

Zinc Finger Readers of Methylated DNA

Nicholas O. Hudson and Bethany A. Buck-Koehntop *

Department of Chemistry, University of Utah, Salt Lake City, UT 84112-0850, USA; u1082423@utah.edu

* Correspondence: koehntop@chem.utah.edu; Tel.: +1-801-581-3186

Academic Editor: Junji Iwahara

Received: 14 September 2018; Accepted: 5 October 2018; Published: 7 October 2018

Abstract: DNA methylation is a prevalent epigenetic modification involved in regulating a number of essential cellular processes, including genomic accessibility and transcriptional outcomes. As such, aberrant alterations in global DNA methylation patterns have been associated with a growing number of disease conditions. Nevertheless, the full mechanisms by which DNA methylation information is interpreted and translated into genomic responses is not yet fully understood. Methyl-CpG binding proteins (MBPs) function as important mediators of this essential process by selectively reading DNA methylation signals and translating this information into down-stream cellular outcomes. The Cys₂His₂ zinc finger scaffold is one of the most abundant DNA binding motifs found within human transcription factors, yet only a few zinc finger containing proteins capable of conferring selectivity for mCpG over CpG sites have been characterized. This review summarizes our current structural understanding for the mechanisms by which the zinc finger MBPs evaluated to date read this essential epigenetic mark. Further, some of the biological implications for mCpG readout elicited by this family of MBPs are discussed.

Keywords: zinc finger; methyl-CpG binding proteins; DNA methylation; epigenetics; protein-DNA interactions

1. Introduction

The most prevalent epigenetic modification found on DNA is the addition of a methyl group at the 5-position of cytosine bases. In mammals, the addition of this reversible chemical modification most commonly occurs in the context of CpG dinucleotides and is referred to as DNA methylation. Genomic DNA methylation patterns are primarily established during development by the *de novo* DNA methyltransferases (DNMT), 3A and 3B, and heritably maintained during DNA replication by DNMT1 [1–4]. Across the genome, most CpG sites are methylated (mCpG), with the exception of GC-rich stretches termed CpG islands (CGIs), found in ~70% of all gene promoters [5]. While DNA methylation inherently represents a seemingly small alteration in the chemical signature, this epigenetic alteration plays a significant role in regulating a number of key processes in both normal cell function as well as disease conditions. In the early stages of development, DNA methylation is essential for regulating genomic imprinting, X-chromosome inactivation, and cellular differentiation during embryogenesis [6–8]. Misregulation of DNA methylation patterns has thus been linked to several neuro-developmental diseases, including Rett, fragile X, and ICF (immunodeficiency, centromeric instability and facial anomalies) syndromes [9,10].

Beyond its role in development, DNA methylation functions to maintain overall genomic stability, preserve cellular identity, and control gene expression by working in conjunction with specific patterns of post-translational modifications (PTMs) on histone proteins to regulate chromatin accessibility and nucleosomal positioning [6,9,11,12]. Aberrant alterations in genomic DNA methylation patterns have consequently been associated with a multitude of disease conditions. In particular, it has become evident that nearly all cancers harbor significant alterations in their DNA

methylation patterns that correlate with disease phenotype [13–16]. Specifically, in many cancers, a global decrease in genomic CpG methylation concomitant with hyper-methylation of promoter CGIs is associated with abnormal proliferation of cells preceding tumorigenesis [9,14,17,18]. Further, DNA methylation patterns have been observed to vary with tumor type and grade [19], promote tumor heterogeneity [20], and be clonally inherited [21]. In addition to cancer, there is increasing evidence that misappropriated DNA methylation patterns are associated with a variety of other disease conditions [22–24], including: Metabolic disorders (e.g., obesity and type II diabetes) [25–28]; autoimmune diseases (e.g., type I diabetes, lupus, and rheumatoid arthritis) [29–31]; and neurological diseases (e.g., Parkinson's, Huntington's, and Alzheimer's) [32]; as well as psychiatric disorders (e.g., depression, anxiety, and schizophrenia) [33,34]. Further, evidence of age related causative DNA methylation alterations in disease promotion have been proposed [35]. While there is significant evidence implicating a role for DNA methylation in a multitude of disease conditions, the full mechanisms by which DNA methylation information is interpreted and translated into genomic responses is not yet fully understood.

In the context of promoter CGIs, DNA methylation is typically associated with inducing transcriptional repression at the target site. Two primary mechanisms by which DNA methylation is believed to elicit gene silencing is by either preventing or promoting transcription factor (TF) binding [11,36]. Specifically, the addition of a methyl group on cytosine bases can affect TF readout at CpG sites by introducing local perturbations to the DNA structure as well as presenting a hydrophobic edge within the major groove binding interface [37–39]. These combined alterations to the DNA can have an overall repulsive effect on the ability of TFs to recognize their cognate DNA binding sequences, leading to transcriptional inhibition [40–44]. Alternatively, these modifications can facilitate the recruitment of TFs that have preferential selectivity for mCpG sites, termed methyl-CpG binding proteins (MBPs); a subject that has been extensively reviewed [45–53]. There is evidence to suggest that direct mCpG occupation by some MBPs can recruit chromatin remodeling enzyme complexes to the target site that subsequently leads to alterations in chromatin architecture and transcriptional outcomes [54–60]. In this capacity, MBPs function as essential functional intermediaries in the gene regulation process. As such, there has been significant interest in mechanistically discerning how these MBPs select and interpret DNA methylation signals, as well as evaluating their potential as novel therapeutic targets. Indeed, it has been demonstrated that the depletion of MBPs in the cancerous context can re-activate transcriptional activity at repressed promoters without altering the methylation status [61–64]. This implies that when mediated by MBPs, DNA methylation may be necessary, but not sufficient, to elicit gene silencing.

In contrast to most hyper-methylated promoter CGIs [65,66], high levels of DNA methylation at gene body CGIs have been correlated with actively transcribed genes [67–69], suggesting a more intricate role for DNA methylation in gene regulation. Gene body DNA methylation sites have been proposed to function as alternative promoters during transcription and to regulate differential splicing [70–75]. Further, MBPs have been shown to associate with methylated CGIs at both promoters and exons in embryonic stem and cancer cells [76,77]. Interestingly, gene body binding did not impede transcriptional elongation when preceded by an active promoter [76,77]. More recently, it has been proposed that MBP occupation at non-promoter mCpG loci may indirectly guide TFs that are insensitive to CpG methylation to their consensus motifs within non-methylated promoter CGIs [78,79]. Combined, it is evident that MBPs are essential factors in mediating methyl-dependent gene regulation, though mechanistic insight for many of their functions remains to be fully elucidated.

For many years, the identified TF MBPs have been confined to the methyl-CpG binding domain (MBD) and Cys₂His₂ zinc finger (ZF) containing families. However, more recently, a series of in vitro high-throughput strategies have been utilized to identify additional TF protein families that appear to harbor mCpG binding capabilities [80–83]. Structural insight for some of these newly discovered MBP classes, including the basic leucine-zipper (bZIP) [84] and homeodomain [82] proteins, have become available, though further in vitro and in cell characterization/validation for most of the discovered systems is still needed. The MBD family was the first identified class of MBPs discovered

[85,86] and, as such, a significant body of research has provided insight for the mechanisms of mCpG recognition, roles in methyl-dependent transcriptional regulation, as well as the overall impact of this protein family on cellular functions. Several reviews have been dedicated to these topics [87–94]. The overall contributions from ZF MBPs in reading this essential epigenetic mark and translating the encoded information into down-stream cellular responses have been comparatively reviewed far less. Hence, our goal here is to provide a current review of the structural evidence that has afforded mechanistic insight for how ZF MBPs distinguish mCpG target sites. Further, we will summarize some of the biological implications for mCpG readout elicited by the currently structurally validated ZF MBPs.

2. Mechanistic Recognition of mCpG Sites by Zinc Finger Proteins—Structural Perspective

2.1. Structural Insight for ZF Recognition of Methylated DNA

The Cys₂His₂ ZF scaffold is one of the most abundant DNA binding motifs found within the human TF proteome. Each modular Cys₂His₂ ZF has a core domain comprised of two β -strands that pack against an α -helix ($\beta\beta\alpha$ -fold), which is stabilized by tetrahedral coordination of a Zn²⁺ atom through the side chains of two cysteine and two histidine residues [95,96]. In specific DNA recognition, the α -helix orients along the major groove and utilizes residues at specific positions (−1, +2, +3, and +6, relative to the start of the helix) to generally make base contacts within a three base pair region [95,96]. Many proteins use tandem arrays of these ZF motifs to expand sequence-specificity for a given DNA consensus site. Irrespective of the abundance of Cys₂His₂ ZF containing proteins, only a few high-resolution structures for TFs containing this structural feature that are capable of conferring selectivity for mCpG over CpG sites have become available in recent years. Currently, these ZF TFs include ZBTB33 (also known as Kaiso) [97,98], Zfp57 [99], Klf4 (Krüppel-like factor 4) [100,101], WT1 (Wilms' tumor protein 1) [102], Egr1 (growth response protein 1; also known as Zif268) [102,103], and CTCF (CCCTC-binding factor) [104] (Figure 1), the structures of which have begun to provide mechanistic insight for how ZF MBPs can recognize this essential epigenetic mark.

From the currently available structures, it has become evident that a common mode of ZF recognition for mCpG sites involves the usage of key arginine and glutamate residues (Figure 2) [105–108]. Specifically, in each of these ZF:methylated DNA complexes, an arginine is involved in classical hydrogen bonding interactions with one (Zfp57, Klf4, WT1, Egr1, CTCF) or two (ZBTB33) of the 3'-G residues. This further positions the arginine side chain to make van der Waals contacts with the 5'-mC methyl group on one DNA strand. This designated 'mC-Arg-G triad' [107] is also utilized by the MBD family of MBPs, and has been shown to play a significant role in stabilizing the binding interaction as well as affording discrimination capability between mC and C [109]. Intriguingly, while all of the ZF MBPs utilize this mC-Arg-G triad in mCpG recognition and in vitro exhibit low nM binding for mCpG containing sequences (<200 nM), their preferential ability to distinguish mCpG over CpG sites is quite variable. Klf4, WT1, and CTCF each have a ZF domain that has been shown in vitro to exhibit marginal (~1.5–1.8-fold) selectivity for mCpG over CpG [101,102,104]. Depending on the study, Egr1 has been shown to be completely insensitive or have an ~3-fold sensitivity for mCpG over CpG sites [102,103]. In contrast, both ZBTB33 and Zfp57 exhibit significantly higher selectivity for their methylated DNA targets (≥ 20 -fold relative to CpG) [98,99]. This affinity differential between the ZF MBPs appears to be significantly correlated with the extent to which the conserved glutamate residue is involved in recognition of the mC.

ZBTB33 and Zfp57 exhibit the highest selectivity for mCpG sites and in each case, the glutamate residue is involved in classical hydrogen bonding interactions between the glutamate carbonyl oxygens and the N4 atom of one (Zfp57) or both (ZBTB33) of the mCs, as well as interactions between the glutamate side chain and the mC methyl (Figure 2a,b). Of particular note, the glutamate residue is able to make CH...O type hydrogen bonds with the methyl group of one (Zfp57) or both (ZBTB33) of the mCs. Due to the significant number of contacts between the core glutamate and two cross-strand mCs, it is unsurprising that mutation of this residue to an alanine in ZBTB33 abolishes DNA binding [98]. In contrast, mutation of the correlative glutamate in Zfp57 to an alanine showed no

difference in DNA binding capability [99]. Closer examination of the Zfp57:methylated DNA structure reveals that the glutamate residue adopts two conformations, one that is more idealized for mC recognition, and one that is better suited for positioning another arginine residue for making hydrogen bonds with a guanine base outside of the mCpG core (Figure 2b). It may be that ZBTB33 is overall better able to spatially coordinate optimal recognition of mCpG sites by utilizing two different α -helices to contribute the arginine and glutamate residues, unlike Zfp57 where both residues are positioned within the same α -helix.

