Cys_2His_2 Zinc Finger Methyl-CpG Binding Proteins: Getting a Handle on Methylated DNA

Amelia J. Hodges, Nicholas O. Hudson and Bethany A. Buck-Koehntop

Department of Chemistry, *University of Utah, 315 South 1400 East, Salt Lake City, UT, 84112, USA*

Correspondence to Bethany A. Buck-Koehntop: *Department of Chemistry, University of Utah, 315 South 1400 East, Salt Lake City, UT, 84112, USA. Fax: +1 801 581 8433. koehntop@chem.utah.edu*
https://doi.org/10.1016/j.jmb.2019.09.012
Edited by P. Defossez

Abstract

DNA methylation is an essential epigenetic modification involved in the maintenance of genomic stability, preservation of cellular identity, and regulation of the transcriptional landscape needed to maintain cellular function. In an increasing number of disease conditions, DNA methylation patterns are inappropriately distributed in a manner that supports the disease phenotype. Methyl-CpG binding proteins (MBPs) are specialized transcription factors that read and translate methylated DNA signals into recruitment of protein assemblies that can alter local chromatin architecture and transcription. MBPs thus play a key intermediary role in gene regulation for both normal and diseased cells. Here, we highlight established and potential structure function relationships for the best characterized members of the zinc finger (ZF) family of MBPs in propagating DNA methylation signals into downstream cellular responses. Current and future investigations aimed toward expanding our understanding of ZF MBP cellular roles will provide needed mechanistic insight into normal and disease state functions, as well as afford evaluation for the potential of these proteins as epigenetic-based therapeutic targets.

Introduction

DNA methylation, the addition of a methyl group at the 5-position of cytosine bases most commonly in the context of CpG dinucleotides, plays an essential role in regulating the transcriptional landscape needed to maintain normal cellular function. Specifically, DNA methylation in association with distinct patterns of histone protein posttranslational modifications (PTMs) regulates chromatin accessibility and transcriptional outcomes [1,2]. In addition to controlling gene expression, DNA methylation also plays a key role in development, maintenance of genomic stability, and preservation of cellular identity [3 6]. It is therefore unsurprising that misappropriation of genomic DNA methylation distributions has been correlated with an expanding list of disease conditions [7 10]. Nevertheless, the full mechanisms by which DNA methylation information is converted into genomic responses that support specific cellular phenotypes is not yet fully understood.

In humans, the majority of genomic CpG sites are methylated (mCpG), the exception being GC-rich stretches termed CpG islands (CGIs), found in ~70% of all gene promoters [11]. Consequently, highly methylated promoter CGIs are typically associated with transcriptional repression [12,13]. Mechanistically, DNA methylation at promoter CGIs induces gene silencing by preventing or promoting transcription factor (TF) binding [1,14]. Indeed, local perturbations to the DNA structure and increased hydrophobicity of the DNA major groove at mCpG sites [15,16] can either serve to repel [17 19] or attract TF binding. TFs that exhibit recognition for mCpG containing DNA sequences have been classified as methyl-CpG binding proteins (MBPs) [20 24]. Evidence suggests that many MBPs play critical intermediary roles in regulating gene activity by reading DNA methylation signals and recruiting protein assemblies that subsequently alter local chromatin architecture and transcriptional states [25 29].

In contrast to promoter CGIs, DNA methylation at gene body CGIs has been correlated with active

transcription [30 32], possibly by creating alternative promoters that regulate differential splicing [33 36]. Intriguingly, MBPs have also been observed to occupy methylated DNA target sites within gene bodies without significantly altering transcriptional elongation when preceded by an active promoter [37,38]. Thus, MBPs have emerged as key factors in mediating a multitude of methyl-dependent gene regulation processes within the cell. As such, there has been significant interest in defining the mechanisms by which MBPs select and interpret DNA methylation signals, as well as evaluating their potential as novel therapeutic targets [39 41].

Cys_2His_2 (C_2H_2) zinc finger (ZF) containing MBPs represent one of the three originally identified MBP classes [21,23,42]. Within this class, ZBTB33 (zinc finger and BTB domain-containing protein 33; also known as Kaiso) was identified nearly two decades ago as the first ZF-containing protein to exhibit selective methylated DNA-binding capability [43,44], which it does through a set of three C_2H_2 ZFs. Using this three ZF domain, BLAST (Basic Local Alignment Search Tool) analysis identified two additional proteins, ZBTB4 and ZBTB38, which were also confirmed to be selective readers of mCpG sites [45]. All three of these proteins belong to the larger BTB/POZ (*b*road complex, *t*ramtrack, *b*ric-á-brac, *po*x virus and *z*inc finger) TF family, and for many years constituted the entire subset of ZF-containing methylated DNA reader proteins (collectively termed the ZBTB MBPs). In recent years, several more verified ZF MBPs have been identified and characterized, including ZFP57 (zinc finger protein 57), KLF4 (Krüppel-like factor 4), WT1 (Wilms' tumor protein 1), EGR1 (growth response protein 1; also known as Zif268), and CTCF (CCCTC-binding factor) [46,47].

In this review, we discuss structure function relationships for the best characterized ZF MBP family members in reading DNA methylation and modulating cellular responses, with an emphasis on indicating areas for future investigation. We further highlight established and potential roles of these proteins in impacting both normal and diseased cellular function within the context of their ability to read mCpG sites. In addition, we summarize the mode of mCpG recognition by ZF MBPs discerned from high-resolution structural evidence. Finally, we provide future perspectives for what remains to be evaluated in the field of ZF MBP structure and function. In short, growing evidence implicates the cellular activities for many of the ZF MBPs in disease conditions, suggesting that some of these proteins may represent targets for epigenetic-based therapeutics. Though additional research for many of the ZF MBPs is still needed to fully elucidate the mechanisms by which mCpG recognition mediates biomolecular interactions, transcriptional outcomes, and cellular phenotype. Furthermore, much remains to be investigated for how the functions of the proteins are regulated within both normal and disease cell contexts.

The Zinc Finger Family of Methyl-CpG Binding Proteins

ZBTB33

Of the three ZBTB MBP family members, ZBTB33 is by far the best characterized. It is also the smallest of the three proteins, composed of an N-terminal BTB/POZ protein interaction domain and a three C_2H_2 ZF DNA binding domain separated by an intrinsically disordered linker (Fig. 1). ZBTB33 was initially identified during a search for novel protein interacting partners for the cell adhesion maintenance protein p120-catenin ($p120^{ctn}$) [48]. It has since been demonstrated that $p120^{ctn}$ binds within the ZF domain of ZBTB33, competitively alleviates DNA binding, and can impact the cellular localization of ZBTB33 [49 51]. The ZBTB33-$p120^{ctn}$ protein association has also been implicated in the ability of ZBTB33 to modulate Wnt signaling, a pathway that is misregulated in a number of cancerous conditions (see Ref. [52] for a recent review covering this topic).

