Proteins That Read DNA Methylation

11

Ke Liu, Takashi Shimbo, Xiaosheng Song, Paul A. Wade, and Jinrong Min

Abstract

Covalent modification of DNA via deposition of a methyl group at the 5′ position on cytosine residues alters the chemical groups available for interaction in the major groove of DNA. This modification, thereby, alters the affinity and specificity of DNA-binding proteins; some of them favor interaction with methylated DNA, and others disfavor it. Molecular recognition of cytosine methylation by proteins often initiates sequential regulatory events that impact gene expression and chromatin structure. The known methyl-DNA-binding proteins have unique domains responsible for DNA methylation recognition: (1) the methyl-CpG-binding domain (MBD), (2) the SET- and RING finger-associated domain (SRA), and (3) some of TF families, such as the C2H2 zinc finger domain, basic helix-loop-helix (bHLH), basic leucine-zipper (bZIP), and homeodomain proteins. Structural analyses have revealed that each domain has a characteristic methylated DNA-binding pattern, and the difference in the recognition mechanisms renders the DNA methylation mark able to transmit complicated biological information. Recent genetic and genomic studies have revealed novel functions of methyl-DNA-binding proteins. These emerging data have also provided glimpses into how methyl-DNA-binding proteins possess unique features and, presumably, functions. In this chapter, we summarize structural and biochemical analyses elucidating the mechanisms for recognition of DNA methylation and correlate this information with emerging genomic and functional data.

Keywords

DNA methylation · MBD · Methylcytosine-binding proteins (MBPs) · Epigenetics

Abbreviations

5caC	5-carboxylcytosine
5fC	5-formylcytosine
5hmC	5-hydroxymethylcytosine
5mC	5-methylcytosine
bHLH	Basic helix-loop-helix
bZIP	Basic leucine-zipper
CTCF	Multidomain CCCTC-binding factor

K. Liu · X. Song · J. Min (✉)
Hubei Key Laboratory of Genetic Regulation and Integrative Biology, School of Life Sciences, Central China Normal University, Wuhan, PR China
e-mail: keliu2015@mail.ccnu.edu.cn; minjinrong@ccnu.edu.cn

T. Shimbo · P. A. Wade
Epigenetics and Stem Cell Biology Laboratory, National Institute of Environmental Health Sciences, Durham, NC, USA
e-mail: wadep2@niehs.nih.gov

A. Jeltsch, R. Z. Jurkowska (eds.), *DNA Methyltransferases - Role and Function*, Advances in Experimental Medicine and Biology 1389, https://doi.org/10.1007/978-3-031-11454-0_11

DNMT	DNA methyltransferase
GD	Glycosylase domain
MARs/SARs	Matrix/scaffold attachment regions
MBD	Methyl-CpG-binding domain
NSC	Neural stem cell
NuRD	Nucleosome remodeling deacetylase
PHD	Plant homeodomain
RING	Really interesting new gene domain
SRA	SET- and RING finger-associated domain
TF	Transcriptional factor
TRD	Transcriptional repression domain
TTD	Tandem Tudor domain
UBL	Ubiquitin-like domain
ZF	Zinc finger

11.1 Introduction

DNA methylation serves as a fundamental component of epigenetic regulation; dysregulation of DNA methylation impacts multiple biological processes, including tumorigenesis (Schubeler 2015). In mammals, most DNA methylation occurs in the context of the CpG dinucleotide. In general, 70–80% of the CpGs in mammalian genomes are methylated (Bird 2002). Nevertheless, cytosine methylation is also present at CpH (non-CpG methylation, H=A, T, or C) sites (Ramsahoye et al. 2000; Woodcock et al. 1987), which accounts for about 25% of the total cytosine methylation in both neurons and embryonic stem cells (*ESCs*). Similar to mCpG, mCpH mainly contributes to transcriptional repression and imprinting (Guo et al. 2014; Sanchez-Mut et al. 2016). mCpH is mainly located in a region with low CpG density, and is established and maintained by DNMT3A (Guo et al. 2014; Ramsahoye et al. 2000). Non-CpG methylation, although less abundant than CpG methylation, occurs in virtually all human tissues and is involved in the repression of development-related genes during stem cell differentiation (Schultz et al. 2015).

Although DNA methylation has historically been depicted as a relatively static modification, recent studies have revealed that the methyl group on cytosine can be further modified by oxidation; Fe (II)- and α-ketoglutarate-dependent oxidation mediated by ten-eleven translocation (TET) dioxygenases converts 5-methylcytosine (5mC) into 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC), and 5-carboxylcytosine (5caC) (Kohli and Zhang 2013; Kriaucionis and Tahiliani 2014; Tahiliani et al. 2009).

The "reader proteins," referred to as methylcytosine-binding proteins (MBPs), specifically recognize DNA methylation marks and initiate signaling pathways. MBPs often interact with other proteins and serve as hubs to recruit effector proteins to particular loci. It is the particular collection of effector proteins associated with each MBP and not the act of binding methylated CpG per se that typically elicits downstream transcriptional effects. The MBPs can be classified using structural information into three major families, each characterized by the presence of a critical recognition domain: the methyl-CpG-binding domain (MBD), the SET- and RING finger-associated domain (SRA), and some of TF families (Fig. 11.1). Although DNA methylation precludes the interaction of many TFs, such as MYC, CREB, NRF1, and members of the E2F family, with their specific DNA recognition sequences (Domcke et al. 2015; Tate and Bird 1993), some TFs, such as the extended homeodomain family, prefer methylated CpG sequences (Kribelbauer et al. 2017; Yin et al. 2017). Each MBP has unique features, including DNA-binding preferences, expression patterns, or protein–protein interaction partners, and has critical roles in various biological contexts. The domain architecture of each protein family is unique, and comparisons between these structures enable insights into the similarities and differences in recognition of methylcytosine, presenting opportunities for a single modification, CpG methylation, to nucleate different effectors (Fig. 11.1).

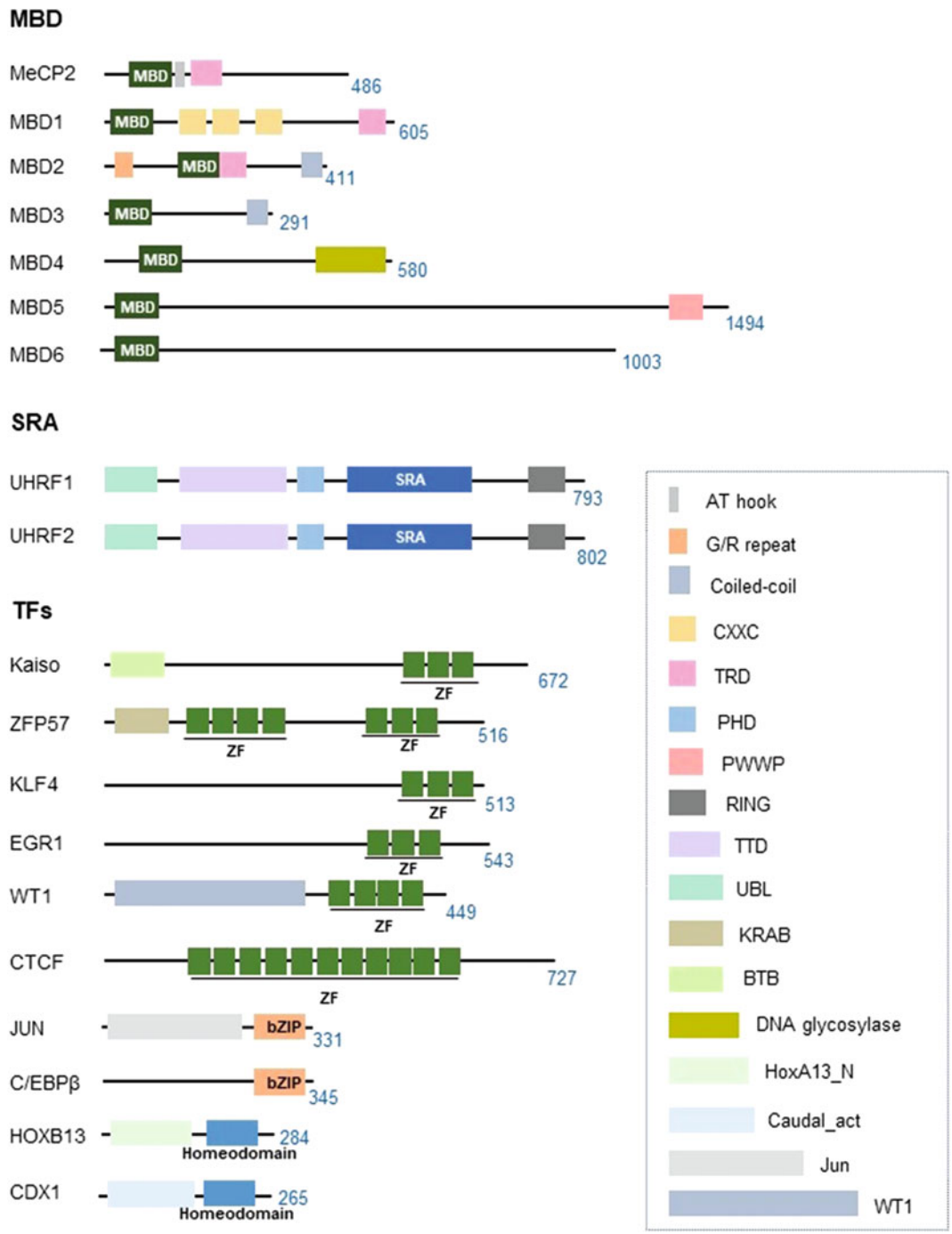

Fig. 11.1 Domain structures of methylated DNA-binding proteins. MeCP2 (NG_007107.2), MBD1 (NP_001191065.1), MBD2 (NP_003918.1), MBD3 (NP_001268382.1), MBD4 (NP_001263199.1), MBD5 (NP_060798.2), MBD6 (NP_060798.2), UHRF1 (NP_001041666.1), UHRF2 (NP_690856.1), Kaiso (NP_001171671.1), ZFP57 (Q9NU63-2), KLF4 (NP_001300981.1), EGR1 (NP_001955.1), WT1 (P19544.2), CTCF (XP_016878357.1), JUN (NP_002219.1), C/EBPβ (NP_005185.2), HOXB13 (NP_006352.2), CDX1 (NP_001795.2), CDX2 (NP_001256.3). *TRD* transcriptional repression domain, *CXXC* CXXC type zinc finger domain, *UBL* ubiquitin-like domain, *PHD* plant homeodomain, *SRA* SET- and RING finger-associated domain, *PWWP* Pro-Trp-Trp-Pro domain, *RING* really interesting new gene finger domain, *TTD* tandem tudor domain, *BTB* BR-C, ttk, and bab domain, *C2H2* C2H2 type zinc finger domain, *KRAB* Kruppel-associated box domain, *bZIP* basic leucine zipper domain, *WT1* Wilms' tumor 1

11.2 The Methyl-CpG-Binding Domain Family

11.2.1 MeCP2

MeCP2 was the first MBP to be purified biochemically and its cDNA cloned and sequenced (Lewis et al. 1992; Meehan et al. 1992; Meehan et al. 1989). The cDNA initially cloned by Bird and colleagues codes for a protein of 492 amino acids that contains an N-terminal MBD domain and a transcriptional repression domain in the C-terminal region. Surprisingly, MeCP2 was subsequently found to be homologous to a matrix attachment binding protein from chicken known as ARBP (attachment region binding protein), a protein identified by biochemical assays based on its binding to a sequence motif (5′-GGTGT-3′) found in matrix/scaffold attachment regions (MARs/SARs) (von Kries et al. 1991; Weitzel et al. 1997). The functional domain responsible for the binding of methylated CpG sites in MeCP2 was subsequently identified and termed the methyl-CpG-binding domain (MBD) that became the archetypal methyl-CpG-binding domain. Subsequent homology searches using the MBD from MeCP2 led to the identification of the remaining MBD family proteins (Hendrich and Bird 1998; Nan et al. 1993). MeCP2 thus represents the founding member of the MBD protein family.

Structure analysis revealed that MeCP2 recognizes the fully methylated CpG dinucleotide using a 5mC-Arg-Gua triad (Fig. 11.2a). Two arginine residues within the MBD (R111 and R133) each bind to a guanine with bidentate hydrogen bonds and to the 5mC with cation–π interactions, where cytosine methylation expands the aromatic ring structure and strengthens the cation–π interactions between the methylcytosine and the guanidinium group of the arginine residues (Ho et al. 2008; Lei et al. 2019; Liu et al. 2013, 2018; Zou et al. 2012). In addition, a tyrosine residue forms water-mediated hydrogen bonds with one of the two cytosine methyl groups. Although the SELEX experiments refined the model, stipulating that high-affinity interaction with methylated DNA was facilitated if the methylated CpG dinucleotide was flanked by A/T base pairs on each side (Ghosh et al. 2010; Klose et al. 2005), ~~s~~tructural studies by different laboratories showed that the specific DNA recognition is largely confined to the mCpG dinucleotide, and no base-specific interaction was observed outside the mCpG dinucleotide (Fig. 11.2a) (Ho et al. 2008; Liu et al. 2018; Ohki et al. 2001; Otani et al. 2013). In addition, structural analysis and binding studies also confirmed that MeCP2 recognizes TpG DNA with a preference for GTG DNA (Fig. 11.2b) (Lei et al. 2019; Liu et al. 2018). The MeCP2 MBD binds to mCAC or CAC DNA by recognizing their complementary GTG trinucleotide, which explains why MeCP2 has a comparable binding affinity to both mCAC and hmCAC (hydroxymethylated CAC) DNAs (Fig. 11.2b) (Kinde et al. 2015; Lei et al. 2019). The above findings are also consistent with the original finding that chicken ARBP (or cMeCP2) binds to a conserved GGTGT DNA motif found in the MARs/SARs DNA regions that lacks any methylated CpG dinucleotides (Buhrmester et al. 1995; von Kries et al. 1991; Weitzel et al. 1997). The central GTG trinucleotide of the MARs DNA elements is crucial for its binding to cMeCP2, and mutating either of the guanine bases in this trinucleotide significantly reduced this binding (Buhrmester et al. 1995; Weitzel et al. 1997).

MeCP2 displays relatively high expression in neurons, where this level is approximate to that of histone octamers (Skene et al. 2010). MeCP2 plays important role in the normal chromatin architecture in neurons, and its mutations result in the neurodevelopmental disorder Rett syndrome (RTT) in humans (Amir et al. 1999; Chen et al. 2001; Guy et al. 2001). However, despite the important roles of MeCP2 mutations in development and disease, MeCP2 deletion in mice has only minimal impact on global gene regulation. Thus, detailed mechanistic insights into how disruption of MeCP2 causes developmental failure or Rett syndrome are currently lacking. MeCP2 has been found to bind to methylated DNA at CpG and CpA dinucleotides broadly throughout the genome (Gabel et al.