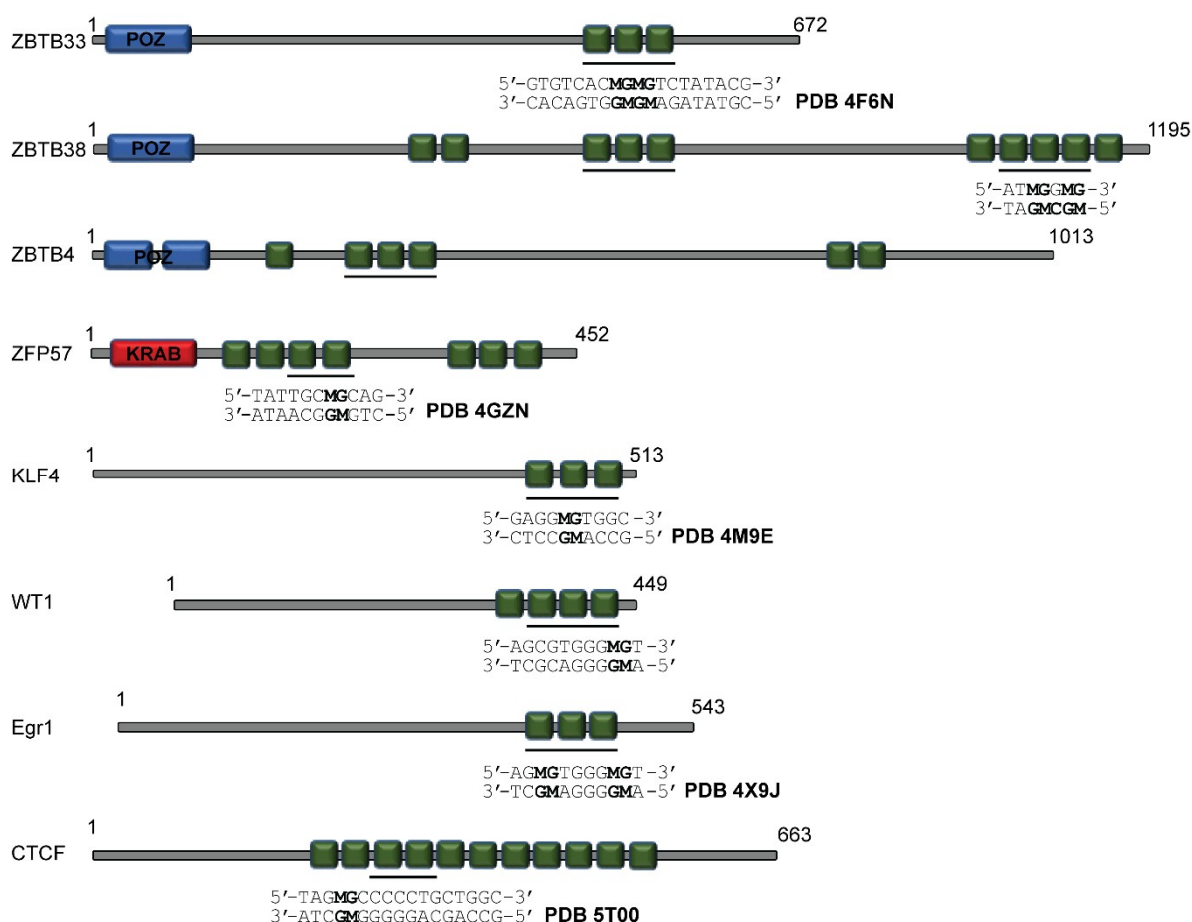


Figure 1. Domain organization for the characterized zinc finger (ZF) methyl-CpG binding proteins (MBPs). For each protein, the ZFs shown to have methyl sensitivity are underlined and the target DNA sequence utilized for high-resolution structural investigations are depicted alongside the Protein Data Bank (PDB) identification for the structures shown in Figure 2. M refers to methylated cytosine. POZ: Pox virus and zing finger domain; KRAB: Krüppel-associated box.

As discussed above, all of the remaining ZF MBPs, including Klf4, WT1, Egr1, and CTCF, have at least one ZF that is either indifferent or has a marginal selectivity for mCpG over CpG sites. For each of these proteins, the glutamate residue is positioned such that it is not capable of forming classical hydrogen bonding interactions with the N4 atoms of the mCs, and is only able to contribute either van der Waals interactions from the side chain and/or CH...O type hydrogen bonding interactions via the carbonyl oxygens with the mC methyl groups (Figure 2c–e). In the cases of Klf4, WT1, and Egr1, there is a conserved aspartate residue preceding the glutamate that simultaneously stabilizes the arginine side chain for recognition of the 3'-G and provides a weaker electrostatic interaction with the mC N4 atom through one of its carbonyl oxygens (Figure 2c,d). Similar to ZBTB33, CTCF also utilizes two different ZF helices to provide the key arginine and glutamate residues for mCpG recognition, however, the glutamate side chain position is fixed such that it is unable to make a classical hydrogen bond with the mC N4 atom (Figure 2e). This seems to be in part due to interactions from a neighboring tyrosine, which positions the glutamate for making a CH...O

hydrogen bond with the mC methyl group, but only affords the glutamate an opportunity to form a weaker electrostatic interaction with the mC N4 atom. Combined, these structures suggest that the loss of a base-specific hydrogen bonding interaction between the glutamate and mC plays a significant role in the observed reduced selectivity for mC over C. Furthermore, Klf4 has been shown to bind equally well to targets that contain interchangeable mCpG or TpG base steps [100]. Given that the glutamate interactions are limited to aliphatic side chain contacts with the methyl at the 5-position of either the mC or T base, the inability to discriminate between these bases is unsurprising. In contrast, ZBTB33 is also capable of recognizing both mCpG containing targets as well as a sequence specific motif containing a TpG site (termed the Kaiso binding site (KBS)) [97,110,111]. However, in this case, the glutamate is still within hydrogen bonding distance of the thymine O4, suggesting that the protonation state of the glutamate must be altered to accommodate recognition of the thymine base, and that ZBTB33 specifically selects for both mCpG and TpG containing targets [97].

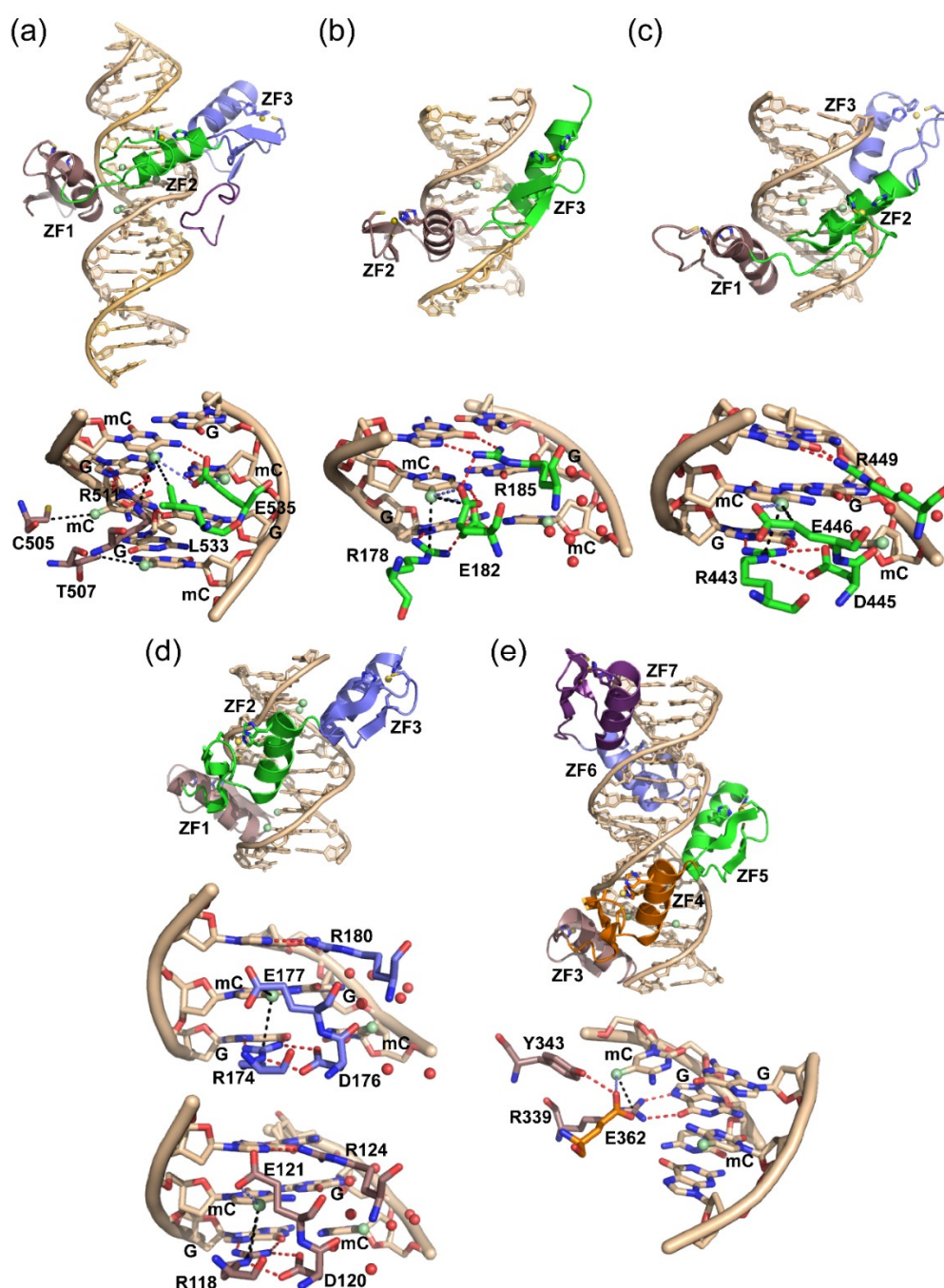


Figure 2. High-resolution crystal structures for members of the ZF MBP family in complex with their respective methylated DNA targets. (a) Human ZBTB33 in complex with a methylated DNA sequence

derived from the *E-cadherin* promoter (PDB 4F6N); (b) mouse Zfp57 in complex with a methylated DNA sequence present in imprinting control regions (PDB 4GZN); (c) mouse Klf4 (Krüppel-like factor 4) in complex with its cognate methylated DNA sequence (PDB 4M9E); (d) human Egr1 (growth response protein 1) in complex with its cognate methylated DNA sequence (PDB 4X91); (e) human CTCF (CCCTC-binding factor) in complex with a methylated version of its core recognition sequence (PDB 5T00). Red spheres indicate water molecules. Red dotted lines denote classical hydrogen bond interactions; blue dotted lines indicate CH...O type hydrogen bonds, and black dotted lines designate van der Waals interactions. For each zoomed-in image, the amino acid side chain color designation matches that of the ZF from which it is derived in the full structural image depicted above.

Another emerging trend observed for ZF MBPs is that the binding at mCpG palindromic sites is asymmetric [106]. For each of the ZF MBPs, except ZBTB33, all the arginine and glutamate interactions predominate at one mCpG (Figure 2b–e). In each of these cases, the remaining palindromic mC is surrounded by an ordered water layer. It has been established that the methyls of mCs are significantly hydrated [112,113], and that this hydration can contribute to the overall binding energetics for MBP recognition of mCpG sites [99,114,115]. In the case of ZBTB33, the glutamate makes cross-strand interactions with both mCs displacing any ordered water molecules around the mC methyl groups. However, the arginine side chain and a proximal leucine residue contribute additional van der Waals contacts for only one mC (Figure 2a), which appears to establish the overall strand preference. Indeed, each of the ZF MBPs has been shown to have similar binding affinities for strand specific hemi-methylated DNA relative to symmetrically methylated DNA sites [97–99,101,102]. The physiological consequence of this asymmetric binding capability for each of these proteins is not yet fully understood.

In addition to base specific readout, the role of local DNA shape is an essential factor in defining the selectivity of a protein for a given DNA target [39,116–118]. Recently, it has been demonstrated that methyl addition to cytosine bases within the context of duplex B-form DNA results in significant perturbations to the local DNA shape predominated by changes in roll and propeller twist [38]. Furthermore, these induced local DNA structural changes have been shown to influence protein recruitment and binding to target genomic sites [37,38]. Thus, based on the available structural information, we sought to determine what impact DNA shape alterations may have on mCpG recognition by ZF MBPs. Our analysis was limited to evaluation of the ZBTB33, Zfp57, and Klf4 methylated DNA complexes as the mCpGs were centrally localized within the target DNA sites for these systems. As structures for the DNA targets in their free forms are not available, methyl-DNA shape [38] was first utilized to determine the effect of methylation on each of the target DNAs utilized in the complex structures. As previously observed [38], methylation at each of the CpG sites resulted in minimal changes to the helical twist and minor groove widths (MGWs), but induced larger changes in both roll and propeller twist (Figure 3). In particular, for all of the DNA sequences, methylation induced a site specific decrease and subsequent increase of the roll at the 5'-mC and 3'-G positions, respectively (Figure 3). Further, the degree of propeller twist was decreased upon methylation across the CpG site for all three DNA sequences evaluated.