Shortly after being identified as a selective reader of methylated DNA [43,44], ZBTB33 was also found to recognize a sequence-specific TpG-containing site, termed the KBS (Kaiso binding sequence; TCCTGCNA) [53]. Notably, this ability to preferentially select for both mCpG and TpG over CpG containing sites using the same ZF domain is recapitulated by ZBTB4 and ZBTB38 [45,54 56]. In terms of specific methylated DNA targets, it has been stated that ZBTB33 requires two consecutive mCpG sites, whereas both ZBTB4 and ZBTB38 only need one for selective methylated DNA recognition. High-resolution structures and supporting biophysical characterization indicate that ZBTB33 requires a 5′-mCGC-3′ or 5′-TGC-3′ core for making base-specific contacts with its consensus DNA motifs [57,58]. Although a GC-step in the 3′ position of these sequences is required for Arg-511 to hydrogen bond with the guanosine base, the methylation status of the pairing cytosine has a moderate impact on the overall binding [57,58]. Notably, the identified consensus motifs for ZBTB4 also harbor a 5′-mCGC-3′ or 5′-TGC-3′ core [56], suggesting a similar mode of DNA recognition between these two proteins. Nevertheless, the genome-wide preference for single or consecutive mCpG steps for each ZBTB MBP remains to be determined.

Furthermore, the impact of bimodal DNA recognition on the cellular transcriptional activities for each ZBTB MBP family member remains to be fully elucidated. Emerging evidence reveals bimodal

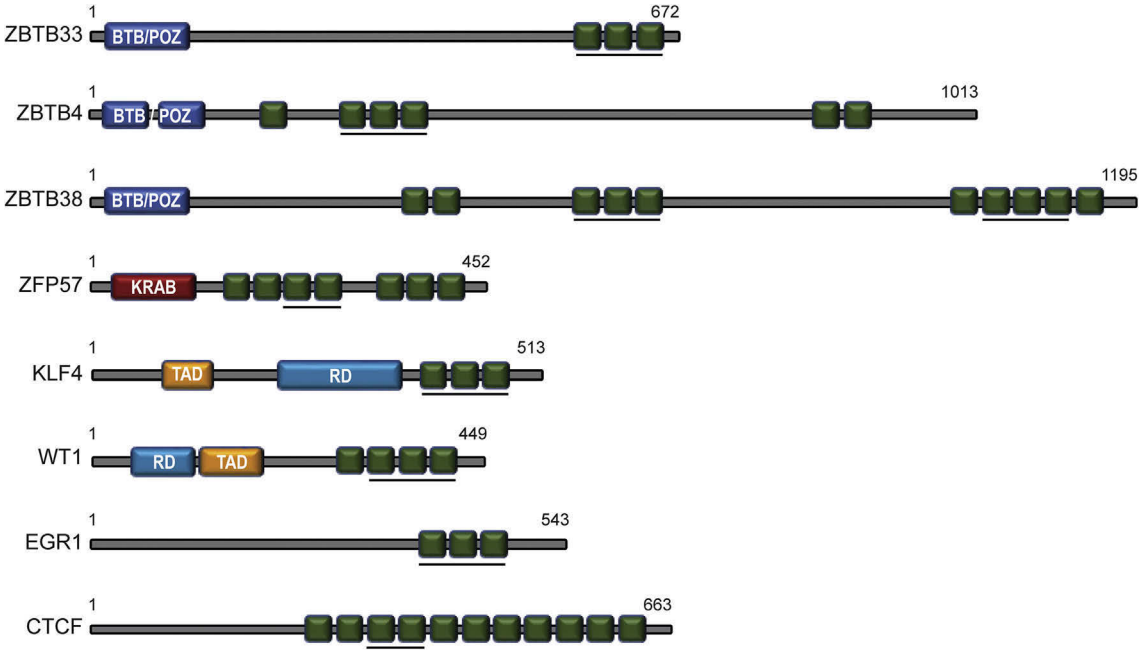

Fig. 1. Domain organization for human versions of the C_2H_2 ZF MBPs discussed in this review. For each protein, the ZFs involved in mCpG binding are underlined. Green boxes indicate individual ZFs. BTB/POZ: Broad complex, Tramtrack, Bric-á-brac/Pox virus and ZF domain; KRAB: Krüppel-associated box; TAD: Transactivation domain; RD: Repression domain.

DNA recognition by ZBTB33 may direct the activity of this protein to either activate or repress transcription depending on the consensus sequence bound, methylation status of target sites, and cellular context [59 65]. Intriguingly, it was also recently demonstrated that deSUMOylation of ZBTB33 altered it from being a transcriptional activator to a repressor at the same promoter [66]. These findings suggest that SUMOylation dynamics, likely through altering ZBTB33 protein protein interactions, may influence transcriptional outcomes at some targeted genomic loci. In addition to SUMOylation, ZBTB33 has also been shown to be phosphorylated at specific residues [52,67], though the overall impact of PTMs on mediating protein interactions and ZBTB33 biological functions remains to be discerned.

In addition to investigating DNA-binding capabilities, a few studies have evaluated ZBTB33 protein interactions beyond $p120^{ctn}$, highlighting the diverse roles of this protein in mediating cellular function. Notably, ZBTB33 has been demonstrated to directly recruit the nuclear receptor corepressor 1 (NCoR1) deacetylase co-repressor complex through its BTB/POZ domain in a DNA methylation-dependent manner, resulting in a repressive chromatin state [29]. More recently, it was determined that ZBTB33 also associates with silencing mediator of retinoic acid and thyroid hormone receptor (SMRT; also known as NCoR2) to repress transcription at target sites [68]. In addition, direct association of the C-terminal region of CTCF with the BTB/POZ domain of ZBTB33 in the context of KBS-dependent DNA recognition leads to negative regulation of CTCF insulator activity [69]. However, there is also evidence to suggest CTCF and ZBTB33 compete for binding sites depending on DNA methylation status of the target site [70]. Similarly, ZBTB33 heterodimerization with the BTB/POZ protein Znf131 results in negative regulation of Znf131 transcriptional activation [71]. Though the relationship between these two proteins may be more complex in that Znf131 has been shown to reciprocally regulate ZBTB33-mediated transcription [72]. Finally, all three ZBTB MBPs have been identified as binding partners for the corepressor protein myeloid translocation gene target 16 (MTG16) through their conserved three C_2H_2 ZF domains [73]. It was demonstrated that ZBTB33:MTG16-mediated transcriptional repression was dependent on KBS and not mCpG recognition.

Nevertheless, a complete determination of the protein interactome for ZBTB33 has not yet been completed. This information is necessary to fully elucidate the mechanisms by which ZBTB33 mediates normal and disease-state cellular functions. Indeed, it is becoming clear that a number of gene promoters harbor both mCpG and KBS containing target sites, and that ZBTB33 occupation of only one or both of these positions can alter

downstream transcriptional outcomes [60,63]. Although this can be in part regulated by methylation status of the CpG site, the evidence presented earlier demonstrates that differential ZBTB33 genomic occupation patterns also result in recruitment of specified protein complexes that facilitate the downstream response.