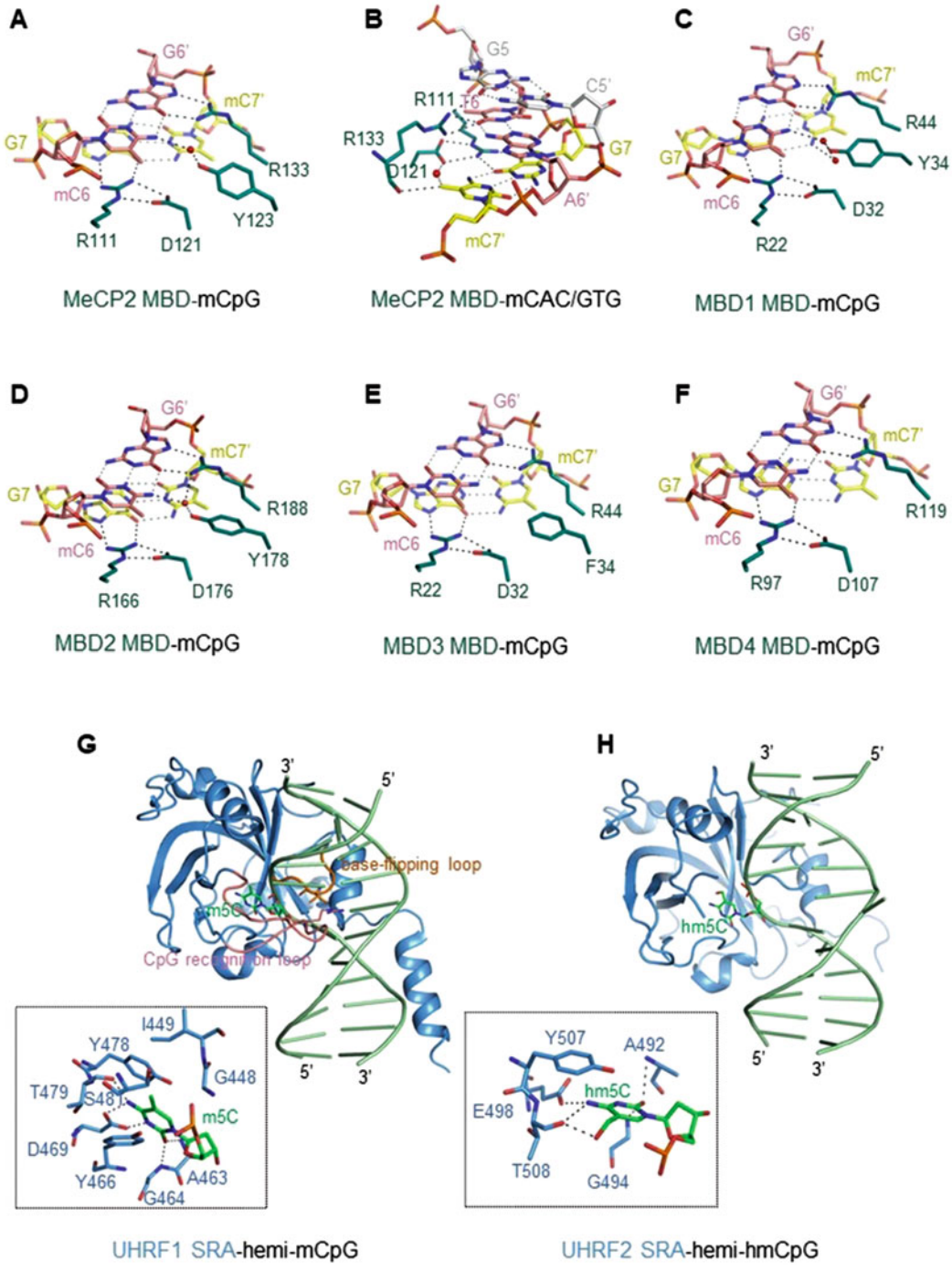

Fig. 11.2 Structures of human MBDs and SRA domains bound to methylated DNA. (**a**) and (**b**) Structures of MeCP2-MBD bound to mCpG (PDB: 6C1Y) and mCAC DNA (PDB: 6OGK), respectively. (**c**–**f**) Structures of MBDs from MBD1-4 in complex, respectively, with mCpG DNA (PDB: 6D1T, 6CNQ, 6CC8, and 4LG7). The protein residues are shown as green sticks, while nucleotides involved in base interactions are shown in sticks and colored in gray (G5-C5′), red (mC6-G6′ or T6-A6′), and yellow (G7-mC7′ or G7-C7′). The hydrogen bonds formed between residues and DNA bases are marked as black dashed lines; the DNA base pair interactions are shown as gray dashed lines. (**g**) Structure of UHRF1 SRA domain bound to the flipping-out 5mC of

2015; Kinde et al. 2016; Skene et al. 2010), which also complicates the connection between the MeCP2 binding and specific genes repression. In addition, studies in human neuronal SH-SY5Y cells (Yasui et al. 2007) or mouse hypothalamus (Chahrour et al. 2008) revealed that most genes bound by MeCP2 at their sparsely methylated promoters are actively expressed. In addition to the N-terminal MBD, the transcriptional repression domain of MeCP2 interacts with DNMT3A and inhibits its activity in vitro (Rajavelu et al. 2018). Moreover, recent structural revelation of the MeCP2 MBD bound to methylated and unmethylated CAC-containing DNA in a similar binding pattern might imply that MeCP2 also functions as a transcription activator and its binding to GTG DNA could provide disease implications in RTT syndrome (Lei et al. 2019; Liu et al. 2018).

11.2.2 MBD1

MBD1, like other MBD proteins, was initially discovered in homology screens using the MeCP2 MBD as a template. MBD1 is distinctive among the MBD proteins in that, in addition to the MBD, the protein has either two or three CXXC zinc finger domains resulting from alternative splicing (Fujita et al. 1999). Like MeCP2, MBD1 also contains a transcriptional repression domain (TRD) near its C-terminus (Fig. 11.1). The structures of MBD1 MBD in complex with methylated DNA revealed a conserved mCpG binding mode with that of MeCP2 (Fig. 11.2c) (Liu et al. 2018; Ohki et al. 2001). Although biochemical analysis previously indicated that nucleotides flanking the methylated CpG dinucleotide exert an influence on binding affinity of the MBD1 MBD (Clouaire et al. 2010), the complex structure of the MBD1 MBD with mCpG DNA does not support DNA base selectivity outside the mCpG dinucleotide (Liu et al. 2018).

In addition to the MBD, the CXXC domains of MBD1 add an additional DNA-binding interface. The CXXC domain selectively recognizes unmethylated DNA sites (Lee and Skalnik 2005; Liu and Min 2019; Xu et al. 2011, 2018). MBD1 isoform with all three CXXC domains can repress genes regardless of their promoter methylation status, while the MBD1 lacking the third CXXC domain (CXXC3) can only suppress gene expression when the promoter is methylated, suggesting that CXXC3 is essential for binding to unmethylated templates (Fujita et al. 2000; Jorgensen et al. 2004). Consistently, structure analysis and binding assays revealed that the CXXC3 domain of MBD1 specifically recognizes unmethylated CpG sites and cytosine methylation abrogates this binding; however, the first two CXXC domains of MBD1 (CXXC1 and CXXC2) do not bind CpG DNA (Jorgensen et al. 2004; Xu et al. 2018). This finding also explains why the increase of heterochromatin localization of TET1 and oxidation of 5mC require the CXXC3 domain-containing MBD1 (Zhang et al. 2017). In ESCs, biotin-tagged MBD1 is enriched at highly methylated regions, and this enrichment was lost after the depletion of DNA methylation. In addition, targeting of MBD1 to unmethylated DNA was observed only when the MBD was deleted, suggesting that the recruiting mechanism of MBD1 is dominated by the MBD–methyl-CpG interaction (Baubec et al. 2013).

Fig. 11.2 (continued) the hemi-methylated DNA (PDB: 3CLZ). The two DNA binding loops are colored in pink and orange, respectively. The interaction details of 5mC recognized by the unique binding pocket of UHRF1 are shown in the left illustration. (**h**) Structure of UHRF2 SRA domain in complex with flipping-out 5-hydroxymethylcytosine of DNA (PDB: 4PW6). The interaction details of 5hmC recognized by the unique binding pocket of UHRF2 are shown in the left illustration. The protein residues and modified cytosines are shown as sticks and colored in blue and green, respectively. The black dashed lines represent the hydrogen bonds between protein residues and base pairs

MBD1 is involved in neurodevelopment. MBD1 maintains the multipotency of neural stem cells (NSCs) by repressing neural cell differentiation-related genes (Jobe et al. 2017). Mutations or polymorphisms of MBD1 are associated with autism spectrum disorder (ASD), and also result in the accumulation of undifferentiated NSCs, impaired neurogenesis, and learning deficits in mice (Cukier et al. 2010; Jobe et al. 2017; Li et al. 2005).

11.2.3 MBD2

MBD2 is a member of the chromatin remodeling complex, nucleosome remodeling deacetylase (NuRD), which functions as a repressor connecting DNA methylation with histone deacetylation (Feng and Zhang 2001). Like MeCP2 and MBD1, MBD2 contains a TRD in addition to an MBD. MBD2 also has a glycine-arginine repeat and a coiled-coil domain (Fig. 11.1), which is essential for binding to the Mi-2/NuRD complex (Gnanapragasam et al. 2011). MBD2 was also reported to be an integral component of the MeCP1 complex, which was subsequently shown to have biochemical similarities to NuRD (Ng et al. 1999; Zhang et al. 1999). Single-molecule fluorescence techniques revealed that MBD2 spreads more quickly in the regions of CpG-rich sequences than that of CpG-free DNA, while MBD2 binding is static or with a slow exchange when interacting with DNA regions enriched for mCpG. It was thought that MBD2 facilitates the rapid movement and nucleosome remodeling of the NuRD complex in the CpG-rich regions, and this movement is limited when those loci are methylated, which further contributes to gene silencing (Pan et al. 2017).

The interaction of the MBD2 MBD with mCpG DNA was first defined for chicken MBD2 (which is >95% identical to human MBD2) by Williams and colleagues (Scarsdale et al. 2011). Subsequently, two human MBD2–mCpG complex structures were determined (Liu et al. 2018). Similar to MeCP2 and MBD1, base-specific contacts with the methylated CpG palindrome are mediated by a pair of arginine residues (R166 and R188), as well as a tyrosine residue (Y176) in MBD2 (Fig. 11.2d). Although MeCP2 and MBD2 MBDs exhibit a conserved binding mode, knock-in mice expressing a chimeric protein (MM2) by swapping the MBDs of MeCP2 and MBD2 exhibit severe phenotypic features that are largely similar to those seen in mouse models of RTT, suggesting that the conserved MBD of MBD2 could not functionally replace that of MeCP2 (Tillotson et al. 2021).

MBD2 is expressed in most somatic cells and is particularly abundant in ESCs in mice (Hendrich and Bird 1998). Despite this intriguing expression pattern, MBD2 knockout mice are viable and fertile (Hendrich et al. 2001). MBD2 has two predominant isoforms, MBD2a and MBD2c; MBD2c lacks the carboxyl-terminal region including the coiled-coil domain integral to the interaction with the NuRD complex (Hendrich and Bird 1998). Differential expression of MBD2a and MBD2c was shown in human pluripotent stem cells (hPSCs); MBD2c is dominant in hPSCs, while MBD2a is dominant in fibroblasts (Lu et al. 2014). Interestingly, MBD2a, but not MBD2c, can interact with the NuRD complex and promote differentiation, while MBD2c enhances reprogramming efficiency when overexpressed in fibroblasts.

11.2.4 MBD3

As predicted from its high sequence similarity with MBD2 (Fig. 11.1), MBD3 is also a member of the NuRD complex (Le Guezennec et al. 2006). MBD3 from mammals has been reported to lack the capacity for high-affinity interaction with methylated DNA in conventional biochemical assays (Fraga et al. 2003; Hendrich and Bird 1998). In contrast, the amphibian protein displays a strong preference for methylated substrates (Wade et al. 1999). Close inspection of the amino acid sequence of MBD3 MBDs from multiple species reveals that mammals differ from Amphibia, fish, reptiles, and birds at a critical position encoding the conserved tyrosine residue involved in specific contacts with the methylated

cytosine (changed to phenylalanine in mammals). Recently, crystal structures of human MBD3 MBD bound to mCpG DNA were determined (Liu et al. 2019). Similar to MBD2, the MBD3 MBD binds to mCpG via two conserved arginine fingers (Fig. 11.2e). By structural comparison to that of MBD2, the replacement of tyrosine by phenylalanine at F34 of MBD3 results in weaker mCpG DNA binding compared to MBD2, due to the loss of a solvent-mediated interaction of the hydroxyl group in tyrosine with the N4-amino group of the methylcytosine (Liu et al. 2019). The complex structure of the MBD3 MBD bound to a non-palindromic DNA also revealed that the MBDs recognize the mCpG DNA without orientation preference (Liu et al. 2019), consistent with the observations that the sequences flanking the mCpG dinucleotide do not influence the mCpG DNA binding significantly. Therefore, MBD3 is able to bind to methylated CpG DNA, albeit with a reduced binding affinity (Gunther et al. 2013; Liu et al. 2019).

MBD2 and MBD3 form mutually exclusive complexes (Le Guezennec et al. 2006). In contrast to the mild phenotypes of the MBD2 knockout mice, MBD3 deletion causes early embryonic lethality. Furthermore, MBD3-null ES cells can maintain stemness, even in the absence of leukemia inhibitory factor (LIF) (Hendrich et al. 2001; Kaji et al. 2006). These striking phenotypes suggest that MBD2 and MBD3 have nonredundant roles. Importantly, deletion of MBD3 can significantly enhance reprogramming efficiency, suggesting that MBD3 functions as a barrier to reprogramming (Luo et al. 2013; Rais et al. 2013). However, the field is not in complete agreement with the role of MBD3 in reprogramming, suggesting a possibility that the role of MBD3 is highly context-dependent (Dos Santos et al. 2014). Moreover, in uterine serous carcinoma patients, a small segment of chromosome 19 containing MBD3 is frequently deleted, suggesting a critical role of MBD3 in tumorigenesis or tumor progression (Zhao et al. 2013). Despite the accumulation of evidence on the critical function of MBD3 in biology, the detailed molecular mechanism of how MBD3, or MBD3/NuRD, regulates gene expression and chromatin structure is still unclear. Most cells express a splice variant of MBD3 that disrupts the canonical MBD sequence (Hendrich and Tweedie 2003). Genomic localization analyses revealed that MBD3 preferentially localizes at unmethylated CpG-rich regions, including CpG islands, while MBD2 distributes across the genome in a methylation-dependent manner (Baubec et al. 2013).

11.2.5 MBD4

MBD4 contains two functional domains, an N-terminal MBD and a C-terminal glycosylase domain (GD), separated by a long link of unknown function (Fig. 11.1). The presence of a glycosylase domain makes MBD4 a unique member of the MBD family. The MBD4 MBD has high affinity for methylated CpG-containing substrates. It has a similar affinity for the deamination product of that substrate, i.e., methylated CpG base-paired (mismatch) with TpG (Hendrich et al. 1999; Otani et al. 2013). Not surprisingly, MBD4 possesses enzymatic activity that can repair mCpG/TpG or mCpG/hmUpG double-stranded mismatches generated by spontaneous deamination of 5mC (Hendrich and Bird 1998; Hendrich et al. 1999). MBD4 knockout mice are viable and fertile with minor phenotypes, including a slight increase in C to T mutations at CpG sites (Millar et al. 2002; Wong et al. 2002). Although the deletion of MBD4 itself does not impact tumorigenesis, it increases tumor incidence in a susceptible genetic background (mutation in the adenomatous polyposis coli (APC) gene). In addition, mutations in MBD4 have been observed in human colorectal tumors with microsatellite instability (Riccio et al. 1999). Taken together, these observations suggest that MBD4 plays an important role in tumor progression by regulating DNA mismatch repair.