To determine the impact of protein binding, CURVES+ [119] was utilized to discern how these DNA shape parameters were altered upon complex formation. For each protein:DNA complex, the magnitude of the roll at the 5'-mC and 3'-G positions were substantially increased and marginally reduced, respectively (Figure 3). These combined observations suggest that the specific directional alterations in the degree of roll upon cytosine methylation may be a common mode of DNA shape alteration needed to facilitate ZF MBP binding. No clear pattern for propeller twist alterations were evident across the three protein:DNA complexes, while all three bound DNA targets exhibited a noticeable increase in MGW. For the Zfp57 and Klf4 methylated DNA complex structures, alterations in helical twist were minimal, whereas in the ZBTB33 methylated DNA complex, there was a noticeable increase in helical twist angles for the 5'-mCs and a decrease for the 3'-Gs (Figure 3). This differential impact on helical twist angles upon protein binding may be correlated with the fact that the key arginine and glutamate residues of ZBTB33 make cross-strand base-specific contacts, while these same residues in Zfp57 and Klf4 contact only one half of the mCpG palindrome. Finally, from

the CURVES+ analysis, it was determined that the overall deformation in DNA bend from B-form was 21.9°, 11.2°, and 4.2° after ZBTB33, Zfp57, and Klf4 binding, respectively. This observation suggests that the proteins exhibiting a higher methyl selectivity induce a larger global bend in the DNA, likely to facilitate base specific contacts at the core mCpG site relative to the less selective binders. Though more high-resolution structures are required to confirm this.

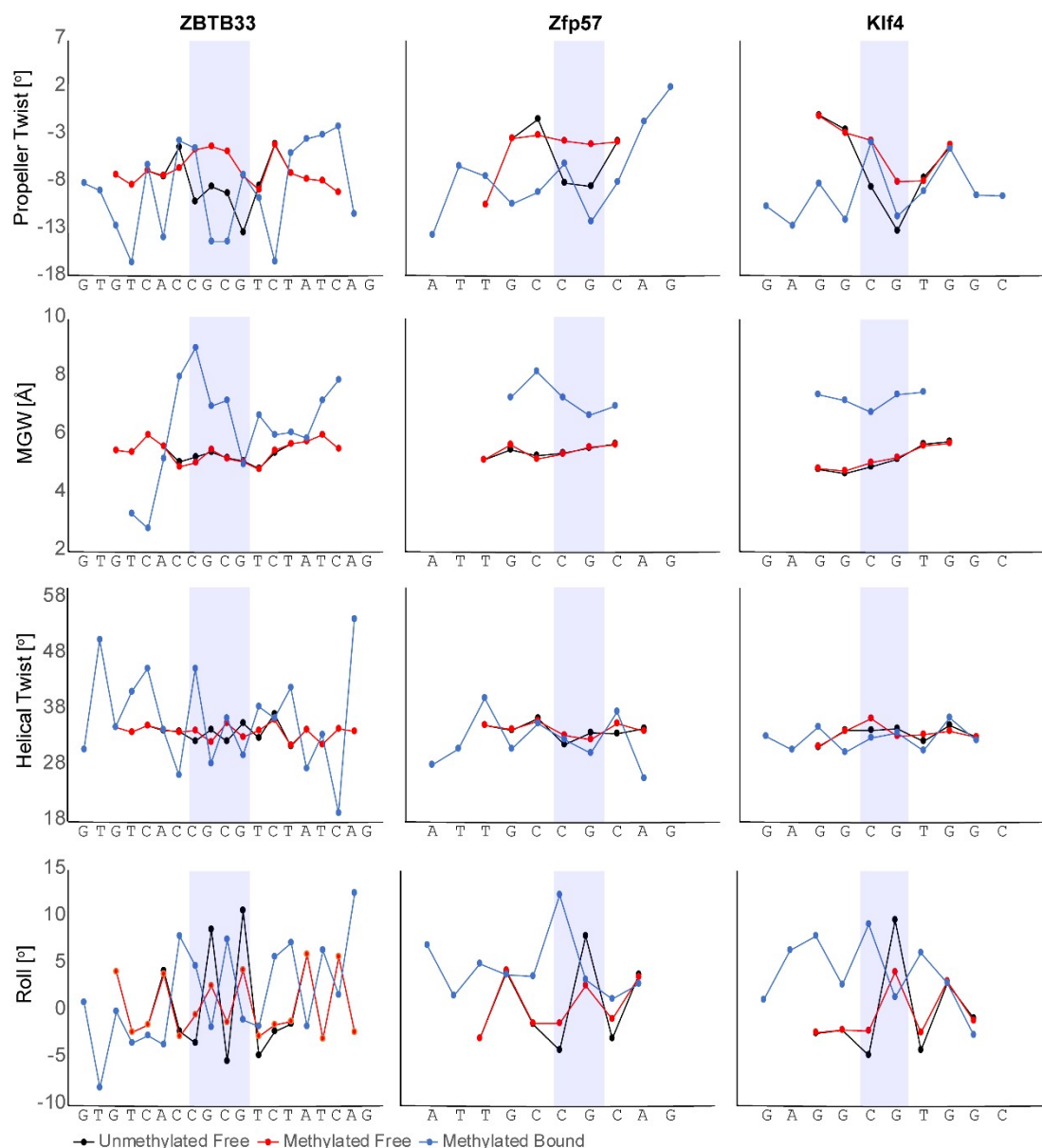


Figure 3. DNA shape analyses for the DNA targets of ZBTB33, Zfp57, and Klf4 used for crystallography in their free unmodified CpG (black), free methyl-CpG (red), and methyl-CpG protein complexed (blue) forms. The blue rectangles are centered on the core CpG dinucleotides.

2.2. Additional Methyl-Selective ZF TFs and Alternative Modes of mCpG Recognition

ZBTB33 represents the founding member of the ZBTB family of MBPs, which also includes ZBTB4 and ZBTB38 [120]. All three proteins share a conserved set of three Cys₂His₂ ZFs that are responsible for methylated DNA recognition [120] (Figure 1). Importantly, while there are not yet high-resolution structures for ZBTB4 and ZBTB38 in complex with their respective methylated DNA binding targets, all of the core residues involved in ZBTB33 recognition of methylated DNA are conserved. Indeed, similar to ZBTB33, mutation of the key glutamate within ZBTB4 to an alanine abolished binding to methylated DNA [121]. Further, in vitro and in cell evidence suggests that both ZBTB4 and ZBTB38 are also highly selective for mCpG sites over CpG [120–124]. Finally, it has been

established that similar to ZBTB33, both ZBTB4 and ZBTB38 exhibit bimodal DNA binding by also recognizing a sequence-specific TpG containing motif [120,121,125].

In addition to the three conserved ZFs, ZBTB4 and ZBTB38 each have multiple N- and C-terminal ZFs (Figure 1), the functions of which have yet to be fully characterized. Recently it has been demonstrated that a sub-set of the five C-terminal ZFs of ZBTB38 also participate in methyl-selective DNA recognition both in vitro and within the cell [124] (Figure 1). Specifically, it was determined that the middle three ZFs (ZFs 7–9) are necessary and sufficient for mCpG recognition, though one or both of the additional ZFs (ZF6 and/or ZF10) are needed to confer optimal affinity during DNA binding. However, unlike any of the other ZF MBPs characterized to date, it was determined that while ZFs 7 and 8 had the key glutamate residue, they were lacking the core arginine required for mCpG recognition. Rather, a lysine residue in ZF7 was found to be substituted at the +6 position typically occupied by the arginine (Figure 4). Systematic mutation of both the lysine and glutamate residues significantly impacted DNA binding, indicating that these two residues are required for mCpG recognition [124]. This is the first evidence of a lysine residue playing a surrogate role for arginine in ZF recognition of mCpG sites, and suggests that an alternative mode for ZF recognition of mCpG exists.

To determine whether other ZF containing proteins harbor this same lysine/glutamate pair, a Basic Local Alignment Search Tool (BLAST) search using ZFs 7–8 of ZBTB38 was conducted. Three other ZF proteins, including one Krüppel-associated box (KRAB) domain protein (ZNF282), were identified (Figure 4). The biological functions for these three factors are not well characterized, though it would be interesting to evaluate their capacity for methyl-selective DNA recognition. Finally, it is important to note that a number of additional ZF proteins have been identified from recent high-throughput studies to potentially have preferential mCpG binding capacity [80–83]. Characterization for these additional ZF proteins is needed to determine whether the mechanisms by which ZF proteins can preferentially select for methylated DNA targets needs to be expanded further.

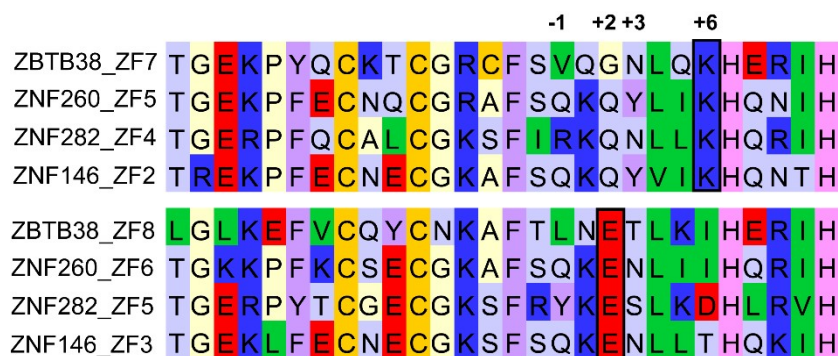


Figure 4. Sequence alignment for additional ZF proteins that share the conserved lysine/glutamate residue pair (boxed) that was shown for the C-terminal ZFs of ZBTB38 to participate in selective mCpG recognition.

3. Physiological Consequence of Methyl Sensitivity

From in vitro biochemical and structural insight, it is evident that the ZF MBPs bind to mCpGs in a sequence-specific context. However, these essential pieces of information are unable to inform on the cellular genomic distributions and functions for these TFs. In recent years, there has been an increasing interest in correlating mechanistic in vitro findings with target selectivity within the cell [118]. In general, determining TF distributions at target sites in the cell is far more complicated due to the fact that typically numerous copies of a consensus motif are present in the genome, but accessibility to these regions can be differential between cellular contexts. Further, binding to a given loci may or may not lead to a cellular response, which is again likely to be cell specific. Nevertheless, the rapidly increasing advances in ChIP-based technologies as well as high-throughput sequencing methodologies have afforded the capability to begin evaluating the cellular readout of methylated

DNA targets for many of the structurally characterized ZF MBPs. Here, we briefly correlate the in vitro structural findings for mCpG sensitivity with the potential cellular relevance for each ZF MBP.

3.1. ZBTB33, ZBTB4, and ZBTB38

Through their conserved set of three Cys²His² set of ZFs, each of the three members of the ZBTB MBP family has been shown to bind to both mCpG as well as sequence specific TpG containing sites [111,120–127]. For ZBTB33, this bimodal DNA binding capacity has provided the capability of functioning to either up- or down-regulate transcription depending on the gene context [128–130]. Further, each protein appears to function in an analogous manner to the MBD family of TFs in that direct promoter occupation recruits chromatin remodeling complexes, which subsequently alters transcriptional outcomes [58–60,131]. A significant body of work combining in vitro and in cell analyses has implicated the functions of these proteins in transcriptionally regulating genes associated with a number of core cellular functions; particularly in the context of cancer. Of the three family members, ZBTB33 is the best characterized. In particular, ZBTB33 transcriptional activities have been associated with regulating cellular proliferation [128,129,132–134], apoptosis [134–136], and migration/invasion [137–140] in various cancer models. Further, ZBTB33 has been shown to be overexpressed in aggressive forms of breast and prostate cancer tumors [137,139,141–143]. In several cancers, down-regulation of ZBTB4 correlates with disease advancement [59,144,145] through several mechanisms, including supporting p53-dependent anti-apoptotic activity [59], up-regulating oncogenic proteins [144–146], and increasing genomic instability [147]. The evaluation of ZBTB38 in cancer has been limited, though several studies have correlated ZBTB38 transcriptional activities with disease relevant pathways, such as apoptosis [122,148], proliferation [149], and differentiation [123,125,150].