Although continued investigative efforts are needed to deconvolute the cellular DNA targets of ZBTB33, identify the full spectrum of protein interacting partners, and evaluate the impact of PTMs on regulating these interactions, it is becoming increasingly evident that ZBTB33-mediated transcriptional activities are correlated with a number of cancer types, including breast [74 76], chronic myeloid leukemia [77], colon [61,78], glioma [79], intestinal [80 82], lung [83,84], pancreatic [85], and prostate cancers [86 88]. Indeed, from combined studies using cancer cell line and mouse models, ZBTB33 has been implicated in regulating cellular proliferation [60,63,77,79,84,88], apoptosis [89,90], migration/invasion [79,83,87,88,91], and epithelial-mesenchymal transition (EMT) [75,79,86,92] in a manner that supports the disease condition. Furthermore, from studies in human tumor samples, two important trends for ZBTB33 have begun to emerge. First, the subcellular localization of ZBTB33 appears to vary with tumor type and can be predictive of patient outcome [75,76,83,85,87]. Second, the overexpression and subcellular localization trends for ZBTB33 can be variable in patient tumors depending on ethnicity [74 76,85,87] and gender [85]. Finally, ZBTB33 transcriptional activities have more recently been associated with other disease conditions recognized to be in part driven by disrupted DNA methylation patterns including inflammatory bowel disease [93,94] and behavioral disorders [95,96]. In sum, it is now relatively well established that ZBTB33 plays a key role in supporting several disease conditions and should be considered as a potential target for new epigenetic-based therapeutic interventions.

ZBTB4

ZBTB4 was first identified from a BLAST search seeking to identify additional proteins containing ZFs similar to those of ZBTB33 [45]. In addition to the three C_2H_2 ZFs shared with ZBTB33 (ZFs 2 4), ZBTB4 also has another N-terminal and two C-terminally localized C_2H_2 ZFs; the functions of which remain to be defined (Fig. 1). Interestingly, unlike other ZBTB family members, the BTB/POZ protein interaction domain of human ZBTB4 contains a noncanonical ~60 residue Ser/Ala insertion [45]. From comparison with high-resolution structures of other BTB/POZ protein domains [97], the linker insertion within ZBTB4 is not predicted to disrupt the overall tertiary structure or dimerization interface. However, the role of this flexible polypeptide region in regulating ZBTB4 activities and biomolecular interactions is unknown.

On identification, ZBTB4 was confirmed to also selectively recognize methylated DNA as well as the TpG-containing KBS [45] through ZFs 2 4. Further attempts to refine the sequence context for ZBTB4 TpG and mCpG consensus sites using *in vitro* methodologies led to identification of the Z4BS (ZBTB4 binding sequence; C(C/T)GCCATC) and methyl-dependent Z4BS (meZ4BS; (C/A)^{m}CGC(C/T)AT), respectively [56]. Although no structural data for ZBTB4 in complex with either its methylated DNA target or the Z4BS are currently available, the core residues required for DNA recognition in ZBTB33 [57,58] are conserved in ZBTB4. Indeed, substitution of the comparable ZBTB4 Arg-326 with alanine abolished methylated DNA-binding capability [56], suggesting a similar mode of DNA recognition. Notably, the sequence context differences between the KBS and Z4BS are minimal, though the cellular consequences of ZBTB4 and ZBTB33 reading similar TpG-containing DNA consensus motifs remains to be evaluated. Intriguingly, ZBTB4 was found to be able to heterodimerize with ZBTB38 but not ZBTB33 [45], suggesting that ZBTB4 and ZBTB33 do not coregulate target genes. Furthermore, current literature has not identified overlap in the genomic targeting of ZBTB4 and ZBTB33, although a direct comparison of global genomic occupancy and subsequent transcriptional alterations within the same cellular context is not yet available. Additional studies will be required to evaluate the extent of functional redundancy between these two proteins.

Although many more investigations are needed to define the full spectrum of ZBTB4 functions, several studies have begun to provide insight for some cellular roles of this protein. Specifically, ZBTB4 has been shown to function as a DNA methylation-dependent transcriptional repressor [28,45], which in certain contexts is facilitated by recruitment of Sin3/histone deacetylase (HDAC) corepressor complexes [28]. Furthermore, in mouse cells ZBTB4 exhibits DNA methylation-dependent localization at centromeric and pericentromeric repeats [27], where it may play a role in regulating chromosomal stability. Indeed, ZBTB4 depletion in human cancer cell lines and mouse models resulted in an overall instability of the genome characterized by increased miotic abnormalities and incidences of lagging chromosomes [98]. Consistent with these findings, a decrease in ZBTB4 levels either through mRNA targeting by miRNAs [99 102], genomic deletion [103], or protein degradation following phosphorylation by human protein kinase 2 (HIPK2) [104] have been correlated with a number of cancer types [28,98,100 102,105 107]. Furthermore, elevated ZBTB4 mRNA levels coincided with improved

patient relapse-free survival outcomes in breast and prostate cancers [101,102].

Based on these observations, it is unsurprising that ZBTB4 depletion has been found to provide a mechanistic survival advantage for cancerous cells. Specifically, ZBTB4 depletion has been correlated with increased proliferation [100,107], cell survival [28,107], and migration/invasion [100] in cancer cells. These observed cancer-promoting phenotypes have been in part correlated with direct transcriptional repression by ZBTB4 for several members of the oncogenic specificity protein (Sp) TF family [101]. In addition to directly repressing Sp1, ZBTB4 and Sp1 compete for binding at GC-rich sites within the promoters of protumorigenic genes [101,102] and chromatin remodelers [106]. Combined, these findings imply that many cancer cells selectively reduce ZBTB4 levels to amplify oncogenic Sp TF activities. Consequently, over the last several years the interplay between ZBTB4 and Sp TFs has been evaluated for therapeutic feasibility. Indeed, a number of anticancer agents have been demonstrated to simultaneously increase Sp-targeting miRNAs while decreasing ZBTB4-targeting miRNAs, significantly impacting the survival potential of targeted cancer cells [108 110]. In sum, current literature evidence suggests that ZBTB4 primarily functions as a tumor suppressor; however, more studies are required to fully define the cellular functions of this protein. Furthermore, the genome-wide evaluation of ZBTB4 DNA targets, protein interacting partners, and roles of PTMs on mediating ZBTB4 cellular activities remains to be fully characterized.

ZBTB38

ZBTB38 was first identified by two independent laboratories due to its ability to recognize the TpG-containing E-box dyad within the tyrosine hydroxylase promoter using its N-terminal ZFs [54], as well as its interaction with the corepressor C-terminal binding protein (CtBP) [27]. Among the three ZBTB MBPs, ZBTB38 is the largest and most structurally complex. In addition to its BTB/POZ protein interaction domain, ZBTB38 has two distinct sets of five ZF clusters that are separated by an intrinsically disordered linker region (Fig. 1). Of the ten ZFs, all are C_2H_2 with the exception of ZF 2, which is a predicted C_2HC type. Similar to ZBTB33, early investigations showed that the shared N-terminal ZBTB38 ZFs (ZFs 3 5) participate in bimodal DNA recognition, whereby this ZF set can specifically bind to either mCpG or TpG-containing DNA targets [45,54,55]. It was further shown that the C-terminal set of ZBTB38 ZFs (ZFs 6 10) were incapable of specifically binding DNA consensus sites belonging to the N-terminal ZFs [54,55]. Consequently, the functional role for the C-terminal ZF domain was for many years undefined.