Similar to MeCP2 and other MBD family members, the crystal structures of the MBD4 MBD with different DNA sequences revealed that MBD4 recognizes mCpG or mismatched DNA by the 5mC (or T)-Arg-Gua triad interactions (Fig. 11.2f) (Liu et al. 2018; Otani

et al. 2013; Wu et al. 2021). While there are some minor alterations, the overall structure of the MBD4 MBD bound to a mismatched DNA is highly similar to that of the MBD4 MBD bound to symmetric methylated CpG DNA. In contrast to the MBD, the glycosylase domain of MBD4 binds to DNA containing a G:X mismatch in a very different manner (Hashimoto et al. 2012). The target nucleotide is flipped out from the DNA strand and an arginine residue from the MBD fills that space. The flipped base is associated with the active-site cleft. Importantly, the crystal structure of the full-length MBD4 containing both the MBD and GD is yet to be solved, so it is unclear to what level the two domains communicate. This question has been partially approached via analysis of solution structures; although MBD4 shows a slow exchange rate between different DNA molecules (intermolecular exchange), it has rapid exchange rate between the two binding sites on the same dsDNA molecule (intramolecular exchange) (Walavalkar et al. 2014). These data support a local hopping model in which the MBD of MBD4 rapidly scans multiple methylated CpG sites and supports the mismatch repair conducted by the GD (Walavalkar et al. 2014).

11.2.6 MBD5 and MBD6

MBD5 and MBD6 are characterized to be associated with neurodevelopmental disorders (Cukier et al. 2012; Talkowski et al. 2012). MBD5-null mice develop growth defects and preweaning lethality, exhibiting several phenotypic features seen in patients with 2q23.1 microdeletion (Du et al. 2012). Like other MBD proteins, MBD5 and MBD6 contain an MBD that is required for their heterochromatin localization (Laget et al. 2010). However, electrophoretic mobility shift assay (EMSA) experiments using purified MBDs from MBD5 and MBD6 indicated that these domains display no methylcytosine-binding capacity. This property may be explained by the loss of a characteristic loop structure and the conserved arginine fingers that contribute to DNA base-specific binding, which is critical for methylated mCpG DNA contacts in the MBDs (Hendrich and Tweedie 2003; Liu et al. 2018). Interestingly, although the incomplete MBDs of these proteins have lost the methylcytosine-binding affinity, they can interact with mammalian PR-DUB polycomb protein complex, which is known as a histone H2A deubiquitinase (Baymaz et al. 2014). These distinct differences from other MBD proteins may assign a specialized function to MBD5 and MBD6.

11.3 SET- and RING-Associated (SRA) Domain

11.3.1 UHRF1

UHRF1 (ubiquitin-like, containing PHD and RING finger domains 1, also known as ICBP90 or Np95) contains a ubiquitin-like domain (UBL), tandem Tudor domain (TTD), plant homeodomain (PHD), SET- and RING-associated (SRA) domain, and really interesting new gene (RING) domain, which interdependently coordinate epigenetic functions of UHRF1 (Fig. 11.1). UHRF1 was originally identified as a potential regulator of topoisomerase IIα (Hopfner et al. 2000). UHRF1 is essential for the maintenance of proper DNA methylation levels by recruiting DNA methyltransferase 1 (DNMT1) to replication foci (Fang et al. 2016; Li et al. 2018). Deletion of UHRF1 in mice causes genome-wide DNA hypomethylation and results in embryonic lethality, presumably due to the dysfunction of DNMT1. UHRF1 was also observed to contribute to DNA damage repair by binding directly to interstrand crosslink (ICL)-containing DNA and facilitating the recruitment of other DNA repair factors (Liang et al. 2015; Mancini et al. 2021; Tian et al. 2015).

UHRF1 recognizes hemi-methylated DNA with its SRA domain employing the same base-flipping mechanism, which is commonly found in DNA methyltransferases (Arita et al. 2008; Avvakumov et al. 2008; Hashimoto et al. 2008; Song et al. 2012). The SRA domain uses the CpG recognition loop and the base-flipping loop approach the major groove and minor groove of

DNA, respectively. The flipped-out methylcytosine of the duplex DNA is stabilized in a binding pocket with van der Waals interactions, planar stacking contacts, and Watson–Crick polar hydrogen bonds (Fig. 11.2g). The UHRF1 SRA-DNA complex structure showed that cytosine methylation from the complementary strand interferes with the conformation of a conserved asparagine, e.g., human N489, which explains why UHRF1 SRA prefers binding to hemi-methylated DNA rather than symmetric mCpG DNA (Arita et al. 2008; Avvakumov et al. 2008; Hashimoto et al. 2008). The use of this base-flipping mechanism uniquely positions UHRF proteins in the MBP family; the SRA domain was the first domain that conducts base flipping without enzymatic activity (Song et al. 2012).

In addition to hemi-methylated DNA, UHRF1 also recognizes histone modifications, such as H3K9me3, unmodified H3K9, and H3R2 through its histone reader domains (Hu et al. 2011; Nady et al. 2011; Rajakumara et al. 2011). This interaction is allosterically regulated by phosphatidylinositol 5-phospate (PI5P), which alters the local structure around the Tudor and PHD domains (Gelato et al. 2014). Interestingly, it has been demonstrated that the interaction between UHRF1 and H3K9me3 is essential for the maintenance of DNA methylation (Rothbart et al. 2012). Furthermore, UHRF1 binding to hemi-methylated CpGs activates the ubiquitylation activity of the UHRF1 RING domain toward H3K18 and/or K23 adjacent to the H3 binding site of UHRF1, suggesting a role for UHRF1 as a molecular hub connecting DNA methylation and histone modifications (Harrison et al. 2016).

11.3.2 UHRF2

UHRF2 (also known as NIRF) has the same domain structure and high sequence similarity as its paralog UHRF1 (Fig. 11.1). They are the only two proteins with an SRA domain in humans (Mori et al. 2002). Based on their sequence similarity, UHRF2 and UHRF1 appear to share the same functions; UHRF2 also recognizes hemi-methylated DNA and interacts with DNMT1 (Zhang et al. 2011). However, there are critical differences between UHRF1 and UHRF2. Most importantly, unlike UHRF1, the SRA domain of UHRF2 creates a larger pocket to specifically bind to 5hmC by base flipping (Fig. 11.2h) (Zhou et al. 2014). In addition, UHRF2 and UHRF1 are differentially expressed; UHRF2 is downregulated in ESCs and gradually upregulated upon differentiation, whereas UHRF1 shows an opposite pattern (Pichler et al. 2011). Moreover, the introduction of UHRF2 into UHRF1-null ESCs cannot rescue the hypomethylation phenotype, suggesting a differential functionality of UHRF2, at least in ESCs (Zhang et al. 2011). The 5hmC binding can stimulate the E3 ligase activity of the UHRF2 RING domain to regulate the K33-linked polyubiquitination of the BER component XRCC1 (Liu et al. 2021). However, unlike UHRF1, UHRF2 could not ubiquitinate histones in the context of nucleosomes, implying that fragments outside the chromatin binding region also contribute to UHRF1 location in a productive conformation for nucleosomal histone ubiquitination (Vaughan et al. 2018). With these similarities and differences, it is yet unclear how UHRF1 and UHRF2 cooperatively (or distinctively) function in cells; analysis with UHRF2 knockout mice would provide more information. UHRF2 knockout mice exhibit frequent spontaneous seizures and abnormal electrical activities during adulthood. In addition, UHRF2 knockout mice only display a decreased 5mC level at certain genomic loci in brains. Therefore, UHRF2 might play a unique role differing from that of its paralog UHRF1 in the maintenance of 5mC levels (Liu et al. 2017).

11.4 Transcription Factors

In addition to the canonical MBD and SRA domains, an increasing number of transcriptional regulators have been identified to instruct downstream events depending on the recognition of different cytosine modification states, including

the members of C2H2 zinc finger (ZF) proteins, basic helix-loop-helix (bHLH), basic leucine-zipper (bZIP), and homeodomain transcriptional factor families.

11.4.1 Kaiso and ZBTB38

The C2H2 zinc finger is one of the most abundant DNA binding motifs. Each C2H2 zinc finger contains a ββα-fold core comprised of two β-strands packing against an α-helix, and this core is stabilized by a tetrahedrally coordinated Zn^{2+} with two cysteines and two histidine residues (Klug 2010). Being a member of the BTB/POZ (broad complex, Tramtrack, and bric-a-brac/poxvirus and zinc finger) family, Kaiso (also known as ZBTB33) contains three C2H2 zinc fingers and was originally identified as a binding partner of p120 catenin (Daniel and Reynolds 1999). In addition to p120 catenin, Kaiso also interacts with a repression complex, N-CoR (nuclear receptor corepressor), and suppresses the expression of MTA2, a member of the NuRD complex, in a methylation-dependent manner (Yoon et al. 2003). As the N-CoR complex contains histone deacetylases, the recruitment of the N-CoR complex mediated by Kaiso is proposed as a potential mechanism of DNA methylation-dependent gene repression. However, Kaiso is also known to associate with p53 and upregulate apoptosis-related genes, suggesting pleiotropic roles of Kaiso in different biological contexts (Koh et al. 2015; Koh et al. 2014). Moreover, Kasio showed both pro- and antitumorigenic activities, which also implies that Kaiso is a context-dependent regulator (Koh et al. 2014; Prokhortchouk et al. 2006; Soubry et al. 2010).

Kaiso preferentially binds to two consecutively methylated CpG dinucleotides or to a chemically similar, albeit unmethylated, TpG-containing sites, TCCTGCCA (also called the Kaiso binding sequence, KBS) (Daniel et al. 2002; Prokhortchouk et al. 2001). The crystal structures of Kaiso have been solved in complex with two different DNA templates: a methylated template, MeECad (promoter region of E-cadherin) containing two methylated CpG dinucleotides, and an unmethylated sequence TCCTGCCA (Buck-Koehntop et al. 2012; Daniel et al. 2002). The two structures are almost identical; Kaiso recognizes the methyl group, either of mCpG or TpG dinucleotides, using the 5mC-Arg-Gua triad structure (Fig. 11.3a) (Daniel et al. 2002). The first two zinc fingers hold the major groove of DNA, and the third zinc finger (together with the C-terminal extension) enables high-affinity binding (Fig. 11.3a) (Buck-Koehntop et al. 2012). A recent study suggests that E535 of Kaiso adopts different conformations to determine the distinct recognition of methylated and KBS motifs by Kaiso (Fig. 11.3a) (Nikolova et al. 2020). Recently, Kaiso was found to immunoprecipitate with the de novo DNA methyltransferases DNMT3A/3B, suggesting that Kaiso may recruit the DNA methyltransferases to modulate genome methylation apart from being a methyl-DNA-binding protein (Kaplun et al. 2021).

In addition to Kasio, ZBTB4 and ZBTB38 were also found to bind methylated DNA in vitro and in vivo. ZBTB4 and ZBTB38 specifically bind to the methylated allele of imprinting gene H19/Igf2 and become delocalized with loss of DNA methylation (Filion et al. 2006). ZBTB38 is involved in cellular proliferation, apoptosis, and genomic stability through modulating transcriptional activity (Miotto et al. 2014; Nishii et al. 2012; Oikawa et al. 2008; Pozner et al. 2018). ZBTB38 specifically binds to a DNA sequence ((A/G)TmCG(G/A)(mC/T)(G/A)) through its C-terminal ZF6-10 (Pozner et al. 2018). The crystal structure of ZBTB38 in complex with a DNA sequence of ATmCGGmCG revealed that ZF7 and ZF8 contribute to base-specific DNA interactions using a 5mC-Arg-Gua triad, while ZF6 and ZF9 mainly stabilize ZF7 and ZF8 to form base-specific interactions (Hudson et al. 2018).

11.4.2 CTCF

The multidomain CCCTC-binding factor (CTCF) is crucial for chromatin architecture organization

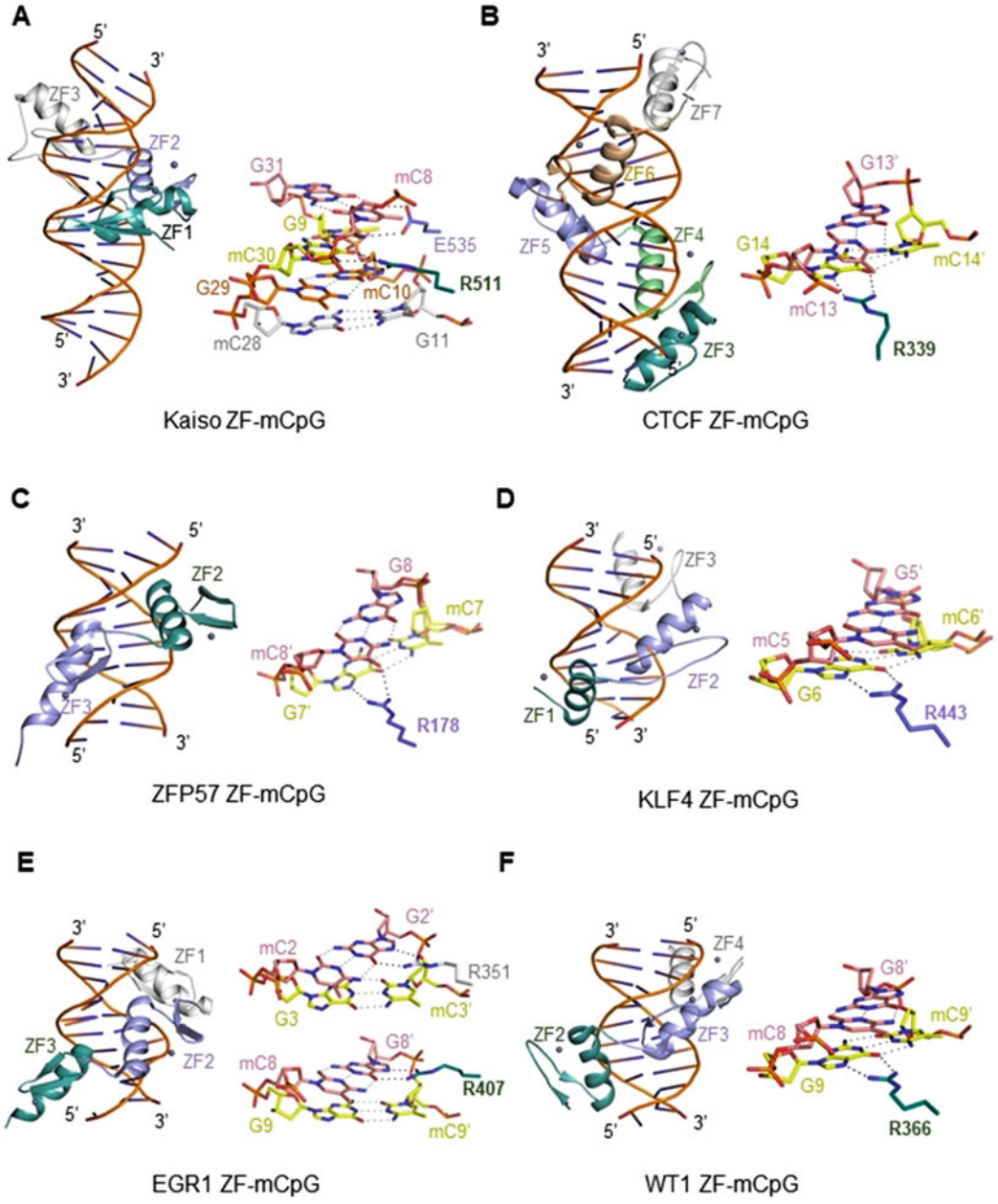

Fig. 11.3 Structural basis of C2H2 ZFs binding to methylated CpG DNA. (**a**) Structure of Kaiso ZF1-3 in complex with DNA containing two consecutively methylated CpG sites (PDB: 4F6N). (**b**) Structure of CTCF ZF3-7 in complex with methylated CpG DNA (PDB: 5T00). (**c**) Structure of ZFP57 ZF2-3 in complex with methylated CpG DNA (PDB: 4GZN). (**d**) Structure of KLF4 ZF1-3 in complex with methylated CpG DNA (PDB: 4M9E). (**e**) Structure of EGR1 ZF1-3 in complex with methylated CpG DNA (PDB: 4X9J). (**f**) Structure of WT1 ZF2-4 in complex with methylated CpG DNA (PDB: 4R2E). The zinc ions are shown as gray balls. The hydrogen bonds formed between protein residues and DNA bases are marked as black dashed lines, while the DNA base pair interactions are shown as gray dashed lines