Nevertheless, the analysis of publicly available chromatin immunoprecipitation sequencing (ChIP-seq) data for ZBTB33 determined that it predominantly localizes to a palindromic CpG containing sequence (5'-TCTCGCGAGA-3') in the non-methylated state [151]. Moreover, no localization for ZBTB33 at any KBS loci was observed, though this sequence motif is highly represented within the human genome, and ZBTB33 binding at this site is not dependent on modification status. These ChIP-seq findings are in contrast to the mode of DNA recognition delineated from structural investigations [97,98,106], as well as the expanding body of literature that demonstrates by ChIP-quantitative PCR (ChIP-qPCR) or -semiquantitative PCR (ChIP-sqPCR) that in cell, ZBTB33 directly localizes to promoters through KBS and/or selective mCpG binding and regulates transcription [61,128–130,134–137,139,143,152]. Notably, most of these ChIP-based studies utilized the same antibody as the ChIP-seq experiments. It has been suggested that the lack of ZBTB33 occupation at mCpG sites from the ChIP-seq data may indicate that the preferred methylated DNA target sites are unavailable due to being present in condensed heterochromatin regions [151,153]. While this may be a possibility for some gene loci depending on the cellular context, there is a notable loss of occupation at known and characterized ZBTB33 gene targets observed across several cell lines. Further, ZBTB33 also functions to directly up-regulate genes [129,130], which would be occurring within open euchromatin regions. An alternative explanation for the ChIP-seq data deviating from the other ChIP-based in cell studies is that this may be a consequence of technique-specific issues during ChIP-seq sample preparation. Indeed, several challenges in sample and library preparation for this widely used method, particularly for capturing occupation events for low abundance transcription factors, have been identified [154–156]. More recently, advanced methods for preparing ChIP-seq samples have become available (for example [157–160]) that may be able to clarify the global in cell genomic occupancies for this ZF MBP.

3.2. Zfp57

Genomic imprinting is a process by which gene expression from only the maternal or paternal allele is dictated, proper management of which is essential for normal development. In most instances, this monoallelic expression is directed by regulating DNA methylation patterns at CpG clusters termed imprinting control regions (ICRs) [161]. As such, it would seem reasonable that

proteins that have evolved to have mCpG binding selectivity would function to mediate proper maintenance of this essential cellular process. Zfp57, a member of the KRAB domain family of ZF proteins, has increased expression levels during early embryonic development that tapers off upon embryonic stem cell differentiation [162]. During embryogenesis, ZFP57 has been shown to be required for DNA methylation maintenance at ICRs during both maternal and paternal DNA imprinting [163,164]. Indeed, the loss of Zfp57 in embryonic stem (ES) cells results in hypomethylation of a number of imprinted gene regions, which cannot be re-established by re-introducing Zfp57 into this null background [164,165]. These findings are consistent with a series of ChIP-based experiments that showed Zfp57 preferentially binds to a CpG-containing hexanucleotide DNA sequence localized within ICRs in a methyl-dependent manner to facilitate proper epigenetic maintenance of chromatin during early embryonic development [164]. The fact that in vitro Zfp57 has been shown to have asymmetric DNA binding [99] suggests that it may be capable of binding hemi-methylated DNA within the cell to mechanistically maintain proper DNA methylation patterns at ICRs. Consistent with this idea, the localization of Zfp57 to its target methylated sites within ICRs induces an association with TRIM28 (also known as KAP1), which leads to further recruitment of DNMTs and UHRF1; all of which serves to maintain methylation during the imprinting process [161,164]. More recently, it has been demonstrated that cancer cells can take advantage of the ES cell function of Zfp57 to promote anchorage-independent growth, suggesting that this protein may have oncogenic potential when inappropriately overexpressed [166].

3.3. *Klf4*

Klf4 is a member of the specificity protein/Krüppel-like factor (SP/Klf) family of TFs [167] that is expressed in several tissues and functions in a number of core cellular processes. Importantly, Klf4 has also been shown to be one of four essential factors required for the induction of pluripotent stem cells [168]. Several reviews have summarized the tissue specific roles of Klf4 in development, normal cell function, and disease [169–173]. Klf4 has been shown in vitro to have comparable binding activity for recognizing DNA sequences harboring CpG, mCpG, or TpG steps [100,101]. It may be that this flexibility for motif recognition expands the mechanisms by which Klf4 is capable of functioning as a mediator of key developmental process [108]. Further, from ChIP-seq analyses in conjunction with bisulphite sequencing, it has been demonstrated that the moderate sensitivity of Klf4 for mCpG relative to CpG sites affords it the ability to localize at target loci regardless of methylation status [81,174]. However, it may be that preference for methylated loci is correlated with the cell context specific functions of this protein. Indeed, in mouse embryonic stem cells, Klf4 was found to localize to a moderate proportion of methylated DNA sites relative to non-methylated [81], while in glioblastoma cells, the occupancy at highly methylated sites was ~60% [174]. In contrast to the classical model of MBP TF binding at methylated promoters leading to repression, direct Klf4 occupation at a subset of its methylated target gene loci has recently been shown to recruit chromatin remodeling complexes that signal for the opening of chromatin and subsequent up-regulation of gene expression [174].

3.4. *WT1* and *Egr1*

WT1 and Egr1 represent two ZF proteins that have highly divergent cellular functions. Egr1 cellular activities are induced within the cell in response to stress and various external stimuli [175]. In this capacity, Egr1 functions to regulate essential processes in both the nervous and cardiovascular systems [176–179]. WT1 is a transcription factor that plays a significant role in ensuring proper embryonic development of the genitourinary system [180–182]. WT1 can exist in a multitude of isoforms, and truncations or mutations of this essential protein have been correlated with several disease conditions [183,184]. Nevertheless, these two TFs have sequence conservation within three of their ZFs (ZFs 2–4 in WT1 and ZFs 1–3 in Egr1) that affords them the capability of binding to the same DNA consensus motif, 5'-GCG(T/G)GGGCG-3'. With two CpG sites, there is a significant potential for various forms of cytosine modifications to be present in this sequence within the cell at any given time.

This shared sequence is highly represented within the human genome [79], particularly at CGIs where Egr1 has been shown to in cell occupy utilizing ChIP-based studies [185,186]. It has generally been asserted that Egr1 functional target sites reside in promoter CGIs, which are typically non-methylated [153,186]. Thus, the relative indifference of Egr1 to occupy CpG relative to mCpG versions of its sequence motif would potentially pose an issue for ensuring that this protein is properly directed to its regulatory genomic targets. As such, mechanisms within the cell would be required to ensure that within the relatively short lifetime of the protein, required Egr1 transcriptional activities are not misappropriated by recruitment to non-functional methylated targets [79]. This is particularly important as it has been demonstrated that the combined number of functional and non-functional sites exceeds the number of typically expressed Egr1 protein molecules, and that a significant presence of ‘decoy’ sites regardless of methylation status kinetically impedes the ability of this protein to localize at relevant functional genomic elements [78,187]. It has recently been demonstrated in vitro that when present at sufficiently high concentrations, competitive occupation by a member of the MBD family of MBPs at methylated versions of the Egr1 DNA consensus serves to simultaneously protect Egr1 from binding at potential ‘decoy’ sites and kinetically drives Egr1 toward non-methylated target CGI regions [78]. WT1 has also been confirmed to localize to the shared Egr1 consensus motif in cell by ChIP-seq [188,189], however, the dependency of these occupation events on methylation status remain to be clarified.

It has been further demonstrated that WT1 and Egr1 have differential sensitivity to the oxidative derivatives of mC, including hydroxymethylC (hmC), formylC (fC), and carboxylC (caC) [102], which are generated in the cell through a series of oxidation reactions catalyzed by the ten-eleven translocation (TET) enzymes [190–193]. Mechanistically, this sensitivity difference appears to correlate with the presence or substitution of the core glutamate residue implicated in ZF MBP mC recognition. Specifically, both Egr1 ZF1 and ZF3 possess the key glutamate and arginine residues necessary for providing methyl sensitivity at the 5'- and 3'-GCG, respectively [102]. While ZF4 of WT1 also has the prerequisite arginine and glutamate residues, the glutamate position in ZF2 is substituted with a glutamine. This seemingly conservative alteration affords WT1 the ability to retain some binding capability for sequences harboring hmC, fC, and, in particular, caC, whereas the presence of these oxidative derivatives comparatively abolished binding for Egr1 [102]. Substitution of caC at only the 3'-GCG site resulted in an increased specificity for caC over mC by ZF2 of WT1, seemingly due to the ability of the glutamine residue to form a classical hydrogen bond with the caC carboxylate oxygen [102]. Intriguingly, it has been shown that in certain contexts, WT1 and Egr1 can function antagonistically [194]. It is interesting to hypothesize that this antagonistic behavior under certain cellular conditions may be in part correlated with the different sensitivities of these two proteins to the various forms of cytosine modifications.

3.5. CTCF and Interplay with Other ZF MBPs

CTCF is a highly expressed, well characterized ZF protein that binds to its specific core sequence through a tandem array of 11 Cys₂His₂ ZFs to elicit a number of cellular functions. Of central importance, CTCF is a key regulator of chromatin architecture that subsequently controls gene expression patterns. Recent high-resolution ChIP-exo experiments have refined the CTCF core motif to a 15-base pair sequence [195–197] that has the potential to acquire methylation on cytosines localized at positions 2 and 12 [198]. In cell, CTCF preferentially localizes to hypo-methylated regions, which serves to protect neighboring regions from acquiring DNA methylation. As an example, CTCF has been shown to protect the retinoblastoma (*Rb*) gene from epigenetic silencing [199]. Intriguingly, hyper-methylation of this promoter leads to a loss of CTCF binding and alternatively recruits ZBTB33, which induces methyl-dependent silencing of *Rb* [199]. While there is evidence to suggest that CTCF and ZBTB33 may compete for binding sites depending on the methylation status, these two TFs have also been shown to directly interact at the 5' β -globin insulator region, which contains a KBS region proximal to the CTCF core binding site [200].

Recently, structural investigations have afforded insight for the impact of cytosine methylation at positions 2 and 12 in the CTCF core sequence [104]. As discussed above, CTCF ZF3 and ZF4 can

contribute the key arginine and glutamate residues necessary to accommodate mCpG binding, and bind proximal to the cytosine at position 12. However, similar to WT1, Klf4, and Egr1, methylation of this site only marginally improved CTCF recognition of its core sequence. In contrast, methylation at the position 2 cytosine significantly impacted the ability of CTCF to bind to its core sequence. Inspection of the residues in ZF7, which is centered at cytosine 2, shows that the key arginine is present, but the glutamate residue is replaced with an aspartate. There is increasing evidence to suggest that while glutamate can provide sensitivity to mC, an aspartate substitution within ZFs confers selectivity for unmodified cytosine [100,107,108,201,202]. This is likely due to the inability of the shorter aspartate side chain to adopt an alternative conformation, which leads to a steric interference with the mC methyl group. Also analogous to Zfp57, Klf4, WT1, and Egr1, CTCF binding at the second cytosine position is asymmetric. This has interesting implications in that the genomic region of the H19 locus that encompasses the second cytosine position overlaps with the Zfp57 consensus site. The methyl sensitivity of CTCF and Zfp57 occurs on opposite strands within this region, suggesting that the status of cytosine modification could mediate binding interactions for these two TFs at this genomic site [104,108]. The cellular consequence of this disparate sensitivity to methylation at the two cytosine positions remains to be fully evaluated, but could offer an additional layer of epigenetic control for directing CTCF activities.

4. Concluding Remarks

In recent years, several structures of ZF MBPs in complex with their methylated DNA targets have become available, affording mechanistic insight into the commonalities and differences in the mode of mCpG recognition between family members. Of particular note, all of the structures to date show the common usage of a mC-Arg-G triad in mCpG recognition as well as a key glutamate that makes specific contacts with the mC. However, it is becoming evident that the number and nature of these contacts between the glutamate and mC dictates the extent for which these proteins are capable of distinguishing between mCpG relative to CpG sites. Further, substitution of the glutamate residue appears to afford increased sensitivity of ZF TFs for either unmodified or various oxidized variants of mC. In addition, recent findings from the C-terminal ZFs of ZBTB38 have demonstrated that a lysine residue can act as a surrogate for the key arginine residue, suggesting that alternative modes of mCpG recognition by ZF proteins are available. Indeed, it is important to reiterate that a number of additional ZF proteins have been identified from recent high-throughput studies to potentially have preferential mCpG binding capacity. Characterization of these additional ZF proteins may consequently illuminate additional mechanisms for mCpG recognition by this common protein scaffold. Finally, advances for in cell high-throughput sequencing strategies have begun to afford a more detailed understanding for how mCpG sensitivity correlates with the intricate biological functions elicited by the ZF MBPs in both normal cell function as well as various disease contexts.

Author Contributions: Conceptualization, N.O.H. and B.A.B.-K.; Writing-Original Draft Preparation, N.O.H. and B.A.B.-K.; Writing-Review & Editing, N.O.H. and B.A.B.-K.; Visualization, N.O.H. and B.A.B.-K.; Funding Acquisition, B.A.B.-K.