Recently, it was discovered that a subset of the C-terminal ZBTB38 ZFs also have preferential selectivity for a methylated DNA consensus motif *in vitro* and within cells, termed the mCZ38BS (methylated C-terminal ZBTB38 binding sequence; (A/G)T^{m}CG(G/A)(^{m}C/T)(G/A)) [111]. Specifically, it was determined that ZFs 7 and 8 are necessary and sufficient for methyl-selective DNA recognition, while ZFs 6 and 9 are required to stabilize the binding interaction [111,112]. ZF 10 does not appear to contribute significantly to DNA binding [111 113]. ZBTB38 is thus unique in that it is the only MBP identified to date that can selectively recognize two different mCpG sites with two distinct ZF DNA-binding domains, though the extent to which these two protein regions interrelate in mediating epigenetic-based transcription remains to be fully deconvoluted. Indeed, preliminary investigation of the interplay between the N- and C-terminal ZBTB38 ZF domains in mediating cellular transcription suggests this process is complex, gene context specific, and warrants further investigation [111].

In addition to DNA binding investigations, several recent reports have expanded our knowledge of ZBTB38 protein interactions. Consistent with its role as an epigenetic-based transcriptional regulator, tandem affinity purification of HA-FLAG-ZBTB38 from HeLa cells followed by mass spectrometry analysis identified a number of ZBTB38 protein interacting partners that include TFs, chromatin remodelers, and histone modifying enzymes [114]. Additionally, as discussed earlier, ZBTB38 was identified as a binding partner for CtBP [27], a component of a corepressor complex known to also include HDACs, histone methyltransferases (HMTs), and histone demethylases (HDMs) [115]. Additionally, the impact of ubiquitination/ubiquitination-like pathways on regulating ZBTB38 cellular abundance has been investigated. Notably, the E3 ubiquitin ligase retinoblastoma binding protein 6 (RBBP6) and deubiquitinase ubiquitin-specific protease 9X (USP9X) function antagonistically to control ZBTB38 cellular protein levels [114,116,117]. Furthermore, in response to cellular infection by herpes simplex virus type 1 (HSV-1) ZBTB38 was identified to be selectively SUMOylated by SUMO-2/3, which targeted it for degradation by the HSV-1 ubiquitin ligase ICP0 [118]. From these combined studies, we have initial insight into how recognition of methylated DNA by ZBTB38 can be translated into the recruitment of protein factors that facilitate downstream chromatin remodeling and transcriptional alterations, as well as how PTMs may regulate cellular abundance of this protein. Though additional research is required to delineate the molecular intricacies linking ZBTB38 gene targeting and specific protein recruitment with chromatin remodeling, transcriptional response, and cellular function.

Furthermore, the cellular functions of ZBTB38 remain to be fully defined. Several studies have correlated various single-nucleotide polymorphisms within the *ZBTB38* gene with disease predispositions such as prostate cancer [119], and organismal characteristics such as height [120]. However, until recently few studies have focused on evaluating the protein functions of ZBTB38. From these reports, there is mounting evidence that ZBTB38 is an important transcriptional regulator in maintenance of numerous essential cellular processes. Indeed, ZBTB38 transcriptional activities have been shown to be required for directing proper proliferation and differentiation in mouse embryonic stem cells [121,122]. Furthermore, ZBTB38-mediated transcription has been correlated with cell survival during oxidative stress [114], negative regulation of apoptosis [123,124], and maintenance of genomic stability during DNA replication [117]. Given these roles, it is unsurprising that ZBTB38 activities have been correlated with various disease conditions. Recently, it was observed that selective upregulation of ZBTB38 led to direct repression of anti-inflammatory genes and promotion of rheumatoid arthritis [125]. In addition, increased ZBTB38 expression has been implicated in promoting migration and invasion of bladder cancer cells as well as poor prognosis in bladder cancer patients [126]. Intriguingly, ZBTB38 depletion was shown to augment efficacy of DNA methyltransferase inhibitors in multiple cancer cell line models [116]. Conversely, it has been demonstrated that reintroduction of ZBTB38 at spinal cord injury sites reverts autophagy and ER stress-induced apoptosis, resulting in partial recovery of damaged tissues [127,128]. Combined, these observations suggest that the cellular activities of ZBTB38 are context dependent, though the full spectrum of ZBTB38 biological functions in normal and disease states remain to be defined.

ZFP57

ZFP57, a member of the Krüpple-associated box (KRAB) domain family, was first identified in a search for ZF-containing genes expressed during early stages of rat peripheral nerve development [129]. KRAB domain ZF proteins broadly function as transcriptional repressors, where DNA recognition by C-terminal C_2H_2 ZFs leads to recruitment of KAP1 (KRAB-associated protein 1; also known as TRIM28 (tripartite motif protein 28), Tif1β, or KRIP-1) to the KRAB domain; followed by scaffolded assembly of chromatin corepressor complexes [130]. In addition to the KRAB domain, human ZFP57 has two sets of four and three C_2H_2 ZF clusters separated by an intrinsically disordered linker region (Fig. 1). From comparison with mouse Zfp57, it was determined that two highly conserved ZFs (ZFs 3 and 4 in human and ZFs 2 and 3 in mouse) were necessary and sufficient to preferentially bind a methylated hexanucleotide sequence (TGC^{m}CGC) [131]. The functions of the remaining ZFP57 ZFs remain to be determined. It is notable that ZFP57 represents the only other member of the ZF MBP class identified thus far to exhibit preferential selectivity for distinguishing mCpG over CpG containing DNA sites *in vitro* and within cells [131 133]. Significantly, within cells it has been shown that the ability of ZFP57 to localize to its target DNA sequences is absolutely dependent on CpG methylation [133].

Although initially determined to function as a transcriptional repressor during early embryonic development of the nervous system [129], ZFP57 is now known to play a key role in DNA methylation maintenance at imprinting germline differentially methylated regions (DMRs) [134]. Several investigations over the years, primarily in mouse ESCs, have illuminated insight into the mechanisms by which Zfp57 regulates this essential developmental process. First, Zfp57 has higher expression levels during early embryonic development that decrease appreciably after ESC differentiation [135]. In adults, the expression of Zfp57 is maintained and localized to testes and ovaries [136], where it is readily available to participate in properly directing embryogenesis [129]. Second, in mouse ESCs it was demonstrated that Zfp57 exclusively binds at germline imprinted DMRs and not at somatically derived or nonimprinting DMRs [133], and that during embryogenesis Zfp57 activities are required to maintain DNA methylation for both maternal and paternal imprinted DMRs [136]. Indeed, Zfp57 loss in mouse ESCs or embryos results in hypomethylation for a number of imprinted germline DMRs [131,136 138] that cannot be restored by reintroducing Zfp57 [138]. Consistently, loss-of-function ZFP57 mutations identified within human cases of transient neonatal diabetes were correlated with loss of DNA methylation at several imprinted regions [139,140].

Third, the ability of Zfp57 to regulate DNA methylation maintenance at imprinting germline DMRs is not only dependent on methyl-selective recognition of its DNA binding sequence, but also its association with KAP1 and subsequent recruitment of multiple factors necessary for inducing heterochromatin. Specifically, the Zfp57/KAP1 complex colocalizes the HMT SETDB1 (Su(var)3-9, Enhancer-of-zeste and trithorax domain bifurcated 1; catalyzes H3K9me3 at imprinted loci [141]) and heterochromatin protein 1 (HP1) to methylated imprinting control regions, which help to establish heterochromatin formation [131,142]. Indeed, KAP1-knockout mouse ESCs were characterized by a significant loss of the repressive H3K9me3 histone mark at imprinted loci known to be occupied by the Zfp57/KAP1 complex [131]. In addition, the Zfp57/KAP1 complex has been shown to associate with

DNMTs and UHRF1; recruitment of which is required for maintaining DNA methylation during the imprinting process [131,143]. This latter process is believed to be facilitated by HP1, which is known to selectively recognize H3K9me3 and recruit the *de novo* DNMTs 3a and 3b [144,145]. Combined, these findings indicate that the Zfp57/KAP1 association plays an integral role in the establishment and retention of DNA methylation and heterochromatin at target imprinted germline DMRs during early development.