and gene expression regulation. CTCF contains 11 tandem C2H2 zinc fingers and binds to a 15-base pair consensus sequence NCANNAG (G/A)NGGC(G/A)(C/G)(T/C) (N=A, C, G or T) that can be methylated on cytosines at positions 2 and 12 (C2 and C12) (Nakahashi et al. 2013; Rhee and Pugh 2011; Wang et al. 2012). The crystal structures of CTCF bound to unmethylated and methylated CpG DNA have been determined (Hashimoto et al. 2017), which showed that ZF3-7 recognize the major groove of DNA, and each ZF interacts with three adjacent DNA base pairs (also called the "triplet" element) (Fig. 11.3b) (Choo and Klug 1997). ZF8 and ZF9 span along the DNA phosphate backbone of the 15-base pair core sequence, and no contact is found for ZF1 and ZF10-11. In the complex structure with the methylated CpG DNA, ZF3 and ZF4 contribute to mCpG binding at position 12 using a 5mC-Arg-G triad (Fig. 11.3b). In contrast, methylation at position 2 cytosine forms a steric clash with D451 of CTCF (Hashimoto et al. 2017), explaining why methylation at C2 significantly abolishes DNA binding, whereas methylation at C12 increases DNA binding of CTCF. Thus, CTCF binds to methylated DNA and recognizes 5mC in a position-dependent manner (Hashimoto et al. 2017).

11.4.3 ZFP57

ZFP57, as a member of the KRAB-ZFP (Kruppel-associated box zinc finger) family, is a maternal-zygotic effect gene for maintaining DNA methylation memory in early mouse embryos and embryonic stem cells. ZFP57, together with its binding cofactor KAP1, binds to methylated hexanucleotides within CpG-rich sequences located in imprinting control regions (ICRs) and regulates imprinted genes (Li et al. 2008; Quenneville et al. 2011). In addition, ZFP57 and KAP1 coimmunoprecipitate with UHRF1 and DNA methyltransferases DNMT1, DNMT 3A, and DNMT 3B to maintain the DNA methylation at ICRs (Quenneville et al. 2011; Zuo et al. 2012). Loss-of-function mutations in human ZFP57 are associated with a global imprinting disorder (ID) and transient neonatal diabetes (TND) (Mackay et al. 2008).

ZFP57 preferentially binds to a methylated hexanucleotide sequence (TGCmCGC) using two classical C2H2 domains (Liu et al. 2012; Quenneville et al. 2011). Like MBD proteins and Kaiso, ZFP57 uses a 5mC-Arg-Gua triad to recognize the methylated cytosine (Fig. 11.3c). In addition to TGCmCGC, the ZFP57 DNA binding domain interacts with the oligonucleotide containing the sequence GGCmCGC in vivo and in vitro, albeit weaker than that of TGCmCGC sequence (Anvar et al. 2016).

11.4.4 KLF4

KLF4 is a member of the Kruppel-like protein family and contains three standard Kruppel-like zinc fingers. KLF4 is well known as one of the Yamanaka factors for reprogramming somatic cells to induce pluripotency (Takahashi and Yamanaka 2006). KLF4 recruits the histone H3K27me3 demethylase JMJD3 to reduce H3K27me3 levels at both enhancers and promoters of epithelial and pluripotency genes, facilitating reprogramming somatic cells to pluripotency (Huang et al. 2020). In addition, KLF4 also promotes somatic cell reprogramming by cooperating with transcription factors OCT4 and SOX2 (Chronis et al. 2017; Takahashi and Yamanaka 2006). OCT4 acts as a pioneer factor that opens heterochromatin and facilitates the binding of KLF4 (Chen et al. 2020). In addition to reprogramming, KLF4 is required for normal skin and colon development, and KLF4 knockout mice die soon after birth (Katz et al. 2002; Segre et al. 1999). The transcriptional activities of KLF4 have also been implicated in regulating genomic stability (El-Karim et al. 2013), cellular proliferation (Chen et al. 2003), DNA damage response, and apoptosis (Yoon et al. 2005).

KLF4 exhibits comparable binding activity to CpG, mCpG, or TpG containing DNA in vitro (Hashimoto et al. 2016; Liu et al. 2014). Interestingly, genome-wide studies showed that about half of the KLF4-binding sites in vivo are highly methylated (Hu et al. 2013). KLF4 binds to a

consensus sequence of GG(T/C)G with a preference for the methylated status (Sharma et al. 2021; Wan et al. 2017). Structural analysis revealed that the tandem C2H2 zinc fingers of KLF4 recognize the major groove of the DNA using a 5mC-Arg-Gua triad (Fig. 11.3d) (Liu et al. 2014; Schuetz et al. 2011). KLF4 progressively loses binding affinity as 5mC is oxidized into 5hmC, 5fC, and 5caC (Liu et al. 2014).

11.4.5 EGR1 and WT1

EGR1 (early growth response protein 1), also called ZIF-268, NGFI-A, KROX 24, or ZENK, is a member of the EGR protein family. EGR1 is involved in several biological processes, including cell proliferation, differentiation, inflammation, and apoptosis (Beckmann and Wilce 1997; Bozon et al. 2003; Duclot and Kabbaj 2017; Lee et al. 2004; Sanchez-Guerrero et al. 2013; Veyrac et al. 2014). EGR1 specifically recognizes and binds target genes using three C2H2 zinc fingers, which either promotes or inhibits the expression of target genes (Kim et al. 2011). The WT1 (Wilms' tumor 1) is a predisposition gene for Wilms' tumor, a pediatric kidney cancer (Call et al. 1990; Charlton and Pritchard-Jones 2016; Gessler et al. 1990). WT1 is involved in the regulation of BMP/pSMAD and FGF pathways in the early kidney anlagen (Motamedi et al. 2014), and also plays critical roles in tissue homeostasis, development, and disease (Hastie 2017). Although there are more than 35 potential mammalian WT1 isoforms generated by splicing or alternative translation start sites, all isoforms contain four C2H2 zinc fingers (Hastie 2017). Like EGR1, WT1 can either repress or activate specific target genes depending on its binding partners (Hastie 2017). EGR1 and WT1 have highly divergent cellular functions, but both TFs recognize the same consensus sequence, GCG (T/G)GGGCG (Hartwig et al. 2010; Pavletich and Pabo 1991; Rauscher et al. 1990; Stoll et al. 2007; Zandarashvili et al. 2015). Consistent with other MBPs, these two proteins interact with a methylated DNA template using a 5mC-Arg-Gua triad (Fig. 11.3e, f) (Hashimoto et al. 2014).

11.4.6 bZIP

Activator protein 1 (AP-1) is a transcription factor family with basic leucine zipper (bZIP), including c-Fos, c-Jun, CREB, C/EBP, and ATF (activating transcription factor), and is involved in multiple biological processes, such as development, metabolism, cell proliferation, and apoptosis (Angel and Karin 1991; Eferl and Wagner 2003; Hess et al. 2004; Karin et al. 1997). The AP-1 transcription factors usually form either heterodimers or homodimers depending on the sequences of their leucine zipper motifs (Miller 2009), and dimerization is considered a prerequisite for DNA binding (Landschulz et al. 1989). c-Fos and c-Jun form homo- or heterodimers to recognize three types of 7-bp 12-*O*-tetradecanoylphorbol-13-acetate (TPA)-response elements: TGAGTCA, mCGAGTCA, and TGAGmCCA (Bhende et al. 2004; Eferl and Wagner 2003; Gustems et al. 2014; Tulchinsky et al. 1996). C-Jun forms a clamp-like structure through its C-terminal leucine zipper, while its N-terminal basic region interacts with the major groove of the target DNA sequence (Fig. 11.4a). In contrast to the other MBPs, which use a 5mC-Arg-G triad to recognize mC or T, AP-1 binds to mCpG and TpG through van der Waals contacts between a conserved di-alanine (Ala265 and Ala266) and the methyl groups of 5mC or thymine (Fig. 11.4b) (Hong et al. 2017).

Another bZIP TF member, CCAAT/enhancer binding protein β (C/EBPβ), is associated with cytokine-mediated macrophage activation and rapid granulopoiesis (Hirai et al. 2006), and regulates gene expression at different developmental stages (Sun et al. 2017; Tsukada et al. 2011). C/EBPβ recognizes a DNA sequence TTGmCGCAA. Instead of the di-alanine, C/EBPβ bZIP has a unique Ala-Val dipeptide (A284-V285) that contributes to van der Waals interactions with the 5-position methyl group of methylated CpG and thymine, and the methylated CpG dinucleotide forms a classic 5mC-Arg-Gua triad with R289 (Yang et al. 2019). Oxidation of 5mC to 5hmC introduces a hydroxyl group and

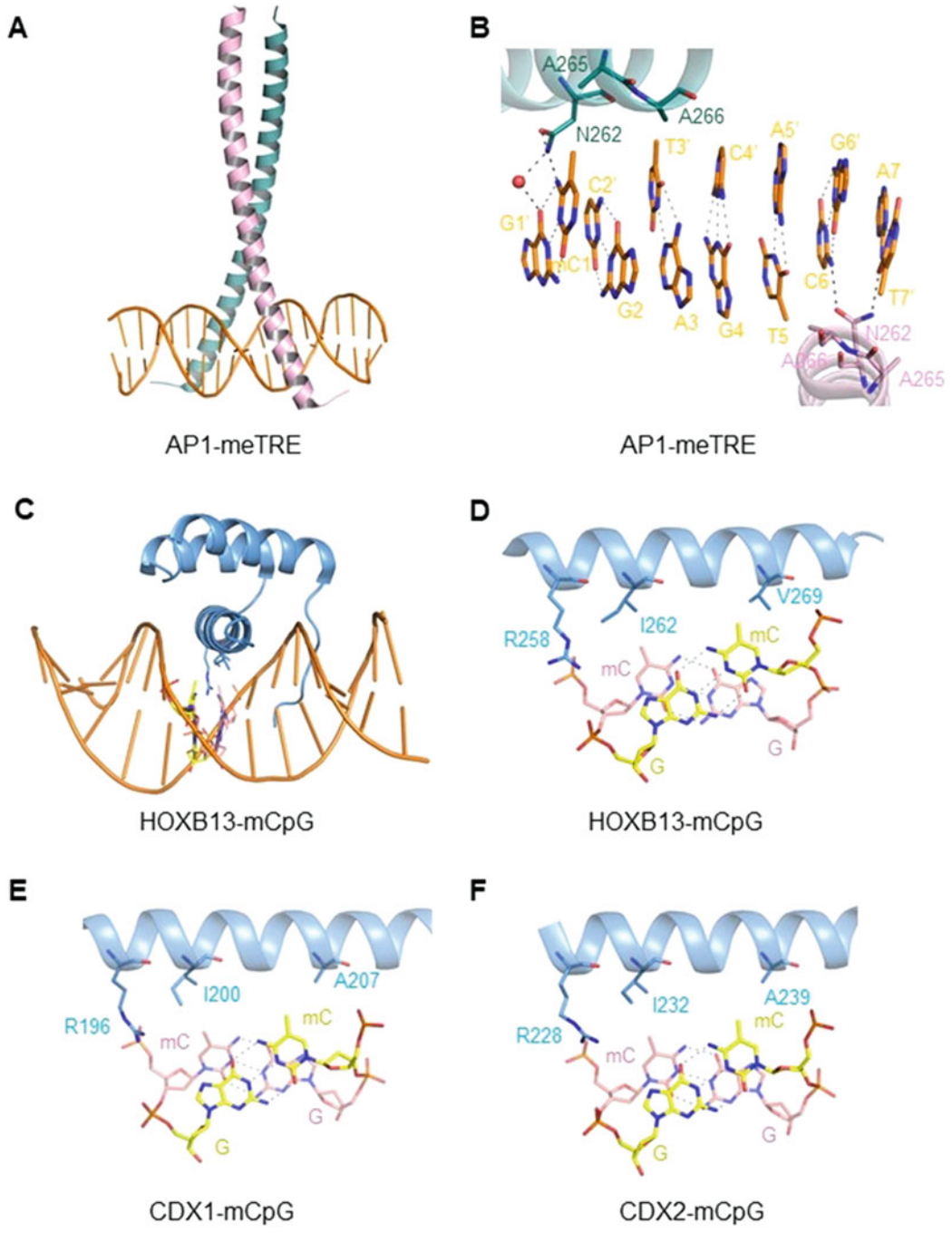

Fig. 11.4 Structures of other TFs binding to methylated DNA. (**a**) AP-1 bZIP domain forms dimer for binding to methylated cytosine DNA (PDB: 5T01). (**b**) AP-1 bZIP protein recognizes the DNA by the van der Waals contacts with the methyl groups of methylated cytosine and thymine. The protein residues are shown as green and pink sticks, respectively, while nucleotides are shown as sticks and colored orange. (**c**) Overall structure of HOXB13 homeodomain in complex with mCpG DNA (PDB: 5EF6). (**d**) HOXB13 homeodomain recognizes the DNA by the van der Waals contacts with the methyl groups of methylated cytosine. The protein residues are shown as blue sticks, while nucleotides are shown in sticks and colored in red and yellow, respectively. (**e**, **f**) CDX1 (PDB: 5LUX) and CDX2 (PDB: 5LTY) homeodomains recognize the DNA by the van der Waals contacts with the methyl groups of methylated cytosine, respectively. The protein residues and nucleotides are shown in the same way as in **d**. The hydrogen bonds formed between protein residues and DNA are marked as black dashed lines, while the DNA base pair interactions are shown as gray dashed lines

reduces its interaction with C/EBPβ (Yang et al. 2019).

