; http://dx.doi.org/10.1126/science.1214453.

28. Prokhortchouk A, Hendrich B, Jørgensen H, Ruzov A, Wilm M, Georgiev G, et al. The p120 catenin partner Kaiso is a DNA methylation-dependent transcriptional repressor. Genes Dev 2001; 15:1613-8; PMID:11445535; http://dx.doi.org/10.1101/gad.198501.

29. Daniel JM, Spring CM, Crawford HC, Reynolds AB, Baig A. The p120(ctn)-binding partner Kaiso is a bi-modal DNA-binding protein that recognizes both a sequence-specific consensus and methylated CpG dinucleotides. Nucleic Acids Res 2002; 30:2911-9; PMID:12087177; http://dx.doi.org/10.1093/nar/gkf398.

30. Sasai N, Defossez PA. Many paths to one goal? The proteins that recognize methylated DNA in eukaryotes. Int J Dev Biol 2009; 53:323-34; PMID:19412889; http://dx.doi.org/10.1387/ijdb.082652ns.

31. Filion GJ, Zhenilo S, Salozhin S, Yamada D, Prokhortchouk E, Defossez PA. A family of human zinc finger proteins that bind methylated DNA and repress transcription. Mol Cell Biol 2006; 26:169-81; PMID:16354688; http://dx.doi.org/10.1128/MCB.26.1.169-181.2006.

32. Sasai N, Nakao M, Defossez PA. Sequence-specific recognition of methylated DNA by human zinc-finger proteins. Nucleic Acids Res 2010; 38:5015-22; PMID:20403812; http://dx.doi.org/10.1093/nar/gkq280.

33. Quenneville S, Verde G, Corsinotti A, Kapopoulou A, Jakobsson J, Offner S, et al. In embryonic stem cells, ZFP57/KAP1 recognize a methylated hexanucleotide to affect chromatin and DNA methylation of imprinting control regions. Mol Cell 2011; 44:361-72; PMID:22055183; http://dx.doi.org/10.1016/j.molcel.2011.08.032.

34. Liu Y, Toh H, Sasaki H, Zhang X, Cheng X. An atomic model of Zfp57 recognition of CpG methylation within a specific DNA sequence. Genes Dev 2012; 26:2374-9; PMID:23059534; http://dx.doi.org/10.1101/gad.202200.112.

35. Buck-Koehntop BA, Stanfield RL, Ekiert DC, Martinez-Yamout MA, Dyson HJ, Wilson IA, et al. Molecular basis for recognition of methylated and specific DNA sequences by the zinc finger protein Kaiso. Proc Natl Acad Sci U S A 2012; 109:15229-34; PMID:22949637; http://dx.doi.org/10.1073/pnas.1213726109.

36. Vaquerizas JM, Kummerfeld SK, Teichmann SA, Luscombe NM. A census of human transcription factors: function, expression and evolution. Nat Rev Genet 2009; 10:252-63; PMID:19274049; http://dx.doi.org/10.1038/nrg2538.

37. Emerson RO, Thomas JH. Adaptive evolution in zinc finger transcription factors. PLoS Genet 2009; 5:e1000325; PMID:19119423; http://dx.doi.org/10.1371/journal.pgen.1000325.

38. Bartels SJ, Spruijt CG, Brinkman AB, Jansen PW, Vermeulen M, Stunnenberg HG. A SILAC-based screen for Methyl-CpG binding proteins identifies RBP-J as a DNA methylation and sequence-specific binding protein. PLoS One 2011; 6:e25884; PMID:21991380; http://dx.doi.org/10.1371/journal.pone.0025884.

39. Isalan M, Choo Y. Engineered zinc finger proteins that respond to DNA modification by HaeIII and HhaI methyltransferase enzymes. J Mol Biol 2000; 295:471-7; PMID:10623539; http://dx.doi.org/10.1006/jmbi.1999.3366.

40. Brinkman AB, Gu H, Bartels SJ, Zhang Y, Matarese F, Simmer F, et al. Sequential ChIP-bisulfite sequencing enables direct genome-scale investigation of chromatin and DNA methylation cross-talk. Genome Res 2012; 22:1128-38; PMID:22466170; http://dx.doi.org/10.1101/gr.133728.111.

41. Prokhortchouk A, Sansom O, Selfridge J, Caballero IM, Salozhin S, Aithozhina D, et al. Kaiso-deficient mice show resistance to intestinal cancer. Mol Cell Biol 2006; 26:199-208; PMID:16354691; http://dx.doi.org/10.1128/MCB.26.1.199-208.2006.

42. Li X, Ito M, Zhou F, Youngson N, Zuo X, Leder P, et al. A maternal-zygotic effect gene, Zfp57, maintains both maternal and paternal imprints. Dev Cell 2008; 15:547-57; PMID:18854139; http://dx.doi.org/10.1016/j.devcel.2008.08.014.

43. Wu SC, Zhang Y. Active DNA demethylation: many roads lead to Rome. Nat Rev Mol Cell Biol 2010; 11:607-20; PMID:20683471; http://dx.doi.org/10.1038/nrm2950.

44. Frauer C, Hoffmann T, Bultmann S, Casa V, Cardoso MC, Antes I, et al. Recognition of 5-hydroxymethyl-cytosine by the Uhrf1 SRA domain. PLoS One 2011; 6:e21306; PMID:21731699; http://dx.doi.org/10.1371/journal.pone.0021306.

45. Valinluck V, Tsai HH, Rogstad DK, Burdzy A, Bird A, Sowers LC. 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 2004; 32:4100-8; PMID:15302911; http://dx.doi.org/10.1093/nar/gkh739.

46. Buck-Koehntop BA, Martinez-Yamout MA, Dyson HJ, Wright PE. Kaiso uses all three zinc fingers and adjacent sequence motifs for high affinity binding to sequence-specific and methyl-CpG DNA targets. FEBS Lett 2012; 586:734-739; PMID:22300642