In addition to its normal function, it has been demonstrated that cancer cells can take advantage of the ES cell and genomic repressive functions of ZFP57 to promote the disease state. Specifically, overexpression of ZFP57 was found to promote anchorage-independent growth in human fibrosarcoma HT1080 cells as well as induce tumor formation of these cells in mouse models [146]. Furthermore, ZFP57 was observed to be overexpressed in patient tumor samples from several different cancer types relative to normal tissues [146]. More recently, ZFP57 was implicated in directly facilitating liver metastasis of colon cancer in humans [147]. However, the role of ZFP57 in breast cancer may be more complex. An initial study reported that 40% of analyzed human breast cancer tumors exhibited elevated levels of ZFP57 relative to normal tissue [146]; though a second study demonstrated that across a cohort of patient breast cancer samples, ZFP57 mRNA and protein levels were decreased relative to adjacent normal tissues [148]. It was further found in the latter study that overexpression of ZFP57 reduced proliferation in breast cancer cells through direct DNA methylation dependent transcriptional repression of target genes [148]. Combined, these initial studies indicate that depending on the tumor context, ZFP57 can either function to promote or suppress the cancerous phenotype. Though more studies are needed to define the role of ZFP57 in cancer as well as other disease conditions, and discern whether there is merit in considering direct therapeutic targeting of this protein.

KLF4

KLF4, a member of the Krüppel-like factor (KLF) family of TFs, was first isolated from NIH 3T3 cells in a search to identify TFs involved in growth regulation [149]. KLF family proteins all share a conserved set of three C-terminally localized DNA recognizing C_2H_2 ZFs, but harbor variances in their N-terminal domains that facilitate differential recruitment of cofactors to mediate distinct cellular processes [150]. The N-terminal region of KLF4 is known to contain both a transactivation (TAD) and repressor domain (RD) [151] (Fig. 1). This overall domain architecture provides KLF4 with the capability of functioning as either a transcriptional activator or repressor depending on the cellular context and associated protein interacting partners [151,152]. Furthermore, KLF4 is known to acquire a variety of PTMs including phosphorylation, acetylation, and methylation that fine-tune its protein interactions and subsequent transcriptional outcomes [153]. In terms of DNA recognition, KLF family members have been observed to recognize a conserved DNA binding motif, termed the CACCC box. Consistently, ChIP-seq analysis within mouse ESCs followed by *de novo* motif searching defined the specific consensus motif for Klf4 as GGG(T/C)G(T/G)G [154]. Intriguingly, several KLF proteins, including KLF4, have been identified in multiple global searches for mCpG-selective readers [155 157]. Consistently, intersection of the abovementioned Klf4 ChIP-seq data with methylome data in mouse ESCs demonstrated that Klf4 localized to both CpG and mCpG sites, where many of the methylated occupation events occurred within the GG^{m}CGTG context [156]. In a separate study, KLF4 was found *in vitro* to preferentially bind to a GG^{m}CGGG motif; cellular occupation of which was verified in human ESCs using combined ChIP and bisulfite (BS) sequencing strategies [155].

Further structural and biochemical analyses demonstrated that *in vitro* KLF4 exhibits comparable recognition for DNA sequences harboring CpG, mCpG, or TpG steps [158,159]. Notably, it was shown that KLF4 binds equally well to targets that contain interchangeable mCpG or TpG base steps [158], though its ability to select mCpG over CpG sites is marginal (<2-fold) [159]. In short, the combined experimental evidence suggests that although DNA recognition by KLF4 is sequence context dependent, its relative insensitivity for distinguishing CpG, mCpG, and TpG sites within this core dramatically expands the DNA sequence space that can be selectively read by KLF4 to mediate cellular processes [47]. Indeed, although DNA methylation typically functions to either repel or recruit TFs and hence regulate their transcriptional activities, ChIP-seq/BS-seq analyses has demonstrated that KLF4 is capable of binding at target loci regardless of methylation status [156,160]. This has significant implications in disease contexts. For example, in glioblastoma cells more than half of KLF4 target sites were methylated, occupation of which led to recruitment of chromatin remodeling complexes that signaled for euchromatin formation and transcriptional upregulation [160]. Nevertheless, correlating cellular KLF4 CpG vs. mCpG vs. TpG occupation patterns with specific normal and disease-state functions requires further investigation.

To date, multiple reviews have provided detailed summaries for the various cellular roles of KLF4 in development, as well as normal and disease-state functions [153,161 164]. Of particular note, KLF4 has been identified as one of the four essential

factors required for inducing pluripotent stem cells [165,166]. Furthermore, KLF4 is expressed in a number of tissues and has been found to be a key regulator for a number of important cellular processes. Specifically, KLF4 transcriptional activities have been implicated in regulating genomic stability [167,168], cellular proliferation [169,170], DNA damage response [171], and apoptosis. However, depending on the cellular context KLF4 has been observed to participate in both anti- [172,173] and proapoptotic activities [174,175]. This ability of KLF4 to have context-specific bimodal regulation of cellular activities has also been observed in cancer [176,177]. Although several reports have demonstrated that many cancers gain a survival advantage by downregulating KLF4 expression [178 181], other evidence suggests KLF4 can also function to promote the cancerous condition [182 184]. Two recent reviews detail the complex roles for KLF4 in cancer [40,153]. Although much is now known about the diverse cellular roles of KLF4, the full mechanisms by which sequence-specific DNA targeting impacts recruitment of differing protein-interacting partners to affect downstream KLF4 cellular activities remain to be delineated.

WT1

Wilms' tumor gene 1 (*WT1*) is so named because it was first identified as a gene associated with development of a pediatric kidney malignancy, termed Wilms' tumor (WT) [185,186]. In general, WT1 is composed of four C-terminal C_2H_2 ZFs as well as designated N-terminal repression and TADs (Fig. 1) that facilitate its ability to function as a bimodal transcriptional regulator [187,188]. However, WT1 is known to exist in multiple functional isoforms that result from alternative splicing. Most notable of these isoforms is the inclusion/exclusion of three amino acids KTS (designated +KTS/−KTS, respectively) in the linker region between ZFs 3 and 4 [189]. Importantly, different subsets of the WT1 ZFs have been implicated in selective RNA and DNA recognition; targeting of which can be impacted by inclusion/exclusion of the KTS sequence. Specifically, structural evidence suggests that ZFs 2 4 are required for base-specific DNA recognition [190,191], whereas ZF 1 can enhance affinity of overall DNA binding through nonspecific interactions [191]. However, ZF 1 appears to be essential in recognition and binding of RNA [192]. In the case of DNA recognition, it is generally observed that the −KTS isoform binds tighter than the +KTS, whereas the isoform required for preferential recognition of RNA remains unclear and may be context specific (see Ref. [193] for a recent review on this topic).