11.4.7 Homeodomain Proteins

Homeodomain transcription factors bind to DNA and regulate the expression of morphogenesis-related target genes in eukaryotes (Burglin and Affolter 2016). Some homeodomain transcription factors have been reported to preferentially bind to methylated DNA (Yin et al. 2017). To illustrate the molecular basis for the methylated DNA binding preference, the structures of several homeodomain proteins, such as HOXB13, CDX1, and CDX2, bound to their cognate DNA sequences have been reported (Fig. 11.4c–f) (Yin et al. 2017). It was found that the homeodomain recognizes mCpG by direct hydrophobic interactions with the methyl groups of both 5mCs from the CpG dinucleotide. For example, the complex structure of HOXB13 bound to the DNA sequence CTmCGTAAA showed that its DNA binding domain consists of three α-helices, and the C-terminal α-helix (α3) lies in the major groove of the DNA and the N-terminal tail interacts with the minor groove (Fig. 11.4c) (Yin et al. 2017). Specifically, I262 has a direct hydrophobic interaction with the methyl group of the methylcytosine, while V269 forms another hydrophobic contact with the methyl group of methylcytosine from the complementary strand. Besides, the aliphatic chain of R258 interacts with I262 to enhance the hydrophobic environment (Fig. 11.4d) (Yin et al. 2017). This interaction is also conserved in the binding of CDX1 or CDX2 to GTmCGTAAAA (Fig. 11.4e, f) (Yin et al. 2017). In contrast to CDX1 and CDX2, homeodomain protein LHX4 displays a slightly weaker binding to TmCGTTA site and no significant binding to the unmethylated TCGTTA (Yin et al. 2017). The complex structure of LHX4 bound to the TAATA site showed that R127, V131, and A138 adopt a similar conformation to the corresponding residues in HOXB13, CDX1, and CDX2, which also engage in hydrophobic interaction with TmCGTTA (Yin et al. 2017).

11.5 Conclusion

In this chapter, we reviewed the research on proteins that recognize methylated DNA. Several themes have emerged from recent biochemical, structural, and genomic studies of MBPs. Among them, MBD and C2H2 ZFs DNA binding proteins use the same classic 5mC-Arg-Gua triad to recognize the methylated CpG and unmethylated TpG, while the bZIP proteins AP-1 and homeodomain proteins characterized in this chapter use van der Waals contacts to recognize the methyl groups of 5mC and thymine (Hong et al. 2017; Yin et al. 2017). The SRA proteins utilize specialized pockets to recognize 5mC or 5hmC. Although the POU and PAX proteins, NFAT proteins, SMAD proteins, and FOX proteins have also been reported to interact with mCpG sites (Yin et al. 2017), their methylcytosine recognition mechanism remains to be characterized. Thymine can replace mC for DNA binding due to their similar chemical structure; thus, the TpG recognition by POU and NFAT can also provide insights into their mC binding (Remenyi et al. 2001; Stroud et al. 2002).

Detailed analyses of structural and biochemical data have also indicated that the MBDs recognize DNA without sequence selectivity outside the methylated CpG sequence. However, it seems that individual MBD family members display some level of functional specificity as revealed by genomic mapping experiments in living cells, so how does a methylated CpG motif specify recruitment of a unique MBD has been puzzling, as a dinucleotide sequence like mCpG lacks the chemical information inherent in the longer binding sites typical for most transcription factors. The structural and biochemical work of some other MBPs reviewed here places this issue in an even wider context: MBD and most C2H2 ZF proteins utilize a common protein feature (arginine) to recognize chemical features of the methylated CpG dinucleotide. One possible explanation would be that additional functional domains of these MBPs recognize flanking sequence information, and the true consensus recognition sequence for these proteins is not a

simple methylated CpG, but the methylated CpG with its flanking sequence context that differs for each MBP. This feature endows the MBP proteins with considerable flexibility to respond to DNA methylation with different outputs at unique loci within the genome. Unraveling how cells utilize this surprising flexibility to resolve epigenetic regulation remains a principal challenge of current genetic and genomic experiments.

Acknowledgments This work was supported by funds from the National Natural Science Foundation of China (Number 31770834 to K.L) and the funds from Central China Normal University (Number 31101200001 to J.M. and CCNU19TS084 to K.L). This work was also supported by the Intramural Research Program of the National Institute of Environmental Health Sciences, NIH (ES101965 to P.A.W.). Due to the limited space, we apologize to those whose original work could not be cited in this chapter.

References

Amir RE, Van den Veyver IB, Wan M, Tran CQ, Francke U, Zoghbi HY (1999) Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. Nat Genet 23:185–188. https://doi.org/10.1038/13810

Angel P, Karin M (1991) The role of Jun, Fos and the AP-1 complex in cell-proliferation and transformation. Biochim Biophys Acta 1072:129–157. https://doi.org/10.1016/0304-419x(91)90011-9

Anvar Z, Cammisa M, Riso V, Baglivo I, Kukreja H, Sparago A, Girardot M, Lad S, De Feis I, Cerrato F, Angelini C, Feil R, Pedone PV, Grimaldi G, Riccio A (2016) ZFP57 recognizes multiple and closely spaced sequence motif variants to maintain repressive epigenetic marks in mouse embryonic stem cells. Nucleic Acids Res 44:1118–1132. https://doi.org/10.1093/nar/gkv1059

Arita K, Ariyoshi M, Tochio H, Nakamura Y, Shirakawa M (2008) Recognition of hemi-methylated DNA by the SRA protein UHRF1 by a base-flipping mechanism. Nature 455:818–821. https://doi.org/10.1038/nature07249

Avvakumov GV, Walker JR, Xue S, Li Y, Duan S, Bronner C, Arrowsmith CH, Dhe-Paganon S (2008) Structural basis for recognition of hemi-methylated DNA by the SRA domain of human UHRF1. Nature 455:822–825. https://doi.org/10.1038/nature07273

Baubec T, Ivanek R, Lienert F, Schubeler D (2013) Methylation-dependent and -independent genomic targeting principles of the MBD protein family. Cell 153:480–492. https://doi.org/10.1016/j.cell.2013.03.011

Baymaz HI, Fournier A, Laget S, Ji Z, Jansen PW, Smits AH, Ferry L, Mensinga A, Poser I, Sharrocks A, Defossez PA, Vermeulen M (2014) MBD5 and MBD6 interact with the human PR-DUB complex through their methyl-CpG-binding domain. Proteomics 14:2179–2189. https://doi.org/10.1002/pmic.201400013

Beckmann AM, Wilce PA (1997) Egr transcription factors in the nervous system. Neurochem Int 31:477–510.; Discussion 517–6. https://doi.org/10.1016/s0197-0186(96)00136-2

Bhende PM, Seaman WT, Delecluse HJ, Kenney SC (2004) The EBV lytic switch protein, Z, preferentially binds to and activates the methylated viral genome. Nat Genet 36:1099–1104. https://doi.org/10.1038/ng1424

Bird A (2002) DNA methylation patterns and epigenetic memory. Genes Dev 16:6–21. https://doi.org/10.1101/gad.947102

Bozon B, Davis S, Laroche S (2003) A requirement for the immediate early gene zif268 in reconsolidation of recognition memory after retrieval. Neuron 40:695–701. https://doi.org/10.1016/s0896-6273(03)00674-3

Buck-Koehntop BA, Stanfield RL, Ekiert DC, Martinez-Yamout MA, Dyson HJ, Wilson IA, Wright PE (2012) Molecular basis for recognition of methylated and specific DNA sequences by the zinc finger protein Kaiso. Proc Natl Acad Sci U S A 109:15229–15234. https://doi.org/10.1073/pnas.1213726109

Buhrmester H, von Kries JP, Stratling WH (1995) Nuclear matrix protein ARBP recognizes a novel DNA sequence motif with high affinity. Biochemistry 34:4108–4117

Burglin TR, Affolter M (2016) Homeodomain proteins: an update. Chromosoma 125:497–521. https://doi.org/10.1007/s00412-015-0543-8

Call KM, Glaser T, Ito CY, Buckler AJ, Pelletier J, Haber DA, Rose EA, Kral A, Yeger H, Lewis WH et al (1990) Isolation and characterization of a zinc finger polypeptide gene at the human chromosome 11 Wilms' tumor locus. Cell 60:509–520. https://doi.org/10.1016/0092-8674(90)90601-a

Chahrour M, Jung SY, Shaw C, Zhou X, Wong ST, Qin J, Zoghbi HY (2008) MeCP2, a key contributor to neurological disease, activates and represses transcription. Science 320:1224–1229. https://doi.org/10.1126/science.1153252

Charlton J, Pritchard-Jones K (2016) WT1 mutation in childhood cancer. Methods Mol Biol 1467:1–14. https://doi.org/10.1007/978-1-4939-4023-3_1

Chen RZ, Akbarian S, Tudor M, Jaenisch R (2001) Deficiency of methyl-CpG binding protein-2 in CNS neurons results in a Rett-like phenotype in mice. Nat Genet 27:327–331. https://doi.org/10.1038/85906

Chen X, Whitney EM, Gao SY, Yang VW (2003) Transcriptional profiling of Kruppel-like factor 4 reveals a function in cell cycle regulation and epithelial

differentiation. J Mol Biol 326:665–677. https://doi.org/10.1016/s0022-2836(02)01449-3

Chen K, Long Q, Xing G, Wang T, Wu Y, Li L, Qi J, Zhou Y, Ma B, Scholer HR, Nie J, Pei D, Liu X (2020) Heterochromatin loosening by the Oct4 linker region facilitates Klf4 binding and iPSC reprogramming. EMBO J 39:e99165. https://doi.org/10.15252/embj.201899165

Choo Y, Klug A (1997) Physical basis of a protein-DNA recognition code. Curr Opin Struct Biol 7:117–125. https://doi.org/10.1016/s0959-440x(97)80015-2

Chronis C, Fiziev P, Papp B, Butz S, Bonora G, Sabri S, Ernst J, Plath K (2017) Cooperative binding of transcription factors orchestrates reprogramming. Cell 168:442–459 e20. https://doi.org/10.1016/j.cell.2016.12.016

Clouaire T, de Las Heras JI, Merusi C, Stancheva I (2010) Recruitment of MBD1 to target genes requires sequence-specific interaction of the MBD domain with methylated DNA. Nucleic Acids Res 38:4620–4634. https://doi.org/10.1093/nar/gkq228

Cukier HN, Rabionet R, Konidari I, Rayner-Evans MY, Baltos ML, Wright HH, Abramson RK, Martin ER, Cuccaro ML, Pericak-Vance MA, Gilbert JR (2010) Novel variants identified in methyl-CpG-binding domain genes in autistic individuals. Neurogenetics 11:291–303. https://doi.org/10.1007/s10048-009-0228-7

Cukier HN, Lee JM, Ma D, Young JI, Mayo V, Butler BL, Ramsook SS, Rantus JA, Abrams AJ, Whitehead PL, Wright HH, Abramson RK, Haines JL, Cuccaro ML, Pericak-Vance MA, Gilbert JR (2012) The expanding role of MBD genes in autism: identification of a MECP2 duplication and novel alterations in MBD5, MBD6, and SETDB1. Autism Res 5:385–397. https://doi.org/10.1002/aur.1251

Daniel JM, Reynolds AB (1999) The catenin p120(ctn) interacts with Kaiso, a novel BTB/POZ domain zinc finger transcription factor. Mol Cell Biol 19:3614–3623. https://doi.org/10.1128/MCB.19.5.3614

Daniel JM, Spring CM, Crawford HC, Reynolds AB, Baig A (2002) 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 30:2911–2919. https://doi.org/10.1093/nar/gkf398

Domcke S, Bardet AF, Adrian Ginno P, Hartl D, Burger L, Schubeler D (2015) Competition between DNA methylation and transcription factors determines binding of NRF1. Nature 528:575–579. https://doi.org/10.1038/nature16462

Dos Santos RL, Tosti L, Radzisheuskaya A, Caballero IM, Kaji K, Hendrich B, Silva JCR (2014) MBD3/NuRD facilitates induction of pluripotency in a context-dependent manner. Cell Stem Cell 15:392. https://doi.org/10.1016/j.stem.2014.08.005

Du Y, Liu B, Guo F, Xu G, Ding Y, Liu Y, Sun X, Xu G (2012) The essential role of Mbd5 in the regulation of somatic growth and glucose homeostasis in mice. PLoS One 7:e47358. https://doi.org/10.1371/journal.pone.0047358

Duclot F, Kabbaj M (2017) The role of early growth response 1 (EGR1) in brain plasticity and neuropsychiatric disorders. Front Behav Neurosci 11:35. https://doi.org/10.3389/fnbeh.2017.00035

Eferl R, Wagner EF (2003) AP-1: a double-edged sword in tumorigenesis. Nat Rev Cancer 3:859–868. https://doi.org/10.1038/nrc1209

El-Karim EA, Hagos EG, Ghaleb AM, Yu B, Yang VW (2013) Kruppel-like factor 4 regulates genetic stability in mouse embryonic fibroblasts. Mol Cancer 12:89. https://doi.org/10.1186/1476-4598-12-89

Fang J, Cheng J, Wang J, Zhang Q, Liu M, Gong R, Wang P, Zhang X, Feng Y, Lan W, Gong Z, Tang C, Wong J, Yang H, Cao C, Xu Y (2016) Hemimethylated DNA opens a closed conformation of UHRF1 to facilitate its histone recognition. Nat Commun 7:11197. https://doi.org/10.1038/ncomms11197

Feng Q, Zhang Y (2001) The MeCP1 complex represses transcription through preferential binding, remodeling, and deacetylating methylated nucleosomes. Genes Dev 15:827–832. https://doi.org/10.1101/gad.876201

Filion GJ, Zhenilo S, Salozhin S, Yamada D, Prokhortchouk E, Defossez PA (2006) A family of human zinc finger proteins that bind methylated DNA and repress transcription. Mol Cell Biol 26:169–181. https://doi.org/10.1128/MCB.26.1.169-181.2006

Fraga MF, Ballestar E, Montoya G, Taysavang P, Wade PA, Esteller M (2003) The affinity of different MBD proteins for a specific methylated locus depends on their intrinsic binding properties. Nucleic Acids Res 31:1765–1774. https://doi.org/10.1093/nar/gkg249

Fujita N, Takebayashi S, Okumura K, Kudo S, Chiba T, Saya H, Nakao M (1999) Methylation-mediated transcriptional silencing in euchromatin by methyl-CpG binding protein MBD1 isoforms. Mol Cell Biol 19:6415–6426. https://doi.org/10.1128/MCB.19.9.6415

Fujita N, Shimotake N, Ohki I, Chiba T, Saya H, Shirakawa M, Nakao M (2000) Mechanism of transcriptional regulation by methyl-CpG binding protein MBD1. Mol Cell Biol 20:5107–5118. https://doi.org/10.1128/MCB.20.14.5107-5118.2000

Gabel HW, Kinde B, Stroud H, Gilbert CS, Harmin DA, Kastan NR, Hemberg M, Ebert DH, Greenberg ME (2015) Disruption of DNA-methylation-dependent long gene repression in Rett syndrome. Nature 522:89–93. https://doi.org/10.1038/nature14319