Funding: This research was funded by the National Science Foundation grant number MCB-1715379.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

References

1. Kim, G.D.; Kelesoglu, N.; Roberts, R.J.; Pradhan, S. Co-operation and communication between the human maintenance and de novo DNA (cytosine-5) methyltransferases. *EMBO J.* **2002**, *21*, 4183–4195.
2. Okano, M.; Bell, D.W.; Haber, D.A.; Li, E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* **1999**, *99*, 247–257.
3. Goll, M.G.; Bestor, T.H. Eukaryotic cytosine methyltransferases. *Annu. Rev. Biochem.* **2005**, *74*, 481–514.

4. Kareta, M.S.; Botello, Z.M.; Ennis, J.J.; Chou, C.; Chédin, F. Reconstitution of the stimulation of *de novo* methylation by human DNMT3L. *J. Biol. Chem.* **2006**, *281*, 25893–25902.
5. Sandelin, A.; Carninci, P.; Lenhard, B.; Ponjavic, J.; Hayashizaki, Y.; Hume, D.A. Mammalian RNA polymerase II core promoters: Insights from genome-wide studies. *Nat. Rev. Genet.* **2007**, *8*, 424–436.
6. Bird, A. DNA methylation patterns and epigenetic memory. *Genes Dev.* **2002**, *16*, 6–21.
7. Augui, S.; Nora, E.P.; Heard, E. Regulation of X-chromosome inactivation by the X-inactivation centre. *Nat. Rev. Genet.* **2011**, *12*, 429–442.
8. Bartolomei, M.S.; Ferguson-Smith, A.C. Mammalian genomic imprinting. *Cold Spring Harbor Perspect. Biol.* **2011**, *3*, a002592.
9. Robertson, K.D.; Wolffe, A.P. DNA methylation in health and disease. *Nat. Rev. Genet.* **2000**, *1*, 11–19.
10. Robertson, K.D., DNA methylation and human disease. *Nat. Rev. Genet.* **2005**, *6*, 587–610.
11. Kass, S.U.; Pruss, D.; Wolffe, A.P. How does DNA methylation repress transcription? *Trends Genet.* **1997**, *13*, 444–449.
12. Sharma, S.; Kelly, T.K.; Jones, P.A. Epigenetics in cancer. *Carcinogenesis* **2009**, *31*, 27–36.
13. Baylin, S.B.; Jones, P.A., A decade of exploring the cancer epigenome—Biological and translational implications. *Nat. Rev. Cancer* **2011**, *11*, 726–734.
14. Esteller, M. Aberrant DNA methylation as a cancer-inducing mechanism. *Annu. Rev. Pharmacol. Toxicol.* **2005**, *45*, 629–656.
15. Esteller, M. Epigenetics in cancer. *N. Engl. J. Med.* **2008**, *358*, 1148–59.
16. Kulis, M.; Esteller, M. DNA methylation and cancer. *Adv. Genet.* **2010**, *70*, 27–56.
17. Jones, P.A.; Baylin, S.B., The epigenomics of cancer. *Cell* **2007**, *128*, 683–692.
18. Feinberg, A.P.; Vogelstein, B. Hypomethylation distinguishes genes of some human cancers from their normal counterparts. *Nature* **1983**, *301*, 89–92.
19. Costello, J.F.; Fruhwald, M.C.; Smiraglia, D.J.; Rush, L.J.; Robertson, G.P.; Gao, X.; Wright, F.A.; Feramisco, J.D.; Peltomäki, P.; Lang, J.C.; et al. Aberrant CpG-island methylation has non-random and tumour-type-specific patterns. *Nat. Genet.* **2000**, *24*, 132–138.
20. Mazon, T.; Pankov, A.; Song, J.S.; Costello, J.F. Intratumoral heterogeneity of the epigenome. *Cancer Cell* **2016**, *29*, 440–451.
21. Massie, C.E.; Mills, I.G.; Lynch, A.G. The importance of DNA methylation in prostate cancer development. *J. Steroid Biochem. and Mol. Biol.* **2017**, *166*, 1–15.
22. Lv, J.; Liu, H.; Su, J.; Wu, X.; Liu, H.; Li, B.; Xiao, X.; Wang, F.; Wu, Q.; Zhang, Y. DiseaseMeth: A human disease methylation database. *Nucleic Acids Res.* **2011**, *40*, D1030–D1035.
23. Xiong, Y.; Wei, Y.; Gu, Y.; Zhang, S.; Lyu, J.; Zhang, B.; Chen, C.; Zhu, J.; Wang, Y.; Liu, H.; Zhang, Y. DiseaseMeth version 2.0 a major expansion and update of the human disease methylation database. *Nucleic Acids Res.* **2017**, *45*, D888–D895.
24. Jin, Z.; Liu, Y. DNA methylation in human diseases. *Genes Dev.* **2018**, *5*, 1–8.
25. Davegårdh, C.; García-Calzón, S.; Bacos, K.; Ling, C. DNA methylation in the pathogenesis of type 2 diabetes in humans. *Mol. Metab.* **2018**, *14*, 12–25.
26. de Mello, V.D.; Pulkkinen, L.; Lalli, M.; Kolehmainen, M.; Pihlajamäki, J.; Uusitupa, M. DNA methylation in obesity and type 2 diabetes. *Ann. Med.* **2014**, *46*, 103–113.
27. Drummond, E.M.; Gibney, E.R. Epigenetic regulation in obesity. *Curr. Opin. Clin. Nutr. Metab. Care* **2013**, *16*, 392–397.
28. Sayols-Baixeras, S.; Subirana, I.; Fernández-Sanlés, A.; Sentí, M.; Lluís-Ganella, C.; Marrugat, J.; Elosua, R. DNA methylation and obesity traits: An epigenome-wide association study. The REGICOR study. *Epigenetics* **2017**, *12*, 909–916.
29. Hewagama, A.; Richardson, B. The genetics and epigenetics of autoimmune disorders. *J. Autoimmun.* **2009**, *33*, 3–11.
30. Quintero-Ronderos, P.; Montoya-Ortiz, G. Epigenetics and Autoimmune Diseases. *Autoimmune Diseases* **2012**, *2012*, 1–16.
31. Sun, B.; Hu, L.; Luo, Z.Y.; Chen, X.P.; Zhou, H.H.; Zhang, W. DNA methylation perspectives in the pathogenesis of autoimmune diseases. *Clin. Immunol.* **2016**, *164*, 21–27.
32. Lu, H.; Liu, X.; Deng, Y.; Qing, H. DNA methylation, a hand behind neurodegenerative diseases. *Front. Aging Neurosci.* **2013**, *5*, 1–16.

33. Grayson, D.R.; Guidotti, A. The Dynamics of DNA Methylation in Schizophrenia and Related Psychiatric Disorders. *Neuropsychopharmacology* **2012**, *38*, 138.
34. Mahgoub, M.; Monteggia, L.M. Epigenetics and Psychiatry. *Neurotherapeutics* **2013**, *10*, 734–741.
35. Issa, J.-P., Aging and epigenetic drift: A vicious cycle. *J. Clin. Invest.* **2014**, *124*, 24–29.
36. Razin, A.; Cedar, H. DNA methylation and gene expression. *Microbiol. Rev.* **1991**, *55*, 451–458.
37. Lazarovici, A.; Zhou, T.; Shafer, A.; Dantas Machado, A.C.; Riley, T.R.; Sandstrom, R.; Sabo, P.J.; Lu, Y.; Rohs, R.; Stamatoyanopoulos, J.A. Probing DNA shape and methylation state on a genomic scale with DNase I. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 6376–6381.
38. Rao, S.; Chiu, T.-P.; Kribelbauer, J.F.; Mann, R.S.; Bussemaker, H.J.; Rohs, R. Systematic prediction of DNA shape changes due to CpG methylation explains epigenetic effects on protein-DNA binding. *Epigenetics Chromatin* **2018**, *11*, 6.
39. Dantas Machado, A.C.; Zhou, T.; Rao, S.; Goel, P.; Rastogi, C.; Lazarovici, A.; Bussemaker, H.J.; Rohs, R. Evolving insights on how cytosine methylation affects protein-DNA binding. *Brief. Func. Genom.* **2014**, *14*, 61–73.
40. Iguchi-Ariaga, S.M.M.; Schaffner, W. CpG methylation of the cAMP-responsive enhancer/promoter sequence TGACGTCA abolishes specific factor binding as well as transcriptional activation. *Genes Dev* **1989**, *3*, 612–619.
41. Pendergast, G.C.; Ziff, E.B. Methylation-sensitive sequence-specific DNA binding by the c-Myc basic region. *Science* **1991**, *251*, 186–189.
42. Watt, F.; Molloy, P.L. Cytosine methylation prevents binding to DNA of a HeLa cell transcription factor required for optimal expression of the adenovirus major late promoter. *Genes Dev.* **1988**, *2*, 1136–1143.
43. Comb, M.; Goodman, H.M. CpG methylation inhibits proenkephalin gene expression and binding of the transcription factor AP-2. *Nucleic Acids Res.* **1990**, *18*, 3975–3982.
44. Gaston, K.; Fried, M. CpG methylation has differential effects on the binding of YY1 and ETS proteins to the bi-directional promoter of the Surf-1 and Surf-2 genes. *Nucleic Acids Res.* **1995**, *23*, 901–909.
45. Defossez, P.-A.; Stancheva, I. Biological functions of methyl-CpG-binding proteins. *Prog. Mol. Biol. Transl. Sci.* **2011**, *101*, 377–398.
46. Lopez-Serra, L.; Estellar, M. Proteins that bind methylated DNA and human cancer: reading the wrong words. *British J. Cancer* **2008**, *98*, 1881–1885.
47. Parry, L.; Clark, S.J. The roles of the methyl-CpG binding proteins in cancer. *Genes and Cancer* **2011**, *2*, 618–630.
48. Sansom, O.J.; Maddison, K.; Clarke, A.R. Mechanism of disease: Methyl-binding domain proteins as potential therapeutic targets in cancer. *Nat. Clin. Pract. Onc.* **2007**, *4*, 305–315.
49. Fournier, A.; Sasai, N.; Nakao, M.; Defossez, P.-A. The role of methyl-binding proteins in chromatin organization and epigenome maintenance. *Brief. Funt. Genom.* **2011**, *11*, 251–264.
50. Joulie, M.; Miotto, B.; Defossez, P.-A. Mammalian methyl-binding proteins: What might they do? *Bioessays* **2010**, *32*, 1025–1032.
51. Sasai, N.; Defossez, P.-A. Many paths to one goal? The proteins that recognize methylated DNA in eukaryotes. *Int. J. Dev. Biol.* **2009**, *53*, 323–334.
52. Bird, A.; Macleod, D. Reading the DNA methylation signal. **2004**, *69*, 113–118.
53. Shimbo, T.; Wade, P.A., Proteins that read DNA methylation. *Adv. Exp. Med. Biol.* **2016**, *945*, 303–320.
54. Nan, X.; Ng, H.H.; Johnson, C.A.; Laherty, C.D.; Turner, B.M.; Eisenmann, R.N.; Bird, A. Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature* **1998**, *393*, 386–389.
55. Jones, P.L.; Veenstra, G.J.; Wade, P.A.; Vermaak, D.; Kass, S.U.; Landsberger, N.; Strouboulis, J.; Wolffe, A.P. Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. *Nat. Genet.* **1998**, *19*, 187–191.
56. Zhang, Y.; Ng, H.H.; Erdjument-Bromage, H.; Tempst, P.; Bird, A.; Reinberg, D. Analysis of the NuRD subunits reveals a histone deacetylases core complex and a connection with DNA methylation. *Genes Dev.* **1999**, *13*, 1924–1935.
57. Ng, H.H.; Zhang, Y.; Hendrich, B.; Johnson, C.A.; Turner, B.M.; Erdjument-Bromage, H.; Tempst, P.; Reinberg, D.; Bird, A. MBD2 is a transcriptional repressor belonging to the MeCP1 histone deacetylase complex. *Nat. Genet.* **1999**, *23*, 58–61.