It is further notable that WT1 ZFs 2 4 share sequence identity with ZFs 1 3 of EGR1, and as such both of these proteins are capable of binding to the same DNA consensus motif GCG(T/G)GGGCG [194]. Since this first observation, several additional WT1 DNA sequence motifs have been identified (see Ref. [193] for a summary); the majority of which conform to the general GNGNGGGNG sequence proposed from initial structural and biochemical studies [191]. Indeed, ChIP-seq analyses have found that within a variety of cellular contexts WT1 occupies genomic loci harboring motifs similar to that of EGR1 [195,196]. More recent ChIP-seq investigations designed to differentiate −KTS and +KTS isoform-specific genomic targeting in leukemia K562 cells demonstrated that the majority of loci containing the EGR1 motif were occupied by the −KTS isoform [197]. This finding is consistent with *in vitro* studies where the −KTS isoform exhibits enhanced binding for the EGR1 motif relative to the +KTS [190,194,198]. Intriguingly, it was also determined that the −KTS isoform was preferentially localized near the transcription start sites of genes, whereas +KTS isoform peaks were primarily enriched within gene bodies [197]. It has been proposed that the variance in genomic distributions for the two WT1 isoforms may be due to isoform-specific associations with protein-binding partners [193]. However, none of the ChIP-seq investigations to date evaluated the dependency of WT1 occupation events on modification status of the CpG cytosines.

Indeed, it has been shown *in vitro* that methylation of both CpG steps in the GCGTGGGGCG sequence context slightly enhanced DNA binding by WT1 −KTS (~1.8-fold relative to CpG), but binding to this sequence by WT1 +KTS was drastically reduced regardless of cytosine methylation status [190]. More importantly, it was found that WT1 exhibits moderate binding capability for sequences containing the oxidative mC derivatives hydroxymethylC (hmC), formylC (fC), and carboxylC (caC) [190], which are generated by the ten-eleven translocation (TET) enzymes [199,200]. Though notably, substitution of caC at only the 3′-GCG site resulted in an increased specificity for caC over mC by ZF 2 of WT1 [190]. Intriguingly, intersection of WT1 ChIP-seq with RNA-seq data after WT1 depletion suggests that WT1 primarily functions as a transcriptional activator [195,196]. Consistently, WT1 has been identified as a binding partner for TET2 [201], and as such may play a direct role in maintaining the euchromatin state at targeted loci. Taken together, it is possible that CpG cytosine modification status may simultaneously influence WT1 isoform-specific genomic targeting as well as associations with various transcriptional cofactors, though this requires investigation.

In terms of biological function, it is well established that WT1 plays a significant role in ensuring proper embryonic development of the genitourinary system [202 204], and that mutations of this essential

protein results in imbalanced ratios of various WT1 isoforms which leads to a multitude of disease conditions [205,206]. Nevertheless, it is important to note that WT1 appears to have functional duality in regulating a number of cellular activities [206], which can differentially influence its role in disease conditions. Indeed, in pediatric Wilms' tumor WT1 was originally found to function as a tumor suppressor [185,186]. In contrast, in leukemia and several solid tumor types, WT1 activities are seemingly oncogenic [207]. Significantly, it has been demonstrated that targeting of WT1 through RNA interference induced apoptosis in multiple cancer cell line models [208], highlighting the therapeutic potential of targeting this protein. Additional studies are needed to mechanistically deconvolute the roles of the various WT1 isoforms in mediating disease-state conditions, including investigating isoform-specific protein and genomic DNA interactions.

EGR1

Discovered over 30 years ago, EGR1 represents a highly characterized member of the immediate early gene class; the cellular roles of which have been evaluated and reviewed in detail [209 211]. Most notably, it is well established that EGR1 primarily functions as a transcriptional activator, where its cellular activities are specifically induced in response to stress signals and external stimuli [212]. Indeed, EGR1 was initially identified in a screen for genes upregulated by nerve growth factor on exposure to the protein synthesis inhibitor cycloheximide in rat PC12 cells [213]. Subsequently, several additional groups observed stimulated expression of EGR1 in response to multiple growth factors in various cell lines [214 218]. In line with being an inducible TF, EGR1 transcriptional activities have been implicated in regulating several essential processes in both nervous and cardiovascular systems [219 222].

EGR1 contains three C-terminal C_2H_2 ZFs and multiple N-terminal activation domains that can interact with various protein binding partners depending on its needed function (Fig. 1). Additionally, EGR1 is highly posttranslationally modified, which regulates cellular abundance of this protein, as well as directs interactions with its biomolecular interacting partners (reviewed in Refs. [209,210]). As discussed earlier for WT1, EGR1 binds the GC-rich DNA motif GCG(G/T)GGGCG and does so equally well regardless of cytosine methylation status at either CpG site [190,223]. However, it was observed that unlike WT1, incorporation of the oxidative mC derivatives hmC, fC, and caC at either CpG location abolished DNA binding by EGR1 [190]. This observed difference in sensing oxidative mC derivatives is believed to be due to both EGR1 ZFs 1 and 3 containing the key arginine and glutamate residues involved in mCpG recognition [46,224,225], whereas the glutamate in ZF 2 of WT1 is substituted with a glutamine that seemingly allows WT1 to retain some binding capability for hmC, fC, and caC containing sequences [190]. Interestingly, WT1 and EGR1 have been shown in certain instances to function antagonistically [226,227], which could possibly be dictated by their different sensitivities to modified cytosines. Nevertheless, more studies are required to decipher the interplay between these two proteins, the functional significance of their disparate binding capacities for chemically modified DNA targets, and the consequences of this disparity on directing cellular function.

Regardless of its indifference to cytosine methylation status *in vitro*, within cells EGR1 has been observed to predominantly occupy nonmethylated versions of its consensus motifs within promoter CGIs [228,229]. It is worth noting that the EGR1 consensus motif is highly represented within the human genome [230], and that the combined number of functional and nonfunctional sites containing this target site exceeds the typical number of available EGR1 protein molecules. Moreover, it has been shown that this significant presence of nonfunctional or 'decoy' sites can kinetically impede the ability of EGR1 to appropriately localize to functional genomic loci [231,232]. These observations pose an interesting question, how within its relatively short lifetime does EGR1 preferentially localize to its target sites? It may be that other factors like WT1, which can also occupy EGR1 target sites regardless of the DNA methylation status, compete for EGR1 binding. However, it has also been demonstrated *in vitro* that when present at sufficiently high concentrations, a member of the MBD family of MBPs can occupy methylated 'decoy' EGR1 sites and kinetically direct EGR1 to its functional nonmethylated CGI targets [232]. More investigations are required to elucidate the mechanisms by which EGR1 kinetically localizes to its target sites and to define the roles that other selective mCpG reader proteins may play in this process.