Gelato KA, Tauber M, Ong MS, Winter S, Hiragami-Hamada K, Sindlinger J, Lemak A, Bultsma Y, Houliston S, Schwarzer D, Divecha N, Arrowsmith CH, Fischle W (2014) Accessibility of different histone H3-binding domains of UHRF1 is allosterically regulated by phosphatidylinositol 5-phosphate. Mol Cell 54:905–919. https://doi.org/10.1016/j.molcel.2014.04.004

Gessler M, Poustka A, Cavenee W, Neve RL, Orkin SH, Bruns GA (1990) Homozygous deletion in Wilms

tumours of a zinc-finger gene identified by chromosome jumping. Nature 343:774–778. https://doi.org/10.1038/343774a0

Ghosh RP, Horowitz-Scherer RA, Nikitina T, Shlyakhtenko LS, Woodcock CL (2010) MeCP2 binds cooperatively to its substrate and competes with histone H1 for chromatin binding sites. Mol Cell Biol 30:4656–4670. https://doi.org/10.1128/MCB.00379-10

Gnanapragasam MN, Scarsdale JN, Amaya ML, Webb HD, Desai MA, Walavalkar NM, Wang SZ, Zu Zhu S, Ginder GD, Williams DC Jr (2011) 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 108:7487–7492. https://doi.org/10.1073/pnas.1015341108

Gunther K, Rust M, Leers J, Boettger T, Scharfe M, Jarek M, Bartkuhn M, Renkawitz R (2013) Differential roles for MBD2 and MBD3 at methylated CpG islands, active promoters and binding to exon sequences. Nucleic Acids Res 41:3010–3021. https://doi.org/10.1093/nar/gkt035

Guo JU, Su Y, Shin JH, Shin J, Li H, Xie B, Zhong C, Hu S, Le T, Fan G, Zhu H, Chang Q, Gao Y, Ming GL, Song H (2014) Distribution, recognition and regulation of non-CpG methylation in the adult mammalian brain. Nat Neurosci 17:215–222. https://doi.org/10.1038/nn.3607

Gustems M, Woellmer A, Rothbauer U, Eck SH, Wieland T, Lutter D, Hammerschmidt W (2014) c-Jun/c-Fos heterodimers regulate cellular genes via a newly identified class of methylated DNA sequence motifs. Nucleic Acids Res 42:3059–3072. https://doi.org/10.1093/nar/gkt1323

Guy J, Hendrich B, Holmes M, Martin JE, Bird A (2001) A mouse Mecp2-null mutation causes neurological symptoms that mimic Rett syndrome. Nat Genet 27:322–326. https://doi.org/10.1038/85899

Harrison JS, Cornett EM, Goldfarb D, DaRosa PA, Li ZM, Yan F, Dickson BM, Guo AH, Cantu DV, Kaustov L, Brown PJ, Arrowsmith CH, Erie DA, Major MB, Klevit RE, Krajewski K, Kuhlman B, Strahl BD, Rothbart SB (2016) Hemi-methylated DNA regulates DNA methylation inheritance through allosteric activation of H3 ubiquitylation by UHRF1. Elife 5. https://doi.org/10.7554/eLife.17101

Hartwig S, Ho J, Pandey P, Macisaac K, Taglienti M, Xiang M, Alterovitz G, Ramoni M, Fraenkel E, Kreidberg JA (2010) Genomic characterization of Wilms' tumor suppressor 1 targets in nephron progenitor cells during kidney development. Development 137:1189–1203. https://doi.org/10.1242/dev.045732

Hashimoto H, Horton JR, Zhang X, Bostick M, Jacobsen SE, Cheng X (2008) The SRA domain of UHRF1 flips 5-methylcytosine out of the DNA helix. Nature 455:826–829. https://doi.org/10.1038/nature07280

Hashimoto H, Zhang X, Cheng X (2012) Excision of thymine and 5-hydroxymethyluracil by the MBD4 DNA glycosylase domain: structural basis and implications for active DNA demethylation. Nucleic Acids Res 40:8276–8284. https://doi.org/10.1093/nar/gks628

Hashimoto H, Olanrewaju YO, Zheng Y, Wilson GG, Zhang X, Cheng X (2014) Wilms tumor protein recognizes 5-carboxylcytosine within a specific DNA sequence. Genes Dev 28:2304–2313. https://doi.org/10.1101/gad.250746.114

Hashimoto H, Wang D, Steves AN, Jin P, Blumenthal RM, Zhang X, Cheng X (2016) Distinctive Klf4 mutants determine preference for DNA methylation status. Nucleic Acids Res 44:10177–10185. https://doi.org/10.1093/nar/gkw774

Hashimoto H, Wang D, Horton JR, Zhang X, Corces VG, Cheng X (2017) Structural basis for the versatile and methylation-dependent binding of CTCF to DNA. Mol Cell 66(711-720):e3. https://doi.org/10.1016/j.molcel.2017.05.004

Hastie ND (2017) Wilms' tumour 1 (WT1) in development, homeostasis and disease. Development 144:2862–2872. https://doi.org/10.1242/dev.153163

Hendrich B, Bird A (1998) Identification and characterization of a family of mammalian methyl-CpG binding proteins. Mol Cell Biol 18:6538–6547. https://doi.org/10.1128/MCB.18.11.6538

Hendrich B, Tweedie S (2003) The methyl-CpG binding domain and the evolving role of DNA methylation in animals. Trends Genet 19:269–277. https://doi.org/10.1016/S0168-9525(03)00080-5

Hendrich B, Hardeland U, Ng HH, Jiricny J, Bird A (1999) The thymine glycosylase MBD4 can bind to the product of deamination at methylated CpG sites. Nature 401:301–304. https://doi.org/10.1038/45843

Hendrich B, Guy J, Ramsahoye B, Wilson VA, Bird A (2001) Closely related proteins MBD2 and MBD3 play distinctive but interacting roles in mouse development. Genes Dev 15:710–723. https://doi.org/10.1101/gad.194101

Hess J, Angel P, Schorpp-Kistner M (2004) AP-1 subunits: quarrel and harmony among siblings. J Cell Sci 117:5965–5973. https://doi.org/10.1242/jcs.01589

Hirai H, Zhang P, Dayaram T, Hetherington CJ, Mizuno S, Imanishi J, Akashi K, Tenen DG (2006) C/EBPbeta is required for 'emergency' granulopoiesis. Nat Immunol 7:732–739. https://doi.org/10.1038/ni1354

Ho KL, McNae IW, Schmiedeberg L, Klose RJ, Bird AP, Walkinshaw MD (2008) MeCP2 binding to DNA depends upon hydration at methyl-CpG. Mol Cell 29:525–531. https://doi.org/10.1016/j.molcel.2007.12.028

Hong S, Wang D, Horton JR, Zhang X, Speck SH, Blumenthal RM, Cheng X (2017) Methyl-dependent and spatial-specific DNA recognition by the orthologous transcription factors human AP-1 and Epstein-Barr virus Zta. Nucleic Acids Res 45:2503–2515. https://doi.org/10.1093/nar/gkx057

Hopfner R, Mousli M, Jeltsch JM, Voulgaris A, Lutz Y, Marin C, Bellocq JP, Oudet P, Bronner C (2000)

ICBP90, a novel human CCAAT binding protein, involved in the regulation of topoisomerase IIalpha expression. Cancer Res 60:121–128

Hu L, Li Z, Wang P, Lin Y, Xu Y (2011) Crystal structure of PHD domain of UHRF1 and insights into recognition of unmodified histone H3 arginine residue 2. Cell Res 21:1374–1378. https://doi.org/10.1038/cr.2011.124

Hu S, Wan J, Su Y, Song Q, Zeng Y, Nguyen HN, Shin J, Cox E, Rho HS, Woodard C, Xia S, Liu S, Lyu H, Ming GL, Wade H, Song H, Qian J, Zhu H (2013) DNA methylation presents distinct binding sites for human transcription factors. Elife 2:e00726. https://doi.org/10.7554/eLife.00726

Huang Y, Zhang H, Wang L, Tang C, Qin X, Wu X, Pan M, Tang Y, Yang Z, Babarinde IA, Lin R, Ji G, Lai Y, Xu X, Su J, Wen X, Satoh T, Ahmed T, Malik V, Ward C, Volpe G, Guo L, Chen J, Sun L, Li Y, Huang X, Bao X, Gao F, Liu B, Zheng H, Jauch R, Lai L, Pan G, Chen J, Testa G, Akira S, Hu J, Pei D, Hutchins AP, Esteban MA, Qin B (2020) JMJD3 acts in tandem with KLF4 to facilitate reprogramming to pluripotency. Nat Commun 11:5061. https://doi.org/10.1038/s41467-020-18900-z

Hudson NO, Whitby FG, Buck-Koehntop BA (2018) Structural insights into methylated DNA recognition by the C-terminal zinc fingers of the DNA reader protein ZBTB38. J Biol Chem 293:19835–19843. https://doi.org/10.1074/jbc.RA118.005147

Jobe EM, Gao Y, Eisinger BE, Mladucky JK, Giuliani CC, Kelnhofer LE, Zhao X (2017) Methyl-CpG-binding protein MBD1 regulates neuronal lineage commitment through maintaining adult neural stem cell identity. J Neurosci 37:523–536. https://doi.org/10.1523/JNEUROSCI.1075-16.2016

Jorgensen HF, Ben-Porath I, Bird AP (2004) Mbd1 is recruited to both methylated and nonmethylated CpGs via distinct DNA binding domains. Mol Cell Biol 24:3387–3395. https://doi.org/10.1128/MCB.24.8.3387-3395.2004

Kaji K, Caballero IM, MacLeod R, Nichols J, Wilson VA, Hendrich B (2006) The NuRD component Mbd3 is required for pluripotency of embryonic stem cells. Nat Cell Biol 8:285–292. https://doi.org/10.1038/ncb1372

Kaplun D, Starshin A, Sharko F, Gainova K, Filonova G, Zhigalova N, Mazur A, Prokhortchouk E, Zhenilo S (2021) Kaiso regulates DNA methylation homeostasis. Int J Mol Sci 22. https://doi.org/10.3390/ijms22147587

Karin M, Liu Z, Zandi E (1997) AP-1 function and regulation. Curr Opin Cell Biol 9:240–246. https://doi.org/10.1016/s0955-0674(97)80068-3

Katz JP, Perreault N, Goldstein BG, Lee CS, Labosky PA, Yang VW, Kaestner KH (2002) The zinc-finger transcription factor Klf4 is required for terminal differentiation of goblet cells in the colon. Development 129:2619–2628

Kim JH, Jeong IY, Lim Y, Lee YH, Shin SY (2011) Estrogen receptor beta stimulates Egr-1 transcription via MEK1/Erk/Elk-1 cascade in C6 glioma cells. BMB Rep 44:452–457. https://doi.org/10.5483/BMBRep.2011.44.7.452

Kinde B, Gabel HW, Gilbert CS, Griffith EC, Greenberg ME (2015) Reading the unique DNA methylation landscape of the brain: non-CpG methylation, hydroxymethylation, and MeCP2. Proc Natl Acad Sci U S A 112:6800–6806. https://doi.org/10.1073/pnas.1411269112

Kinde B, Wu DY, Greenberg ME, Gabel HW (2016) DNA methylation in the gene body influences MeCP2-mediated gene repression. Proc Natl Acad Sci U S A 113:15114–15119. https://doi.org/10.1073/pnas.1618737114

Klose RJ, Sarraf SA, Schmiedeberg L, McDermott SM, Stancheva I, Bird AP (2005) DNA binding selectivity of MeCP2 due to a requirement for A/T sequences adjacent to methyl-CpG. Mol Cell 19:667–678. https://doi.org/10.1016/j.molcel.2005.07.021

Klug A (2010) The discovery of zinc fingers and their applications in gene regulation and genome manipulation. Annu Rev Biochem 79:213–231. https://doi.org/10.1146/annurev-biochem-010909-095056

Koh DI, Han D, Ryu H, Choi WI, Jeon BN, Kim MK, Kim Y, Kim JY, Parry L, Clarke AR, Reynolds AB, Hur MW (2014) KAISO, a critical regulator of p53-mediated transcription of CDKN1A and apoptotic genes. Proc Natl Acad Sci U S A 111:15078–15083. https://doi.org/10.1073/pnas.1318780111

Koh DI, An H, Kim MY, Jeon BN, Choi SH, Hur SS, Hur MW (2015) Transcriptional activation of APAF1 by KAISO (ZBTB33) and p53 is attenuated by RelA/p65. Biochim Biophys Acta 1849:1170–1178. https://doi.org/10.1016/j.bbagrm.2015.07.008

Kohli RM, Zhang Y (2013) TET enzymes, TDG and the dynamics of DNA demethylation. Nature 502:472–479. https://doi.org/10.1038/nature12750

Kriaucionis S, Tahiliani M (2014) Expanding the epigenetic landscape: novel modifications of cytosine in genomic DNA. Cold Spring Harb Perspect Biol 6:a018630. https://doi.org/10.1101/cshperspect.a018630

Kribelbauer JF, Laptenko O, Chen S, Martini GD, Freed-Pastor WA, Prives C, Mann RS, Bussemaker HJ (2017) Quantitative analysis of the DNA methylation sensitivity of transcription factor complexes. Cell Rep 19:2383–2395. https://doi.org/10.1016/j.celrep.2017.05.069

Laget S, Joulie M, Le Masson F, Sasai N, Christians E, Pradhan S, Roberts RJ, Defossez PA (2010) The human proteins MBD5 and MBD6 associate with heterochromatin but they do not bind methylated DNA. PLoS One 5:e11982. https://doi.org/10.1371/journal.pone.0011982

Landschulz WH, Johnson PF, McKnight SL (1989) The DNA binding domain of the rat liver nuclear protein C/EBP is bipartite. Science 243:1681–1688. https://doi.org/10.1126/science.2494700

Lee JH, Skalnik DG (2005) CpG-binding protein (CXXC finger protein 1) is a component of the mammalian Set1 histone H3-Lys4 methyltransferase complex, the analogue of the yeast Set1/COMPASS complex. J Biol Chem 280:41725–41731. https://doi.org/10.1074/jbc.M508312200

Lee JL, Everitt BJ, Thomas KL (2004) Independent cellular processes for hippocampal memory consolidation and reconsolidation. Science 304:839–843. https://doi.org/10.1126/science.1095760

Le Guezennec X, Vermeulen M, Brinkman AB, Hoeijmakers WA, Cohen A, Lasonder E, Stunnenberg HG (2006) MBD2/NuRD and MBD3/NuRD, two distinct complexes with different biochemical and functional properties. Mol Cell Biol 26:843–851. https://doi.org/10.1128/MCB.26.3.843-851.2006

Lei M, Tempel W, Chen S, Liu K, Min J (2019) Plasticity at the DNA recognition site of the MeCP2 mCG-binding domain. Biochim Biophys Acta Gene Regul Mech 1862:194409. https://doi.org/10.1016/j.bbagrm.2019.194409

Lewis JD, Meehan RR, Henzel WJ, Maurer-Fogy I, Jeppesen P, Klein F, Bird A (1992) Purification, sequence, and cellular localization of a novel chromosomal protein that binds to methylated DNA. Cell 69:905–914. https://doi.org/10.1016/0092-8674(92)90610-o