58. Sasai, N.; Matsuda, E.; Sarashina, E.; Ishada, Y.; Kawaichi, M. Identification of a novel BTB-zinc finger transcriptional repressor, CIBZ, that interacts with CtBP corepressor. *Genes to Cells* **2005**, *10*, 871–885.
59. Weber, A.; Marquardt, J.; Elzi, D.; Forster, N.; Starke, S.; Glaum, A.; Yamada, D.; Defossez, P.-A.; Delrow, J.; Eisenman, R.N.; et al. Zbtb4 represses transcription of *P21CIP1* and controls the cellular responses to p53 activation. *EMBO J.* **2008**, *27*, 1563–1574.
60. Yoon, H.G.; Chan, D.W.; Reynolds, A.B.; Qin, J.; Wong, J. N-CoR mediates DNA methylation-dependent repression through a methyl CpG binding protein Kaiso. *Mol. Cell* **2003**, *12*, 723–734.
61. Lopes, E.C.; Valls, E.; Figueroa, M.E.; Mazur, A.; Meng, F.-G.; Chiosis, G.; Laird, P.W.; Schreiber-Agus, N.; Grealley, J.M.; Prokhorchouk, E.; et al. Kaiso contributes to DNA methylation-dependent silencing of tumor suppressor genes in colon cancer cell lines. *Cancer Res.* **2008**, *68*, 7258–7263.
62. Lopez-Serra, L.; Ballestar, E.; Ropero, S.; Setien, F.; Billard, L.-M.; Fraga, M.F.; Lopez-Nieva, P.; Alaminos, M.; Guerrero, D.; Dante, R.; et al. Unmasking of epigenetically silenced candidate tumor suppressor genes by removal of methyl-CpG-binding domain proteins. *Oncogene* **2008**, *27*, 3556–3566.
63. Luo, G.; Jin, C.; Long, J.; Fu, D.; Yang, F.; Xu, J.; Yu, X.; Chen, W.; Ni, Q. RNA interference of MBD1 in BxPC-3 human pancreatic cells delivered by PLGA-polyoxamer nanoparticles. *Cancer Biol. Ther.* **2009**, *8*, 1–5.
64. Fukushima, S.; Kondo, E.; Horii, A. Methyl-CpG targeted transcriptional activation allows re-expression of tumor suppressor genes in human cancer cells. *Biochem. Biophys. Res. Comm.* **2008**, *377*, 600–605.
65. Jones, P.A.; Laird, P.W. Cancer epigenetics comes of age. *Nature Genet.* **1999**, *21*, 163–167.
66. Baylin, S.B.; Herman, J.G. DNA hypermethylation in tumorigenesis: Epigenetics joins genetics. *Trends Genet.* **2000**, *16*, 168–174.
67. Ball, M.P.; Li, J.B.; Gao, Y.; Lee, J.H.; LeProust, E.M.; Park, I.H.; Xie, B.; Daley, G.Q.; Church, G.M. Targeted and genome-scale strategies reveal gene-body methylation signatures in human cells. *Nat. Biotechnol.* **2009**, *27*, 361–368.
68. Maunakea, A.K.; Nagarajan, R.P.; Bilenky, M.; Ballinger, T.J.; D'Souza, C.; Fouse, S.D.; Johnson, B.E.; Hong, C.; Nielsen, C.; Zhao, Y.; et al. Conserved role of intragenic DNA methylation in regulating alternative promoters. *Nature* **2010**, *466*, 253–257.
69. Lister, R.; Pelizzola, M.; Dowen, R.H.; Hawkins, R.D.; Hon, G.; Tonti-Filippini, J.; Nery, J.R.; Lee, L.; Ye, Z.; Ngo, Q.M.; Edsall, L.; Antosiewicz-Bourget, J.; Stewart, R.; Ruotti, V.; Millar, A.H.; Thomson, J.A.; Ren, B.; Ecker, J.R. Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* **2009**, *462*, 315–322.
70. Jjingo, D.; Conley, A.B.; Yi, S.V.; Lunyak, V.V.; Jordan, I.K. On the presence and role of human gene-body DNA methylation. *Oncotarget* **2012**, *3*, 462–474.
71. Jones, P.A. Functions of DNA methylation: Islands, start sites, gene bodies and beyond. *Nat. Rev. Genet.* **2012**, *13*, 484–492.
72. Ndlovu, M.N.; Denis, H.; Fuks, F. Exposing the DNA methylome iceberg. *Trends Biochem. Sci.* **2011**, *36*, 381–387.
73. Sarda, S.; Das, A.; Vinson, C.; Hannenhalli, S. Distal CpG islands can serve as alternative promoters to transcribe genes with silenced proximal promoters. *Gen. Res.* **2017**, *27*, 1–14.
74. Sarda, S.; Hannenhalli, S. Orphan CpG islands as alternative promoters. *Transcription* **2018**, *9*, 171–176.
75. Maor, G.L.; Yearim, A.; Ast, G. The alternative role of DNA methylation in splicing regulation. *Trends Genet.* **2015**, *31*, 274–280.
76. Nguyen, C.T.; Gonzales, F.A.; Jones, P.A. Altered chromatin structure associated with methylation-induced gene silencing in cancer cells: Correlation of accessibility, methylation, MeCP2 binding and acetylation. *Nucleic Acids Res.* **2001**, *29*, 4598–4606.
77. Baubec, T.; Ivanek, R.; Lienert, F.; Schubeler, D. Methylation-dependent and -independent genomic targeting principles of the MBD protein family. *Cell* **2013**, *153*, 480–492.
78. Kemme, C.A.; Marquez, R.; Luu, R.H.; Iwahara, J. Potential role of DNA methylation as a facilitator of target search processes for transcription factors through interplay with methyl-CpG-binding proteins. *Nucleic Acids Res.* **2017**, *45*, 7751–7759.
79. Kemme, C.A.; Nguyen, D.; Chattopadhyay, A.; Iwahara, J. Regulation of transcription factors via natural decoys in genomic DNA. *Transcription* **2016**, *7*, 115–120.
80. Hu, S.; Wan, J.; Su, Y.; Song, Q.; Zeng, Y.; Nguyen, H.N.; Shin, J.; Cox, E.; Rho, H.S.; Woodard, C.; et al. DNA methylation presents distinct binding sites for human transcription factors. *eLife* **2013**, *2*, e00726.

81. Spruijt, C.G.; Gnerlich, F.; Smits, A.H.; Pfaffeneder, T.; Jansen, P.W.; Bauer, C.; Münzel, M.; Wagner, M.; Müller, M.; Khan, F.; et al. Dynamic readers for 5-(hydroxy)methylcytosine and its oxidized derivatives. *Cell* **2013**, *152*, 1146–1159.
82. Yin, Y.; Morgunova, E.; Jolma, A.; Kaasinen, E.; Sahu, B.; Khund-Sayeed, S.; Das, P.K.; Kivioja, T.; Dave, K.; Zhong, F.; et al. Impact of cytosine methylation on DNA binding specificities of human transcription factors. *Science* **2017**, *356*, eaaj2239.
83. Zhu, H.; Wang, G.; Qian, J. Transcription factors as readers and effectors of DNA methylation. *Nat. Rev. Genet.* **2016**, *17*, 551–565.
84. Hong, S.; Wang, D.; Horton, J.R.; Zhang, X.; Speck, S.H.; Blumenthal, R.M.; Cheng, X. Methyl-dependent and spatial-specific DNA recognition by the orthologous transcription factors human AP-1 and Epstein-Barr virus Zta. *Nucleic Acids Res.* **2017**, *45*, 2503–2515.
85. Hendrich, B.; Bird, A. Identification of a family of methyl-CpG binding proteins. *Mol. Cell. Biol.* **1998**, *18*, 6538–6547.
86. Meehan, R.R.; Lewis, J.D.; McKay, S.; Kleiner, E.L.; Bird, A.P. Identification of a mammalian protein that binds specifically to DNA containing methylated CpGs. *Cell* **1989**, *58*, 499–507.
87. Bogdanovic, O.; Veenstra, G.J.C. DNA methylation and methyl-CpG binding proteins: Developmental requirements and function. *Chromosoma* **2009**, *118*, 549–565.
88. Du, Q.; Luu, P.-L.; Stirzaker, C.; Clark, S.J. Methyl-CpG-binding domain proteins: Readers of the epigenome. *Epigenomics* **2015**, *7*, 1051–1073.
89. Fatemi, M.; Wade, P.A. MBD family proteins: Reading the epigenetic code. *J. Cell Sci.* **2006**, *119*, 3033–3037.
90. Ginder, G.D.; Williams, D.C.J. Readers of DNA methylation, the MBD family as potential therapeutic targets. *Pharmacol. Ther.* **2018**, *184*, 98–111.
91. Hendrich, B.; Tweedie, S. The methyl-CpG binding domain and the evolving role of DNA methylation in animals. *Trends Genet.* **2003**, *19*, 269–277.
92. Wade, P.A. Methyl CpG-binding proteins and transcriptional repression. *BioEssays* **2001**, *23*, 1131–1137.
93. Wood, K.H.; Zhou, Z., Emerging molecular and biological functions of MBD2, a reader of DNA methylation. *Front. Genet.* **2016**, *7*, 93.
94. Dhasarathy, A.; Wade, P.A. The MBD protein family—reading and epigenetic mark? *Mutat. Res.* **2008**, *647*, 39–43.
95. Klug, A. The discovery of zinc fingers and their applications in gene regulation and genome manipulation. *Annu. Rev. Biochem.* **2010**, *79*, 213–231.
96. Wolfe, S.A.; Neklodova, L.; Pabo, C.O. DNA recognition by Cys₂His₂ zinc finger proteins. *Annu. Rev. Biophys. Biomol. Struct.* **1999**, *3*, 183–212.
97. Buck-Koehntop, B.A.; Stanfield, R.L.; Ekiert, D.C.; Martinez-Yamout, M.A.; Dyson, H.J.; Wilson, I.A.; Wright, P.E. Molecular Basis for recognition of methylated and specific DNA sequences by the zinc finger protein Kaiso. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 15229–15234.
98. Nikolova, E.N.; Stanfield, R.L.; Dyson, H.J.; Wright, P.E. CH–O hydrogen bonds mediate highly specific recognition of methylated CpG sites by the zinc finger protein Kaiso. *Biochemistry* **2018**, *57*, 2109–2120.
99. Liu, Y.; Toh, H.; Sasaki, H.; Zhang, X.; Cheng, X. An atomic model of Zfp57 recognition of CpG methylation within a specific DNA sequence. *Genes Dev.* **2012**, *26*, 2374–2379.
100. Hashimoto, H.; Wang, D.; Steves, A.N.; Jin, P.; Blumenthal, R.M.; Zhang, X.; Cheng, X. Distinctive Klf4 mutants determine preference for DNA methylation status. *Nucleic Acids Res.* **2016**, *44*, 10177–10185.
101. Liu, Y.; Olanrewaju, Y.O.; Zheng, Y.; Hashimoto, H.; Blumenthal, R.M.; Zhang, X.; Cheng, X. Structural Basis for Klf4 recognition of methylated DNA. *Nucleic Acids Res.* **2014**, *42*, 4859–4867.
102. Hashimoto, H.; Olanrewaju, Y.O.; Zheng, Y.; Wilson, G.G.; Zhang, X.; Cheng, X. Wilms tumor protein recognizes 5-carboxylcytosine within a specific DNA sequence. *Genes Dev.* **2014**, *28*, 2304–2313.
103. Zandarashvili, L.; White, M.A.; Esadze, A.; Iwahara, J. Structural impact of complete CpG methylation within target DNA on specific complex formation of the inducible transcription factor Egr-1. *FEBS Lett.* **2015**, *589*, 1748–1753.
104. Hashimoto, H.; Wang, D.; Horton, J.R.; Zhang, X.; Corces, V.G.; Cheng, X. Structural basis for the versatile and methylation-dependent binding of CTCF to DNA. *Mol. Cell* **2017**, *66*, 711–720.
105. Abramoff, M.D.; Magalhaes, P.J.; Ram, S.J. Image Processing with ImageJ. *Biophotonics Int.* **2004**, *11*, 36–42.
106. Buck-Koehntop, B.A.; Defossez, P.-A., On how mammalian transcription factors recognize methylated DNA. *Epigenetics* **2013**, *8*, 131–137.