CTCF

CTCF is an abundant and well-characterized protein containing 11 tandem C_2H_2 ZFs (Fig. 1) that was first identified as a direct transcriptional repressor of *c-myc* [233]. More generally, CTCF is known to work with a number of protein interacting partners to regulate chromatin architecture and influence transcriptional outcomes. The diverse functions of CTCF have been the topic of several reviews [234 238]. Relative to the other ZF MBPs, the core consensus CTCF DNA-binding motif NC(A/G)NNAG(G/A)NGGC(G/A)(C/G)(T/C) is much larger and predominantly recognized by ZFs 3 7 [239]. Though it has been shown that depending on the genomic site, various combinations of the remaining CTCF ZFs can be used to recognize

additional conserved 5′- and 3′-flanking regions outside the core sequence [240,241]. This modularity in CTCF motif recognition is believed to regulate the association of this protein at various genomic loci by fine-tuning the strength of the DNA binding interaction [241].

Variable CTCF genomic occupancies have also been associated with cytosine methylation status at either position 2 or 12 of the core consensus motif [242]; methylation of which can occur within either the CpA or CpG context. Specifically, several studies have indicated that DNA methylation was disruptive to DNA binding by CTCF [70,243 246]. Importantly, comparative intersection of ChIP-seq and methylome data demonstrated that in a number of cell types, there is indeed an overall indirect correlation between CTCF occupancy and cytosine methylation for the genomic target sites shared between cells [242]. However, it was also shown that differential methylation of the CTCF core consensus motif was implicated in directing the observed variances in genomic occupancy across cell types. Recent structural evidence has provided insight for how this observed differential methylation of the CTCF motif may modulate CTCF DNA targeting [239]. Notably, it was determined that CTCF binding was relatively indifferent to cytosine methylation status at position 12, but was drastically reduced when position 2 was methylated. This difference in binding is believed to be due to substitution of the key glutamate needed for mC recognition with an aspartate in the α-helix of ZF7, which is centralized at position 2. In contrast, CTCF ZFs 3 and 4 can, respectively, contribute the key arginine and glutamate residues used for accommodating recognition of a methylated cytosine at position 12.

Intriguingly, it was also determined that although methylation incidence of the CTCF motif was comparatively increased for a number of sites in immortalized cells relative to normal cells, there was no corresponding decrease in the total number of CTCF occupation events [242]. A possible explanation for this is CTCF transcript levels were also found to be elevated in the immortalized cells, which could drive CTCF toward occupancy of genomic targets harboring its less preferred methylated consensus motif. Nevertheless, the broader mechanisms by which DNA methylation may influence CTCF DNA targeting, protein interactions, and cellular activities within normal and diseased cells requires further investigation.

Mode of Methylated DNA Recognition by C_2H_2 ZF Proteins: Structural Insight

C_2H_2 ZF scaffolds, characterized by two β-strands stabilized against an α-helix via tetrahedral coordination of a Zn^{2+} atom by two cysteine and histidine residues [247,248], represent a common DNA binding domain found in human cells. Most proteins use tandem arrays of C_2H_2 ZFs in selective recognition of their cognate DNA sites, where base-specific contacts are facilitated by amino acids localized at specific positions along the α-helix [247,248]. Despite the cellular abundance of C_2H_2 ZF containing proteins, only a subset of this protein class has been identified to exhibit recognition of mCpG sites. Over the past several years, high-resolution structures for most of the ZF MBPs discussed earlier in complex with their consensus methylated DNA targets have become available [57,58,112,132,158,159,190,223,239], providing mechanistic insight for how this family of MBPs can distinguish this essential epigenetic mark. This topic has been extensively reviewed recently [46] and will only be briefly summarized here.

From the currently available structures, several commonalities in ZF recognition of mCpG sites have emerged. First, for the majority of ZF MBPs a key arginine and glutamate residue are required for recognition of the mCpG steps. These two key residues can be contributed from one (ZFP57, KLF4, WT1) or two ZF α-helices (ZBTB MBPs, CTCF) (Fig. 2). Specifically, the arginine residue participates in hydrogen bonding interactions with one (ZFP57, KLF4, WT1, EGR1, CTCF) or two (ZBTB33) of the 3′-G bases, which further positions the arginine side chain to make van der Waals contacts with the methyl group of the 5′-mC (designated the 'mC-Arg-G triad' [249]) (Fig. 2). More recently, it was demonstrated that the C-terminal ZBTB38 ZFs use a lysine residue to make an analogous water-mediated hydrogen bond with one of the 3′-G bases [112], indicating that an alternative mode for ZF recognition of mCpG sites exists.

Furthermore, current structural evidence indicates that the extent to which the conserved glutamate residue is involved in recognition of the mC contributes to determining whether a ZF MBP will exhibit marginal (<2-fold) or higher selectivity (≥20-fold) for distinguishing mC over C. Indeed, for the more selective ZF MBPs (ZBTB33, ZBTB38 (C-term ZFs), ZFP57) the glutamate residue makes classical hydrogen bonding interactions between the glutamate carbonyl oxygens and the N4 atom of one (ZFP57) or both mCs (ZBTB33, ZBTB38), as well as CH···O type hydrogen bonds with the methyl group of one (ZBTB38, ZFP57) or both (ZBTB33) mCs (Fig. 2). In contrast, for the less selective ZF MBPs (KLF4, WT1, EGR1, CTCF) the glutamate is positioned such that it can only make either van der Waals interactions from its aliphatic chain and/or CH···O type hydrogen bonding interactions via the carboxylate oxygens with the methyl group of the mCs (Fig. 2). Finally, emerging evidence indicates that specific amino acid substitutions at the