Li H, Yamagata T, Mori M, Yasuhara A, Momoi MY (2005) Mutation analysis of methyl-CpG binding protein family genes in autistic patients. Brain Dev 27:321–325. https://doi.org/10.1016/j.braindev.2004.08.003

Li X, Ito M, Zhou F, Youngson N, Zuo X, Leder P, Ferguson-Smith AC (2008) A maternal-zygotic effect gene, Zfp57, maintains both maternal and paternal imprints. Dev Cell 15:547–557. https://doi.org/10.1016/j.devcel.2008.08.014

Li T, Wang L, Du Y, Xie S, Yang X, Lian F, Zhou Z, Qian C (2018) Structural and mechanistic insights into UHRF1-mediated DNMT1 activation in the maintenance DNA methylation. Nucleic Acids Res 46:3218–3231. https://doi.org/10.1093/nar/gky104

Liang CC, Zhan B, Yoshikawa Y, Haas W, Gygi SP, Cohn MA (2015) UHRF1 is a sensor for DNA interstrand crosslinks and recruits FANCD2 to initiate the Fanconi anemia pathway. Cell Rep 10:1947–1956. https://doi.org/10.1016/j.celrep.2015.02.053

Liu K, Min J (2019) Structural basis for the recognition of non-methylated DNA by the CXXC domain. J Mol Biol. https://doi.org/10.1016/j.jmb.2019.09.025

Liu Y, Toh H, Sasaki H, Zhang X, Cheng X (2012) An atomic model of Zfp57 recognition of CpG methylation within a specific DNA sequence. Genes Dev 26:2374–2379. https://doi.org/10.1101/gad.202200.112

Liu Y, Zhang X, Blumenthal RM, Cheng X (2013) A common mode of recognition for methylated CpG. Trends Biochem Sci 38:177–183. https://doi.org/10.1016/j.tibs.2012.12.005

Liu Y, Olanrewaju YO, Zheng Y, Hashimoto H, Blumenthal RM, Zhang X, Cheng X (2014) Structural basis for Klf4 recognition of methylated DNA. Nucleic Acids Res 42:4859–4867. https://doi.org/10.1093/nar/gku134

Liu Y, Zhang B, Meng X, Korn MJ, Parent JM, Lu LY, Yu X (2017) UHRF2 regulates local 5-methylcytosine and suppresses spontaneous seizures. Epigenetics 12:551–560. https://doi.org/10.1080/15592294.2017.1314423

Liu K, Xu C, Lei M, Yang A, Loppnau P, Hughes TR, Min J (2018) Structural basis for the ability of MBD domains to bind methyl-CG and TG sites in DNA. J Biol Chem 293:7344–7354. https://doi.org/10.1074/jbc.RA118.001785

Liu K, Lei M, Wu Z, Gan B, Cheng H, Li Y, Min J (2019) Structural analyses reveal that MBD3 is a methylated CG binder. FEBS J 286:3240–3254. https://doi.org/10.1111/febs.14850

Liu X, Xu B, Yang J, He L, Zhang Z, Cheng X, Yu H, Liu X, Jin T, Peng Y, Huang Y, Xia L, Wang Y, Wu J, Wu X, Liu S, Shan L, Yang X, Sun L, Liang J, Zhang Y, Shang Y (2021) UHRF2 commissions the completion of DNA demethylation through allosteric activation by 5hmC and K33-linked ubiquitination of XRCC1. Mol Cell 81:2960–2974 e7. https://doi.org/10.1016/j.molcel.2021.05.022

Lu Y, Loh YH, Li H, Cesana M, Ficarro SB, Parikh JR, Salomonis N, Toh CX, Andreadis ST, Luckey CJ, Collins JJ, Daley GQ, Marto JA (2014) Alternative splicing of MBD2 supports self-renewal in human pluripotent stem cells. Cell Stem Cell 15:92–101. https://doi.org/10.1016/j.stem.2014.04.002

Luo M, Ling T, Xie W, Sun H, Zhou Y, Zhu Q, Shen M, Zong L, Lyu G, Zhao Y, Ye T, Gu J, Tao W, Lu Z, Grummt I (2013) NuRD blocks reprogramming of mouse somatic cells into pluripotent stem cells. Stem Cells 31:1278–1286. https://doi.org/10.1002/stem.1374

Mackay DJ, Callaway JL, Marks SM, White HE, Acerini CL, Boonen SE, Dayanikli P, Firth HV, Goodship JA, Haemers AP, Hahnemann JM, Kordonouri O, Masoud AF, Oestergaard E, Storr J, Ellard S, Hattersley AT, Robinson DO, Temple IK (2008) Hypomethylation of multiple imprinted loci in individuals with transient neonatal diabetes is associated with mutations in ZFP57. Nat Genet 40:949–951. https://doi.org/10.1038/ng.187

Mancini M, Magnani E, Macchi F, Bonapace IM (2021) The multi-functionality of UHRF1: epigenome maintenance and preservation of genome integrity. Nucleic Acids Res 49:6053–6068. https://doi.org/10.1093/nar/gkab293

Meehan RR, Lewis JD, McKay S, Kleiner EL, Bird AP (1989) Identification of a mammalian protein that binds specifically to DNA containing methylated CpGs. Cell 58:499–507. https://doi.org/10.1016/0092-8674(89)90430-3

Meehan RR, Lewis JD, Bird AP (1992) Characterization of MeCP2, a vertebrate DNA binding protein with

affinity for methylated DNA. Nucleic Acids Res 20: 5085–5092. https://doi.org/10.1093/nar/20.19.5085

Millar CB, Guy J, Sansom OJ, Selfridge J, MacDougall E, Hendrich B, Keightley PD, Bishop SM, Clarke AR, Bird A (2002) Enhanced CpG mutability and tumorigenesis in MBD4-deficient mice. Science 297: 403–405. https://doi.org/10.1126/science.1073354

Miller M (2009) The importance of being flexible: the case of basic region leucine zipper transcriptional regulators. Curr Protein Pept Sci 10:244–269. https://doi.org/10.2174/138920309788452164

Miotto B, Chibi M, Xie P, Koundrioukoff S, Moolman-Smook H, Pugh D, Debatisse M, He F, Zhang L, Defossez PA (2014) The RBBP6/ZBTB38/MCM10 axis regulates DNA replication and common fragile site stability. Cell Rep 7:575–587. https://doi.org/10.1016/j.celrep.2014.03.030

Mori T, Li Y, Hata H, Ono K, Kochi H (2002) NIRF, a novel RING finger protein, is involved in cell-cycle regulation. Biochem Biophys Res Commun 296:530–536. https://doi.org/10.1016/s0006-291x(02)00890-2

Motamedi FJ, Badro DA, Clarkson M, Lecca MR, Bradford ST, Buske FA, Saar K, Hubner N, Brandli AW, Schedl A (2014) WT1 controls antagonistic FGF and BMP-pSMAD pathways in early renal progenitors. Nat Commun 5:4444. https://doi.org/10.1038/ncomms5444

Nady N, Lemak A, Walker JR, Avvakumov GV, Kareta MS, Achour M, Xue S, Duan S, Allali-Hassani A, Zuo X, Wang YX, Bronner C, Chedin F, Arrowsmith CH, Dhe-Paganon S (2011) Recognition of multivalent histone states associated with heterochromatin by UHRF1 protein. J Biol Chem 286:24300–24311. https://doi.org/10.1074/jbc.M111.234104

Nakahashi H, Kieffer Kwon KR, Resch W, Vian L, Dose M, Stavreva D, Hakim O, Pruett N, Nelson S, Yamane A, Qian J, Dubois W, Welsh S, Phair RD, Pugh BF, Lobanenkov V, Hager GL, Casellas R (2013) A genome-wide map of CTCF multivalency redefines the CTCF code. Cell Rep 3:1678–1689. https://doi.org/10.1016/j.celrep.2013.04.024

Nan X, Meehan RR, Bird A (1993) Dissection of the methyl-CpG binding domain from the chromosomal protein MeCP2. Nucleic Acids Res 21:4886–4892. https://doi.org/10.1093/nar/21.21.4886

Ng HH, Zhang Y, Hendrich B, Johnson CA, Turner BM, Erdjument-Bromage H, Tempst P, Reinberg D, Bird A (1999) MBD2 is a transcriptional repressor belonging to the MeCP1 histone deacetylase complex. Nat Genet 23:58–61. https://doi.org/10.1038/12659

Nikolova EN, Stanfield RL, Dyson HJ, Wright PE (2020) A conformational switch in the zinc finger protein kaiso mediates differential readout of specific and methylated DNA sequences. Biochemistry 59:1909–1926. https://doi.org/10.1021/acs.biochem.0c00253

Nishii T, Oikawa Y, Ishida Y, Kawaichi M, Matsuda E (2012) CtBP-interacting BTB zinc finger protein (CIBZ) promotes proliferation and G1/S transition in embryonic stem cells via Nanog. J Biol Chem 287: 12417–12424. https://doi.org/10.1074/jbc.M111.333856

Ohki I, Shimotake N, Fujita N, Jee J, Ikegami T, Nakao M, Shirakawa M (2001) Solution structure of the methyl-CpG binding domain of human MBD1 in complex with methylated DNA. Cell 105:487–497. https://doi.org/10.1016/s0092-8674(01)00324-5

Oikawa Y, Matsuda E, Nishii T, Ishida Y, Kawaichi M (2008) Down-regulation of CIBZ, a novel substrate of caspase-3, induces apoptosis. J Biol Chem 283:14242–14247. https://doi.org/10.1074/jbc.M802257200

Otani J, Arita K, Kato T, Kinoshita M, Kimura H, Suetake I, Tajima S, Ariyoshi M, Shirakawa M (2013) Structural basis of the versatile DNA recognition ability of the methyl-CpG binding domain of methyl-CpG binding domain protein 4. J Biol Chem 288:6351–6362. https://doi.org/10.1074/jbc.M112.431098

Pan H, Bilinovich SM, Kaur P, Riehn R, Wang H, Williams DC Jr (2017) CpG and methylation-dependent DNA binding and dynamics of the methylcytosine binding domain 2 protein at the single-molecule level. Nucleic Acids Res 45:9164–9177. https://doi.org/10.1093/nar/gkx548

Pavletich NP, Pabo CO (1991) Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 A. Science 252:809–817. https://doi.org/10.1126/science.2028256

Pichler G, Wolf P, Schmidt CS, Meilinger D, Schneider K, Frauer C, Fellinger K, Rottach A, Leonhardt H (2011) Cooperative DNA and histone binding by Uhrf2 links the two major repressive epigenetic pathways. J Cell Biochem 112:2585–2593. https://doi.org/10.1002/jcb.23185

Pozner A, Hudson NO, Trewhella J, Terooatea TW, Miller SA, Buck-Koehntop BA (2018) The C-terminal zinc fingers of ZBTB38 are novel selective readers of DNA methylation. J Mol Biol 430:258–271. https://doi.org/10.1016/j.jmb.2017.12.014

Prokhortchouk A, Hendrich B, Jorgensen H, Ruzov A, Wilm M, Georgiev G, Bird A, Prokhortchouk E (2001) The p120 catenin partner Kaiso is a DNA methylation-dependent transcriptional repressor. Genes Dev 15:1613–1618. https://doi.org/10.1101/gad.198501

Prokhortchouk A, Sansom O, Selfridge J, Caballero IM, Salozhin S, Aithozhina D, Cerchietti L, Meng FG, Augenlicht LH, Mariadason JM, Hendrich B, Melnick A, Prokhortchouk E, Clarke A, Bird A (2006) Kaiso-deficient mice show resistance to intestinal cancer. Mol Cell Biol 26:199–208. https://doi.org/10.1128/MCB.26.1.199-208.2006

Quenneville S, Verde G, Corsinotti A, Kapopoulou A, Jakobsson J, Offner S, Baglivo I, Pedone PV, Grimaldi G, Riccio A, Trono D (2011) In embryonic stem cells, ZFP57/KAP1 recognize a methylated hexanucleotide to affect chromatin and DNA methylation of imprinting control regions. Mol Cell 44:361–372. https://doi.org/10.1016/j.molcel.2011.08.032

Rais Y, Zviran A, Geula S, Gafni O, Chomsky E, Viukov S, Mansour AA, Caspi I, Krupalnik V, Zerbib M, Maza I, Mor N, Baran D, Weinberger L, Jaitin DA, Lara-Astiaso D, Blecher-Gonen R, Shipony Z, Mukamel Z, Hagai T, Gilad S, Amann-Zalcenstein D, Tanay A, Amit I, Novershtern N, Hanna JH (2013) Deterministic direct reprogramming of somatic cells to pluripotency. Nature 502:65–70. https://doi.org/10.1038/nature12587

Rajakumara E, Wang Z, Ma H, Hu L, Chen H, Lin Y, Guo R, Wu F, Li H, Lan F, Shi YG, Xu Y, Patel DJ, Shi Y (2011) PHD finger recognition of unmodified histone H3R2 links UHRF1 to regulation of euchromatic gene expression. Mol Cell 43:275–284. https://doi.org/10.1016/j.molcel.2011.07.006

Rajavelu A, Lungu C, Emperle M, Dukatz M, Brohm A, Broche J, Hanelt I, Parsa E, Schiffers S, Karnik R, Meissner A, Carell T, Rathert P, Jurkowska RZ, Jeltsch A (2018) Chromatin-dependent allosteric regulation of DNMT3A activity by MeCP2. Nucleic Acids Res 46:9044–9056. https://doi.org/10.1093/nar/gky715

Ramsahoye BH, Biniszkiewicz D, Lyko F, Clark V, Bird AP, Jaenisch R (2000) Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. Proc Natl Acad Sci U S A 97:5237–5242. https://doi.org/10.1073/pnas.97.10.5237

Rauscher FJ 3rd, Morris JF, Tournay OE, Cook DM, Curran T (1990) Binding of the Wilms' tumor locus zinc finger protein to the EGR-1 consensus sequence. Science 250:1259–1262. https://doi.org/10.1126/science.2244209

Remenyi A, Tomilin A, Pohl E, Lins K, Philippsen A, Reinbold R, Scholer HR, Wilmanns M (2001) Differential dimer activities of the transcription factor Oct-1 by DNA-induced interface swapping. Mol Cell 8:569–580. https://doi.org/10.1016/s1097-2765(01)00336-7

Rhee HS, Pugh BF (2011) Comprehensive genome-wide protein-DNA interactions detected at single-nucleotide resolution. Cell 147:1408–1419. https://doi.org/10.1016/j.cell.2011.11.013

Riccio A, Aaltonen LA, Godwin AK, Loukola A, Percesepe A, Salovaara R, Masciullo V, Genuardi M, Paravatou-Petsotas M, Bassi DE, Ruggeri BA, Klein-Szanto AJ, Testa JR, Neri G, Bellacosa A (1999) The DNA repair gene MBD4 (MED1) is mutated in human carcinomas with microsatellite instability. Nat Genet 23:266–268. https://doi.org/10.1038/15443

Rothbart SB, Krajewski K, Nady N, Tempel W, Xue S, Badeaux AI, Barsyte-Lovejoy D, Martinez JY, Bedford MT, Fuchs SM, Arrowsmith CH, Strahl BD (2012) Association of UHRF1 with methylated H3K9 directs the maintenance of DNA methylation. Nat Struct Mol Biol 19:1155–1160. https://doi.org/10.1038/nsmb.2391