107. Liu, Y.; Blumenthal, R.M.; Cheng, X. A common mode of recognition for methylated CpG. *Trends Biochem. Sci.* **2013**, *38*, 177–183.
108. Ren, R.; Horton, J.R.; Zhang, X.; Blumenthal, R.M.; Cheng, X. Detecting and interpreting DNA methylation marks. *Curr. Opin. Struct. Biol.* **2018**, *53*, 88–99.
109. Zou, X.; Ma, W.; Solov'yov, I.A.; Chipot, C.; Schulten, K. Recognition of methylated DNA through methyl-CpG binding domain proteins. *Nucleic Acids Res.* **2012**, *40*, 2747–2758.
110. Buck-Koehntop, B.A.; Martinez-Yamout, M.A.; Dyson, H.J.; Wright, P.E. Kaiso uses all three zinc fingers and adjacent sequence motifs for high affinity binding to sequence-specific and methyl-CpG DNA targets. *FEBS Lett.* **2012**, *586*, 734–739.
111. Daniel, J.M.; Spring, C.M.; Crawford, H.C.; Reynolds, A.B.; Baig, A. The p120^{cat}-binding partner Kaiso is a bi-modal DNA-binding protein that recognizes both a sequence-specific consensus and methylated CpG dinucleotides. *Nucleic Acids Res.* **2002**, *30*, 2911–2919.
112. Ho, P.S.; Quigley, G.J.; Tilton, R.F.J.; Rich, A. Hydration of methylated and nonmethylated B-DNA and Z-DNA. *J. Phys. Chem.* **1988**, *92*, 939–945.
113. Mayer-Jung, C.; Moras, D.; Timsit, Y. Hydration and regulation of methylated CpG steps in DNA. *EMBO J.* **1998**, *17*, 2709–2718.
114. Ho, K.L.; McNae, I.W.; Schmiedeberg, L.; Klose, R.J.; Bird, A.P.; Walkinshaw, M.D. MeCP2 binding to DNA depends upon hydration of methyl-CpG. *Mol. Cell* **2008**, *29*, 525–531.
115. Otani, J.; Arita, K.; Kato, T.; Kinoshita, M.; Kimura, H.; Suetake, I.; Tajima, S.; Ariyoshi, M.; Shirakawa, M. Structural basis of the versatile DNA recognition ability of the methyl-CpG binding domain of methyl-CpG binding domain protein 4. *J. Biol. Chem.* **2013**, *288*, 6351–6362.
116. Rohs, R.; Jin, X.; West, S.M.; Joshi, R.; Honig, B.; Mann, R.S. Origins of specificity in protein-DNA recognition. *Annu. Rev. Biochem.* **2010**, *79*, 233–269.
117. Rohs, R.; West, S.M.; Sosinsky, A.; Liu, P.; Mann, R.S.; Honig, B. The role of DNA shape in protein-DNA recognition. *Nature* **2009**, *461*, 1248–1254.
118. Slattery, M.; Zhou, T.; Yang, L.; Dantas Machado, A.C.; Gordan, R.; Rohs, R. Absence of a simple code: how transcription factors read the genome. *Trends Biochem. Sci.* **2014**, *39*, 381–399.
119. Lavery, R.; Moakher, M.; Maddocks, J.H.; Petkeviciute, D.; Zakrzewska, K. Conformational analysis of nucleic acids revisited: Curves+. *Nucleic Acids Res.* **2009**, *37*, 5917–5929.
120. Fillion, G.J.P.; Zhenilo, S.; Salozhin, S.; Yamada, D.; Prokhortchouk, E.; Defossez, P.-A. A family of human zinc finger proteins that bind methylated DNA and repress transcription. *Mol. Cell. Biol.* **2006**, *26*, 169–181.
121. Sasai, N.; Nakao, M.; Defossez, P.-A. Sequence-specific recognition of methylated DNA by human zinc-finger proteins. *Nucleic Acids Res.* **2010**, *38*, 5015–5022.
122. Oikawa, Y.; Matsuda, E.; Nishii, T.; Ishida, Y.; Kawaichi, M. Down-regulation of CIBZ, a novel substrate of Caspase-3, induces apoptosis. *J. Biol. Chem.* **2008**, *283*, 14242–14247.
123. Oikawa, Y.; Omon, R.; Nishii, T.; Ishida, Y.; Kawaichi, M.; Matsuda, E. The methyl-CpG-binding protein CIBZ suppresses myogenic differentiation by directly inhibiting *myogenin* expression. *Cell Research* **2011**, *21*, 1578–1590.
124. Pozner, A.; Hudson, N.O.; Trewhella, J.; Terroate, T.W.; Miller, S.A.; Buck-Koehntop, B.A. The C-terminal zinc fingers of ZBTB38 are novel selective readers of DNA methylation. *J. Mol. Biol.* **2018**, *430*, 258–271.
125. Kiefer, H.; Chatail-Hermitte, F.; Ravassard, P.; Bayard, E.; Brunet, I.; Mallet, J. ZENON, a novel POZ Kruppel-like DNA binding protein associated with differentiation and/or survival of late postmitotic neurons. *Mol. Cell. Biol.* **2005**, *25*, 1713–1729.
126. Prokhortchouk, A.; Hendrich, B.; Jorgensen, H.R., A.; Wilm, M.; Georgiev, G.; Bird, A.; Prokhortchouk, E. The p120 catenin partner Kaiso is a DNA methylation-dependent transcriptional repressor. *Genes Dev.* **2001**, *15*, 1613–1618.
127. Prokhortchouk, A.V.; Aitkhozhina, D.S.; Sablina, A.A.; Ruzov, A.S.; Prokhortchouk, E.B. Kaiso, a new protein of the BTB/POZ family, specifically binds to methylated DNA sequences. *Russ. J. Genet.* **2001**, *37*, 603–609.
128. Donaldson, N.S.; Pierre, C.C.; Antsey, M.I.; Robinson, S.C.; Weerawardane, S.M.; Daniel, J.M. Kaiso represses the cell cycle gene *cyclin D1* via sequence-specific and methyl-CpG-dependent mechanisms. *PLoS ONE* **2012**, *7*, e50398.
129. Pozner, A.; Terroate, T.W.; Buck-Koehntop, B.A. Cell specific Kaiso (ZBTB33) regulation of cell cycle through cyclin D1 and cyclin E1. *J. Biol. Chem.* **2016**, *291*, 24538–24550.

130. Rodova, M.; Kelly, K.F.; VanSuan, M.; Daniel, J.M.; Werle, M.J. Regulation of the Rapsyn promoter by Kaiso and d-Catenin. *Mol. Cell. Biol.* **2004**, *24*, 7188–7196.
131. Raghav, S.K.; Waszak, S.M.; Krier, I.; Gubelmann, C.; Isakova, A.; Mikkelsen, T.S.; Deplancke, B. Integrative genomics identifies the corepressor SMRT as a gatekeeper of adipogenesis through the transcription factors C/EBPbeta and KAISO. *Mol. Cell* **2012**, *46*, 335–350.
132. Cofre, J.; Menezes, J.R.; Pizzatti, L.; Abdelhay, E. Knock-down of Kaiso induces proliferation and blocks granulocytic differentiation in blast crisis of chronic myeloid leukemia. *Cancer Cell Int.* **2012**, *12*, 28.
133. Han, L.-H.; Su, Q.-X. Kaiso affects cell proliferation and cycle of lung cancer by inhibiting cyclin D1 transcript. *Prog. Anatom. Sci.* **2014**, *20*, 257–261.
134. Koh, D.-I.; Han, D.; Ryu, H.; Choi, W.-I.; Jeon, B.-N.; Kim, M.-K.; Kim, Y.; Kim, J.Y.; Parry, L.; Clarke, A.R.; et al. KAISO, a critical regulator of p53-mediated transcription of CDKN1A and apoptotic genes. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 15078–15083.
135. Bassey-Archibong, B.I.; Rayner, L.G.A.; Hercules, S.M.; Aarts, C.W.; Dvorkin-Gheva, A.; Bramson, J.L.; Hassell, J.A.; Daniel, J.M. Kaiso depletion attenuates the growth and survival of triple negative breast cancer cells. *Cell Death Dis.* **2017**, *8*, e2689.
136. Xue, X.; Zhang, J.; Lan, H.; Xu, Y.; Wang, H. Kaiso protects human umbilical vein endothelial cells against apoptosis by differentially regulating the expression of B-cell CLL/lymphoma 2 family members. *Sci. Rep.* **2017**, *7*, 7116.
137. Jones, J.; Wang, H.; Zhou, J.; Hardy, S.; Turner, T.; Austin, D.; He, Q.; Wells, A.; Grizzle, W.E.; Yates, C. Nuclear Kaiso indicates aggressive prostate cancers and promotes migration and invasiveness of prostate cancer cells. *Am. J. Pathol.* **2012**, *181*, 1836–1846.
138. Kwiecien, J.M.; Bassey-Archibong, B.I.; Dabrowski, W.; Rayner, L.G.A.; Lucas, A.R.; Daniel, J.M. Loss of Kaiso expression in breast cancer cells prevents intra-vascular invasion in the lung and secondary metastasis. *PLoS ONE* **2017**, *12*, e0183883.
139. Wang, H.; Liu, W.; Black, S.; Turner, O.; Daniel, J.M.; Dean-Colomb, W.; He, Q.P.; Davis, M.; Yates, C. Kaiso, a transcriptional repressor, promotes cell migration and invasion of prostate cancer cells through regulation of miR-31 expression. *Oncotarget* **2016**, *7*, 5677–5689.
140. Wang, L.; Ma, J.; Wang, X.; Peng, F.; Chen, X.; Zheng, B.; Wang, C.; Dai, Z.; Ai, J.; Zhao, S. Kaiso (ZBTB33) downregulation by miR-181a inhibits cell proliferation, invasion, and the epithelial-mesenchymal transition in glioma cells. *Cell Physiol. Biochem.* **2018**, *48*, 947–958.
141. Jones, J.; Wang, H.; Karanam, B.; Theodore, S.; Dean-Colomb, W.; Welch, D.R.; Grizzle, W.; Yates, C. Nuclear localization of Kaiso promotes the poorly differentiated phenotype and EMT in infiltrating ductal carcinomas. *Clin. Exp. Metastasis* **2014**, *31*, 497–510.
142. Vermeulen, J.F.; van de Ven, R.A.H.; Ercan, C.; van der Groep, P.; van der Wall, E.; Bult, P.; Christgen, M.; Lehmann, U.; Daniel, J.; van Diest, P.J.; et al. Nuclear Kaiso Expression is Associated with High Grade and Triple-Negative Invasive Breast Cancer. *PLoS ONE* **2012**, *7*, e37864.
143. Bassey-Archibong, B.I.; Kwiecien, J.M.; Milosavljevic, S.B.; Hallett, R.M.; Rayner, L.G.; Erb, M.J.; Crawford-Brown, C.J.; Stephenson, K.B.; Bédard, P.A.; Hassell, J.A.; Daniel, J.M. Kaiso depletion attenuates transforming growth factor- β signaling and metastatic activity of triple-negative breast cancer cells. *Oncogenesis* **2016**, *5*, e208.
144. Kim, K.; Chadalapaka, G.; Lee, S.-O.; Yamada, D.; Sastre-Garau, X.; Defossez, P.-A.; Park, Y.-Y.; Lee, J.-S.; Safe, S. Identification of oncogenic microRNA-17-92/ZBTB4/specificity protein axis in breast cancer. *Oncogene* **2011**, *31*, 1034–1044.
145. Kim, K.H.; Chadalapaka, G.; Pathi, S.S.; Jin, U.-H.; Lee, J.-S.; Park, Y.-Y.; Cho, S.-G.; Chintharlapalli, S.; Safe, S. Induction of the transcriptional repressor ZBTB4 in prostate cancer cells by drug-induced targeting of microRNA-17-92/106b-25 clusters. *Mol. Cancer Ther.* **2012**, *11*, 1852–1862.
146. Yang, W.W.; Chadalapaka, G.; Cho, S.-G.; Lee, S.; Jin, U.-H.; Jutooru, I.; Choi, K.; Leung, Y.-K.; Ho, S.-M.; Safe, S.; Kim, K. The transcriptional repressor ZBTB4 regulates EZH2 through microRNA-ZBTB4-Specificity protein signaling axis. *Neoplasia* **2014**, *16*, 1059–1069.
147. Roussel-Gervais, A.; Naciri, I.; Kirsh, O.; Kasprzyk, L.; Velasco, G.; Grillo, G.; Dubus, P.; Defossez, P.-A. Loss of the methyl-CpG-binding protein ZBTB4 alters mitotic checkpoint, increases aneuploidy and promotes tumorigenesis. *Cancer Res.* **2017**, *77*, 62–73.
148. Cai, Y.; Li, J.; Yang, S.; Li, P.; Zhang, X.; Liu, H. CIBZ, a novel BTB domain-containing protein, is involved in mouse spinal cord injury via mitochondrial pathway independent of p53 gene. *PLoS ONE* **2012**, *7*, e33156.