Sanchez-Guerrero E, Jo SR, Chong BH, Khachigian LM (2013) EGFR and the complexity of receptor crosstalk in the cardiovascular system. Curr Mol Med 13:3–12

Sanchez-Mut JV, Heyn H, Vidal E, Moran S, Sayols S, Delgado-Morales R, Schultz MD, Ansoleaga B, Garcia-Esparcia P, Pons-Espinal M, de Lagran MM, Dopazo J, Rabano A, Avila J, Dierssen M, Lott I, Ferrer I, Ecker JR, Esteller M (2016) Human DNA methylomes of neurodegenerative diseases show common epigenomic patterns. Transl Psychiatry 6:e718. https://doi.org/10.1038/tp.2015.214

Scarsdale JN, Webb HD, Ginder GD, Williams DC Jr (2011) Solution structure and dynamic analysis of chicken MBD2 methyl binding domain bound to a target-methylated DNA sequence. Nucleic Acids Res 39:6741–6752. https://doi.org/10.1093/nar/gkr262

Schubeler D (2015) Function and information content of DNA methylation. Nature 517:321–326. https://doi.org/10.1038/nature14192

Schuetz A, Nana D, Rose C, Zocher G, Milanovic M, Koenigsmann J, Blasig R, Heinemann U, Carstanjen D (2011) The structure of the Klf4 DNA-binding domain links to self-renewal and macrophage differentiation. Cell Mol Life Sci 68:3121–3131. https://doi.org/10.1007/s00018-010-0618-x

Schultz MD, He Y, Whitaker JW, Hariharan M, Mukamel EA, Leung D, Rajagopal N, Nery JR, Urich MA, Chen H, Lin S, Lin Y, Jung I, Schmitt AD, Selvaraj S, Ren B, Sejnowski TJ, Wang W, Ecker JR (2015) Human body epigenome maps reveal noncanonical DNA methylation variation. Nature 523:212–216. https://doi.org/10.1038/nature14465

Segre JA, Bauer C, Fuchs E (1999) Klf4 is a transcription factor required for establishing the barrier function of the skin. Nat Genet 22:356–360. https://doi.org/10.1038/11926

Sharma R, Choi KJ, Quan MD, Sharma S, Sankaran B, Park H, LaGrone A, Kim JJ, MacKenzie KR, Ferreon ACM, Kim C, Ferreon JC (2021) Liquid condensation of reprogramming factor KLF4 with DNA provides a mechanism for chromatin organization. Nat Commun 12:5579. https://doi.org/10.1038/s41467-021-25761-7

Skene PJ, Illingworth RS, Webb S, Kerr AR, James KD, Turner DJ, Andrews R, Bird AP (2010) Neuronal MeCP2 is expressed at near histone-octamer levels and globally alters the chromatin state. Mol Cell 37:457–468. https://doi.org/10.1016/j.molcel.2010.01.030

Song J, Teplova M, Ishibe-Murakami S, Patel DJ (2012) Structure-based mechanistic insights into DNMT1-mediated maintenance DNA methylation. Science 335:709–712. https://doi.org/10.1126/science.1214453

Soubry A, Staes K, Parthoens E, Noppen S, Stove C, Bogaert P, van Hengel J, van Roy F (2010) The transcriptional repressor Kaiso localizes at the mitotic spindle and is a constituent of the pericentriolar material. PLoS One 5:e9203. https://doi.org/10.1371/journal.pone.0009203

Stoll R, Lee BM, Debler EW, Laity JH, Wilson IA, Dyson HJ, Wright PE (2007) Structure of the Wilms tumor suppressor protein zinc finger domain bound to DNA. J

Mol Biol 372:1227–1245. https://doi.org/10.1016/j.jmb.2007.07.017

Stroud JC, Lopez-Rodriguez C, Rao A, Chen L (2002) Structure of a TonEBP-DNA complex reveals DNA encircled by a transcription factor. Nat Struct Biol 9:90–94. https://doi.org/10.1038/nsb749

Sun C, Duan P, Luan C (2017) CEBP epigenetic dysregulation as a drug target for the treatment of hematologic and gynecologic malignancies. Curr Drug Targets 18:1142–1151. https://doi.org/10.2174/1389450117666161228160455

Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, Agarwal S, Iyer LM, Liu DR, Aravind L, Rao A (2009) Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. Science 324:930–935. https://doi.org/10.1126/science.1170116

Takahashi K, Yamanaka S (2006) Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell 126:663–676. https://doi.org/10.1016/j.cell.2006.07.024

Talkowski ME, Rosenfeld JA, Blumenthal I, Pillalamarri V, Chiang C, Heilbut A, Ernst C, Hanscom C, Rossin E, Lindgren AM, Pereira S, Ruderfer D, Kirby A, Ripke S, Harris DJ, Lee JH, Ha K, Kim HG, Solomon BD, Gropman AL, Lucente D, Sims K, Ohsumi TK, Borowsky ML, Loranger S, Quade B, Lage K, Miles J, Wu BL, Shen Y, Neale B, Shaffer LG, Daly MJ, Morton CC, Gusella JF (2012) Sequencing chromosomal abnormalities reveals neurodevelopmental loci that confer risk across diagnostic boundaries. Cell 149:525–537. https://doi.org/10.1016/j.cell.2012.03.028

Tate PH, Bird AP (1993) Effects of DNA methylation on DNA-binding proteins and gene expression. Curr Opin Genet Dev 3:226–231. https://doi.org/10.1016/0959-437x(93)90027-m

Tian Y, Paramasivam M, Ghosal G, Chen D, Shen X, Huang Y, Akhter S, Legerski R, Chen J, Seidman MM, Qin J, Li L (2015) UHRF1 contributes to DNA damage repair as a lesion recognition factor and nuclease scaffold. Cell Rep 10:1957–1966. https://doi.org/10.1016/j.celrep.2015.03.038

Tillotson R, Cholewa-Waclaw J, Chhatbar K, Connelly JC, Kirschner SA, Webb S, Koerner MV, Selfridge J, Kelly DA, De Sousa D, Brown K, Lyst MJ, Kriaucionis S, Bird A (2021) Neuronal non-CG methylation is an essential target for MeCP2 function. Mol Cell 81(1260-1275):e12. https://doi.org/10.1016/j.molcel.2021.01.011

Tsukada J, Yoshida Y, Kominato Y, Auron PE (2011) The CCAAT/enhancer (C/EBP) family of basic-leucine zipper (bZIP) transcription factors is a multifaceted highly-regulated system for gene regulation. Cytokine 54:6–19. https://doi.org/10.1016/j.cyto.2010.12.019

Tulchinsky EM, Georgiev GP, Lukanidin EM (1996) Novel AP-1 binding site created by DNA-methylation. Oncogene 12:1737–1745

Vaughan RM, Dickson BM, Cornett EM, Harrison JS, Kuhlman B, Rothbart SB (2018) Comparative biochemical analysis of UHRF proteins reveals molecular mechanisms that uncouple UHRF2 from DNA methylation maintenance. Nucleic Acids Res 46:4405–4416. https://doi.org/10.1093/nar/gky151

Veyrac A, Besnard A, Caboche J, Davis S, Laroche S (2014) The transcription factor Zif268/Egr1, brain plasticity, and memory. Prog Mol Biol Transl Sci 122:89–129. https://doi.org/10.1016/B978-0-12-420170-5.00004-0

von Kries JP, Buhrmester H, Stratling WH (1991) A matrix/scaffold attachment region binding protein: identification, purification, and mode of binding. Cell 64:123–135. https://doi.org/10.1016/0092-8674(91)90214-j

Wade PA, Gegonne A, Jones PL, Ballestar E, Aubry F, Wolffe AP (1999) Mi-2 complex couples DNA methylation to chromatin remodelling and histone deacetylation. Nat Genet 23:62–66. https://doi.org/10.1038/12664

Walavalkar NM, Cramer JM, Buchwald WA, Scarsdale JN, Williams DC Jr (2014) Solution structure and intramolecular exchange of methyl-cytosine binding domain protein 4 (MBD4) on DNA suggests a mechanism to scan for mCpG/TpG mismatches. Nucleic Acids Res 42:11218–11232. https://doi.org/10.1093/nar/gku782

Wan J, Su Y, Song Q, Tung B, Oyinlade O, Liu S, Ying M, Ming GL, Song H, Qian J, Zhu H, Xia S (2017) Methylated cis-regulatory elements mediate KLF4-dependent gene transactivation and cell migration. Elife 6. https://doi.org/10.7554/eLife.20068

Wang H, Maurano MT, Qu H, Varley KE, Gertz J, Pauli F, Lee K, Canfield T, Weaver M, Sandstrom R, Thurman RE, Kaul R, Myers RM, Stamatoyannopoulos JA (2012) Widespread plasticity in CTCF occupancy linked to DNA methylation. Genome Res 22:1680–1688. https://doi.org/10.1101/gr.136101.111

Weitzel JM, Buhrmester H, Stratling WH (1997) Chicken MAR-binding protein ARBP is homologous to rat methyl-CpG-binding protein MeCP2. Mol Cell Biol 17:5656–5666. https://doi.org/10.1128/MCB.17.9.5656

Wong E, Yang K, Kuraguchi M, Werling U, Avdievich E, Fan K, Fazzari M, Jin B, Brown AM, Lipkin M, Edelmann W (2002) Mbd4 inactivation increases Cright-arrowT transition mutations and promotes gastrointestinal tumor formation. Proc Natl Acad Sci U S A 99:14937–14942. https://doi.org/10.1073/pnas.232579299

Woodcock DM, Crowther PJ, Diver WP (1987) The majority of methylated deoxycytidines in human DNA are not in the CpG dinucleotide. Biochem Biophys Res Commun 145:888–894. https://doi.org/10.1016/0006-291x(87)91048-5

Wu Z, Chen S, Zhou M, Jia L, Li Z, Zhang X, Min J, Liu K (2021) Family-wide characterization of methylated DNA binding ability of arabidopsis MBDs. J Mol

Biol 434:167404. https://doi.org/10.1016/j.jmb.2021.167404

Xu C, Bian C, Lam R, Dong A, Min J (2011) The structural basis for selective binding of non-methylated CpG islands by the CFP1 CXXC domain. Nat Commun 2:227. https://doi.org/10.1038/ncomms1237

Xu C, Liu K, Lei M, Yang A, Li Y, Hughes TR, Min J (2018) DNA sequence recognition of human CXXC domains and their structural determinants. Structure 26:85–95 e3. https://doi.org/10.1016/j.str.2017.11.022

Yang J, Horton JR, Wang D, Ren R, Li J, Sun D, Huang Y, Zhang X, Blumenthal RM, Cheng X (2019) Structural basis for effects of CpA modifications on C/EBPbeta binding of DNA. Nucleic Acids Res 47:1774–1785. https://doi.org/10.1093/nar/gky1264

Yasui DH, Peddada S, Bieda MC, Vallero RO, Hogart A, Nagarajan RP, Thatcher KN, Farnham PJ, Lasalle JM (2007) Integrated epigenomic analyses of neuronal MeCP2 reveal a role for long-range interaction with active genes. Proc Natl Acad Sci U S A 104:19416–19421. https://doi.org/10.1073/pnas.0707442104

Yin Y, Morgunova E, Jolma A, Kaasinen E, Sahu B, Khund-Sayeed S, Das PK, Kivioja T, Dave K, Zhong F, Nitta KR, Taipale M, Popov A, Ginno PA, Domcke S, Yan J, Schubeler D, Vinson C, Taipale J (2017) Impact of cytosine methylation on DNA binding specificities of human transcription factors. Science 356. https://doi.org/10.1126/science.aaj2239

Yoon HG, Chan DW, Reynolds AB, Qin J, Wong J (2003) N-CoR mediates DNA methylation-dependent repression through a methyl CpG binding protein Kaiso. Mol Cell 12:723–734. https://doi.org/10.1016/j.molcel.2003.08.008

Yoon HS, Ghaleb AM, Nandan MO, Hisamuddin IM, Dalton WB, Yang VW (2005) Kruppel-like factor 4 prevents centrosome amplification following gamma-irradiation-induced DNA damage. Oncogene 24:4017–4025. https://doi.org/10.1038/sj.onc.1208576

Zandarashvili L, White MA, Esadze A, Iwahara J (2015) Structural impact of complete CpG methylation within target DNA on specific complex formation of the inducible transcription factor Egr-1. FEBS Lett 589:1748–1753. https://doi.org/10.1016/j.febslet.2015.05.022

Zhang Y, Ng HH, Erdjument-Bromage H, Tempst P, Bird A, Reinberg D (1999) Analysis of the NuRD subunits reveals a histone deacetylase core complex and a connection with DNA methylation. Genes Dev 13:1924–1935. https://doi.org/10.1101/gad.13.15.1924

Zhang J, Gao Q, Li P, Liu X, Jia Y, Wu W, Li J, Dong S, Koseki H, Wong J (2011) S phase-dependent interaction with DNMT1 dictates the role of UHRF1 but not UHRF2 in DNA methylation maintenance. Cell Res 21:1723–1739. https://doi.org/10.1038/cr.2011.176

Zhang P, Rausch C, Hastert FD, Boneva B, Filatova A, Patil SJ, Nuber UA, Gao Y, Zhao X, Cardoso MC (2017) Methyl-CpG binding domain protein 1 regulates localization and activity of Tet1 in a CXXC3 domain-dependent manner. Nucleic Acids Res 45:7118–7136. https://doi.org/10.1093/nar/gkx281

Zhao S, Choi M, Overton JD, Bellone S, Roque DM, Cocco E, Guzzo F, English DP, Varughese J, Gasparrini S, Bortolomai I, Buza N, Hui P, Abu-Khalaf M, Ravaggi A, Bignotti E, Bandiera E, Romani C, Todeschini P, Tassi R, Zanotti L, Carrara L, Pecorelli S, Silasi DA, Ratner E, Azodi M, Schwartz PE, Rutherford TJ, Stiegler AL, Mane S, Boggon TJ, Schlessinger J, Lifton RP, Santin AD (2013) Landscape of somatic single-nucleotide and copy-number mutations in uterine serous carcinoma. Proc Natl Acad Sci U S A 110:2916–2921. https://doi.org/10.1073/pnas.1222577110

Zhou T, Xiong J, Wang M, Yang N, Wong J, Zhu B, Xu RM (2014) Structural basis for hydroxymethylcytosine recognition by the SRA domain of UHRF2. Mol Cell 54:879–886. https://doi.org/10.1016/j.molcel.2014.04.003

Zou X, Ma W, Solov'yov IA, Chipot C, Schulten K (2012) Recognition of methylated DNA through methyl-CpG binding domain proteins. Nucleic Acids Res 40:2747–2758. https://doi.org/10.1093/nar/gkr1057

Zuo X, Sheng J, Lau HT, McDonald CM, Andrade M, Cullen DE, Bell FT, Iacovino M, Kyba M, Xu G, Li X (2012) Zinc finger protein ZFP57 requires its co-factor to recruit DNA methyltransferases and maintains DNA methylation imprint in embryonic stem cells via its transcriptional repression domain. J Biol Chem 287:2107–2118. https://doi.org/10.1074/jbc.M111.322644