149. Nishii, T.; Oikawa, Y.; Ishida, Y.; Kawaichi, M.; Matsuda, E. CtBP-interacting BTB zinc finger protein (CIBZ) promotes proliferation and G1/S transition in embryonic stem cells via Nanog. *J. Biol. Chem.* **2012**, *287*, 12417–12424.
150. Kotoku, T.; Kosaka, K.; Nishio, M.; Ishida, Y.; Kawaichi, M.; Matsuda, E. CIBZ regulates mesodermal and cardiac differentiation of by suppressing T and Mesp1 expression in mouse embryonic stem cells. *Sci. Rep.* **2016**, *6*, 34188.
151. Blattler, A.; Yao, L.; Wang, Y.; Ye, Z.; Jin, V.X.; Farnham, P.J. ZBTB33 binds unmethylated regions of the genome associated with actively expressed genes. *Epigenetics Chromatin* **2013**, *6*, 13.
152. Qin, S.; Zhang, B.; Tian, W.; Gu, L.; Lu, Z.; Deng, D. Kaiso mainly locates in the nucleus *in vivo* and binds to methylated, but not hydroxymethylated DNA. *Chin. J. Cancer Res.* **2015**, *27*, 148–155.
153. Blattler, A.; Farnham, P.J. Cross-talk between Site-specific Transcription Factors and DNA Methylation States. *J. Biol. Chem.* **2013**, *288*, 34287–34294.
154. Chen, Y.; Negre, N.; Li, Q.; Mieczkowska, J.O.; Slattery, M.; Liu, T.; Zhang, Y.; Kim, T.-K.; He, H.H.; Zieba, J.; et al. Systematic evaluation of factors influencing ChIP-seq fidelity. *Nat. Methods* **2012**, *9*, 609–614.
155. Head, S.R.; Komori, H.K.; LaMere, S.A.; Whisenant, T.; Van Nieuwerburgh, F.; Salomon, D.R.; Ordoukhanian, P. Library construction of next-generation sequencing: Overviews and challenges. *BioTechniques* **2014**, *56*, 61–77.
156. Park, P.J. ChIP-seq: Advantages and challenges of a maturing technology. *Nat. Rev. Genet.* **2009**, *10*, 669–680.
157. Schmidl, C.; Renderio, A.F.; Sheffield, N.C.; Bock, C. ChIPmentation: Fast, robust, low-input ChIP-seq for histones and transcription factors. *Nat. Methods* **2015**, *12*, 963–965.
158. Wallerman, O.; Nord, H.; Bysani, M.; Borghini, L.; Wadelius, C. lobChIP: From cells to sequencing ready ChIP libraries in a single day. *Epigenetics Chromatin* **2015**, *8*, 25.
159. Skene, P.J.; Henikoff, S. A simple method for generating high-resolution maps of genome-wide protein binding. *eLife* **2015**, *4*, e09225.
160. Skene, P.J.; Henikoff, S. An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites. *eLife* **2017**, *6*, e21856.
161. Strogantsev, R.; Ferguson-Smith, A.C. Proteins involved in establishment and maintenance of imprinted methylation marks. *Brief. Func. Genom.* **2012**, *11*, 227–239.
162. Li, X.; Leder, P. Identifying genes preferentially expressed in undifferentiated embryonic stem cells. *BMC Cell Biol.* **2007**, *8*, 37.
163. Li, X.; Ito, M.; Zhou, F.; Youngson, N.; Zuo, X.; Leder, P.; Ferguson-Smith, A.C. A maternal-zygotic effect gene, Zfp57, maintains both maternal and paternal imprints. *Dev. Cell* **2008**, *15*, 547–557.
164. Quenneville, S.; Verde, G.; Corsinotti, A.; Kapopoulou, A.; Jakobsson, J.; Offner, S.; Baglivo, I.; Pedone, P.V.; Grimaldi, G.; Riccio, A.; et al. In Embryonic Stem Cells, ZFP57/KAP1 Recognize a Methylated Hexanucleotide to Affect Chromatin and DNA Methylation of Imprinting Control Regions. *Mol. Cell* **2011**, *44*, 361–372.
165. Zuo, X.; Sheng, J.; Lau, H.-T.; McDonald, C.M.; Andrade, M.; Cullen, D.E.; Bell, F.T.; Iacovino, M.; Kyba, M.; Xu, G.; et al. 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.* **2012**, *287*, 2107–2118.
166. Tada, Y.; Yamaguchi, Y.; Kinjo, T.; Song, X.; Akagi, T.; Takamura, H.; Ohta, T.; Yokota, T.; Koide, H. The stem cell transcription factor ZFP57 induces IGF2 expression to promote anchorage-independent growth in cancer cells. *Oncogene* **2015**, *34*, 752–760.
167. Shields, J.M.; Christy, R.J.; Yang, V.W. Identification and characterization of a gene encoding a gut-enriched Kruppel-like factor expressed during growth arrest. *J. Biol. Chem.* **1996**, *271*, 20009–20017.
168. Takahashi, K.; Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **2006**, *126*, 663–676.
169. Dang, D.T.; Pevsner, J.; Yang, V.W. The biology of the mammalian Krüppel-like family of transcription factors. *Int. J. Biochem. Cell. Biol.* **2000**, *32*, 1103–1121.
170. Evans, P.M.; Liu, C. Role of Krüppel-like factor 4 in normal homeostasis, cancer, and stem cells. *Acta Biochim. Biophys. Sin.* **2008**, *40*, 554–564.
171. Farrugia, M.K.; Vanderbilt, D.B.; Salkeni, M.A.; Ruppert, J.M. Kruppel-like Pluripotency Factors as Modulators of Cancer Cell Therapeutic Responses. *Cancer Res.* **2016**, *76*, 1677–1682.

172. Ghaleb, A.M.; Yang, V.W. Krüppel-like factor 4 (KLF4): What we currently know. *Gene* **2017**, *611*, 27–37.
173. Rowland, B.D.; Pepper, D.S. KLF4, p21 and context-dependent opposing forces in cancer. *Nat. Rev. Cancer* **2006**, *6*, 11–23.
174. Wan, J.; Su, Y.; Song, Q.; Tung, B.; Oyinlade, O.; Liu, S.; Ying, M.; Ming, G.-L.; Song, H.; Qian, J.; et al. Methylated cis-regulatory elements mediate KLF4-dependent gene transactivation and cell migration. *eLife* **2017**, *6*, e20068.
175. Pagel, J.L.; Deindl, E. Early growth response 1 - a transcription factor in the crossfire of signal transduction cascades. *Indian J. Biochem. Biophys.* **2011**, *48*, 226–235.
176. Bozon, B.; Davis, S.; Laroche, S. A requirement for the immediate early gene zif268 in reconsolidation of recognition memory after retrieval. *Neuron* **2003**, *40*, 695–701.
177. Khachigian, L.M.; Lindner, V.; Williams, A.J.; Collins, T. Egr-1-induced endothelial gene expression: A common theme in vascular injury. *Science* **1996**, *271*, 1427–1431.
178. Lee, J.L.; Everitt, B.J.; Thomas, K.L. Independent cellular processes for hippocampal memory consolidation and reconsolidation. *Science* **2004**, *304*, 839–843.
179. Yan, S.F.; Fujita, T.; Lu, J.; Okada, K.; Shan Zou, Y.; Mackman, N.; Pinsky, D.J.; Stern, D.M. Egr-1, a master switch coordinating upregulation of divergent gene families underlying ischemic stress. *Nat. Med.* **2000**, *6*, 1355–1361.
180. Lee, S.B.; Haber, D.A. Wilms tumor and the WT1 gene. *Exp. Cell Res.* **2001**, *264*, 74–99.
181. Niaudet, P.; Gubler, M.C. WT1 and glomerular diseases. *Pediatric. Nephrol.* **2006**, *21*, 1653–1660.
182. Wagner, K.D.; Wagner, N.; Schedl, A. The complex life of WT1. *J. Cell Sci.* **2003**, *116*, 1653–1658.
183. Hastie, N.D. Wilms' tumour 1 (WT1) in development, homeostasis and disease. *Development* **2017**, *144*, 2862–2872.
184. Hohenstein, P.; Hastie, N.D. The many facets of the Wilms' tumor gene, WT1. *Hum. Mol. Genet.* **2006**, *15*, R196–R201.
185. Koldamova, R.; Schug, J.; Lefterova, M.; Cronican, A.A.; Fitz, N.F.; Davenport, F.A.; Carter, A.; Castranio, E.L.; Lefterov, I. Genome-wide approaches reveal EGR1-controlled regulatory networks associated with neurodegeneration. *Neurobiol. Dis.* **2014**, *63*, 107–114.
186. Kubosaki, A.; Tomaru, Y.; Tagami, M.; Arner, E.; Miura, H.; Suzuki, T.; Suzuki, M.; Suzuki, H.; Hayashizaki, Y. Genome-wide investigation of in vivo EGR-1 binding sites in monocytic differentiation. *Genome Biol.* **2009**, *10*, R41.
187. Kemme, C.A.; Esadze, A.; Iwahara, J. Influence of quasi-specific sites on kinetics of target DNA search by a sequence-specific DNA-binding protein. *Biochemistry* **2015**, *54*, 6684–6691.
188. Kann, M.; Ettou, S.; Jung, Y.L.; Lenz, M.O.; Taglienti, M.E.; Park, P.J.; Schermer, B.; Benzing, T.; Kreidberg, J.A. Genome-wide analysis of Wilms' tumor 1-Controlled gene expression in podocytes reveals key regulatory mechanisms. *J. Am. Soc. Nephrol.* **2015**, *26*, 2097–2104.
189. Motamedi, F.J.; Badro, D.A.; Clarkson, M.; Rita Lecca, M.; Bradford, S.T.; Buske, F.A.; Saar, K.; Hübner, N.; Brändli, A.W.; Schedl, A. WT1 controls antagonistic FGF and BMP-pSMAD pathways in early renal progenitors. *Nat. Commun.* **2014**, *5*, 4444.
190. He, Y.-F.; Li, B.-Z.; Li, Z.; Liu, P.; Wang, Y.; Tang, Q.; Ding, J.; Jia, Y.; Chen, Z.; Li, L.; et al. Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science* **2011**, *333*, 1303–1307.
191. Ito, S.; D'Alessio, A.C.; Taranova, O.V.; Hong, K.; Sowers, L.C.; Zhang, Y. Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature* **2010**, *466*, 1129–1133.
192. Ito, S.; Shen, L.; Dai, Q.; Wu, S.C.; Collins, L.B.; Swenberg, J.A.; He, C.; Zhang, Y. Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science* **2011**, *333*, 1300–1303.
193. Tahiliani, M.; Koh, K.P.; Shen, Y.; Pastor, W.A.; Bandukwala, H.; Brudno, Y.; Agarwal, S.; Iyer, L.M.; Liu, D.R.; Aravind, L.; Rao, A. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* **2009**, *324*, 930–935.
194. Ritchie, M.F.; Yue, C.; Zhou, Y.; Houghton, P.J.; Soboloff, J. Wilms tumor suppressor 1 (WT1) and early growth response 1 (EGR1) are regulators of STIM1 expression. *J. Biol. Chem.* **2010**, *285*, 10591–10596.
195. Nakahashi, H.; Kwon, K.-R. K.; Resch, W.; Vian, L.; Dose, M.; Stavreva, D.; Hakim, O.; Pruett, N.; Nelson, S.; Yamane, A.; et al. A genome-wide map of CTCF multivalency redefines the CTCF code. *Cell Reports* **2013**, *3*, 1678–1689.
196. Rhee, H.S.; Pugh, B.J. Comprehensive genome-wide protein-DNA interactions detected at single-nucleotide resolution. *Cell* **2011**, *147*, 1408–1419.

197. Terooatea, T.W.; Pozner, A.; Buck-Koehntop, B.A. PAtChCap: Input Strategy for Improving Analysis of ChIP-exo Data and Beyond. *Nucleic Acids Res.* **2016**, *44*, e159.
198. Wang, H.; Maurano, M.T.; Qu, H.; Varley, K.E.; Gertz, J.; Pauli, F.; Lee, K.; Canfield, T.; Weaver, M.; Sanstrom, R.; et al. Widespread plasticity in CTCF occupancy linked to DNA methylation. *Genome Res.* **2012**, *22*, 1680–1688.
199. De La Rosa-Velazquez, I.A.; Rincon-Arango, H.; Benitez-Bribiesca, L.; Recillas-Targa, F. Epigenetic regulation of the human retinoblastoma tumor suppressor gene promoter by CTCF. *Cancer Res.* **2007**, *67*, 2577–2585.
200. Defossez, P.-A.; Kelly, K.F.; Filion, G.J.P.; Perez-Torrado, R.; Magdinier, F.; Menoni, H.; Nordgaard, C.L.; Daniel, J.M.; Gilson, E. The human enhancer blocker CTC-binding factor interacts with the transcription factor Kaiso. *J. Biol. Chem.* **2005**, *280*, 43017–43023.
201. Choo, Y. Recognition of DNA methylation by zinc fingers. *Nat. Struct. Biol.* **1998**, *5*, 264–265.
202. Choo, Y.; Klug, A. Physical basis of a protein-DNA recognition code. *Curr. Opin. Struct. Biol.* **1997**, *7*, 117–125.



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).