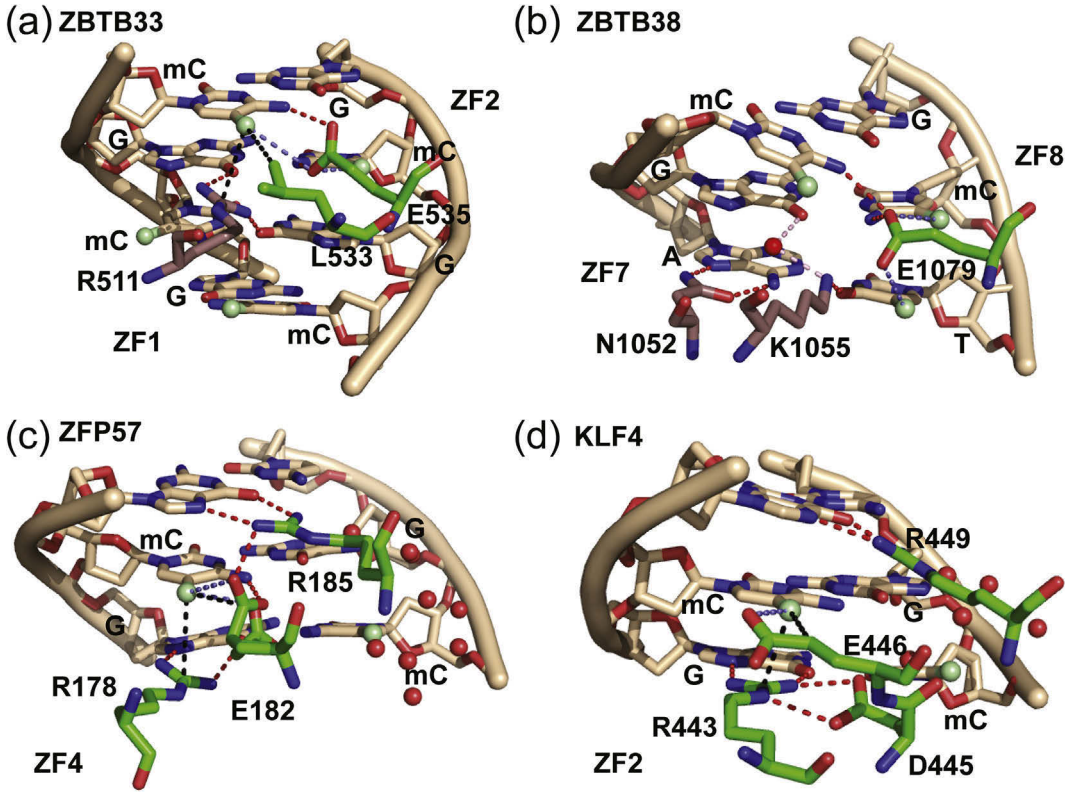

Fig. 2. Representative crystal structures for members of the ZF MBP family in complex with their respective methylated DNA targets. Key residues for recognition of the mCpG palindrome are depicted for: (a) human ZBTB33 (PDB 4F6N); (b) human ZBTB38 (C-terminal ZFs; PDB 6E93); (c) mouse Zfp57 (PDB 4GZN); and (d) mouse Klf4 (PDB 4M9E) methylated DNA complex structures. Red spheres indicate water molecules. Red dotted lines denote classical hydrogen bond interactions; blue dotted lines indicate CH···O type hydrogen bonds; pink dotted lines represent water-mediated hydrogen bond interactions; and black dotted lines designate van der Waals interactions.

glutamate position can fine-tune mC selectivity or provide variable sensitivity to oxidative mC intermediates [158,190,250,251].

Second, all the ZF MBPs exhibit asymmetric binding at the mCpG palindrome, where the majority of arginine/lysine and glutamate interactions predominate at one mCpG (Fig. 2) [224]. In most cases, the remaining cross-strand mC has a hydration shell of ordered water molecules [252,253] that is believed to contribute to the overall binding energetics [132,254,255]. Notably, each of the ZF MBPs have similar binding affinities for symmetrically methylated and strand-specific hemi-methylated DNA [57,58,132,159,190], though the cellular significance of this asymmetric binding capability for each protein is not yet fully understood.

Third, there is emerging evidence that local DNA shape alterations induced by cytosine methylation are required for recruitment of ZF MBPs to their target methylated DNA sites [16,46,256]. Specifically, it was observed that cytosine methylation within the context of duplex B-form DNA results in significant perturbations to the roll and propeller twist localized at the mCpG palindrome [16]. Analysis of DNA shape parameters [16,257] for the three ZF MBP:meDNA structures where the mCpG palindrome is centralized within the bound DNA sequence (ZBTB33, ZFP57 and KLF4) demonstrated that the induced base-specific alterations in roll at the 5′-mC and 3′-G positions observed in the free DNA were effectively inverted after protein binding [46]. Thus, the ZF MBPs may use a common mode of local DNA shape recognition to facilitate binding to their methylated DNA targets. Finally, CURVES+ analysis [257] 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. Similar analysis for the recent C-terminal ZBTB38 ZFs in complex with methylated DNA [112] resulted in a total 15.8° bend deformation from B-form. These trends suggest that the more selective ZF MBPs may induce a larger global DNA bend on binding than the less selective ZF MBPs. Nevertheless, more ZF MBP:meDNA structures are required to confirm the overall role of DNA shape in facilitating selective recruitment of ZF MBPs to methylated DNA loci.

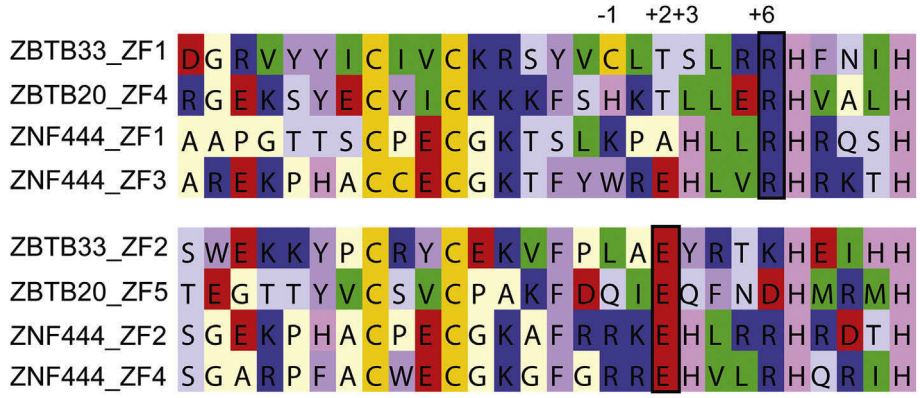

Fig. 3. Sequence alignment showing two potential ZF MBPs that share sequence similarity with ZBTB33 ZFs 1 and 2 at the key arginine and glutamate residues (boxed) involved in mCpG recognition. Note that ZNF444 contains two sets of ZFs that share the ZBTB33 mCpG binding sequence elements. Sequences were identified by using the following search motifs: ZF 1: LXRH and ZF 2: EXXXXH or FXLXE, where X represents any amino acid.

Future Perspectives

The ZF MBPs discussed represent only a subset of what is emerging as a larger MBP family. Indeed, several recent high-throughput studies have identified over 50 additional ZF-containing proteins that appear to have mCpG selective binding capabilities, but remain to be fully characterized [155 157] (reviewed in Ref. [258]). Included within this compiled list is 11 additional members of the KLF family, the closely related SP1 and SP2 proteins, as well as EGR4. This finding is not surprising and was predicted [159] given that all of these proteins share high sequence conservation with KLF4 in one or more of their ZF α-helices; particularly for positioning of the key arginine and glutamate residues needed for mCpG recognition. Further motif searching (MEME suite [259]) for proteins on this list that share sequence similarity with the ZF α-helices of ZBTB33 identified ZBTB20 and ZNF444 (Fig. 3). Nevertheless, the remaining potential ZF MBPs do not readily appear to share the core required sequence elements currently believed to be needed for mCpG recognition. This suggests ZF proteins may preferentially select for methylated DNA sites through additional mechanisms, though further studies on these identified ZF proteins are needed to confirm this possibility.

In addition to structural insight, advances in cell-based high-throughput sequencing strategies have begun to provide a more detailed understanding for how mCpG sensitivity correlates with ZF MBP normal and disease state biological functions. Though the level of information in this regard is highly variable among the discussed ZF MBPs. Indeed, across current ZF MBP family members, there is disparate mechanistic detail for defining how methyl-dependent DNA binding capability impacts biomolecular interactions and downstream cellular responses. Furthermore, several of the ZF MBPs have preferential selectivity for more than one DNA binding sequence; however, the consequence of this sequence modality in mechanistically driving normal and disease-state cellular functions remains unclear.

Furthermore, we feel that it will be important to define the extent of functional redundancy not only within but also across ZF MBP families in directing cellular activities. It is evident that members within families (e.g. ZBTB MBPs, KLF proteins, EGR proteins) share protein sequence similarities that afford many of them a level of commonality in identifying DNA targets. However, there are also unique protein sequence elements that distinguish their overall cellular activities; mechanistic aspects of which remain to be determined. Finally, it is evident that the various ZF MBP classes exhibit diversity in their primary cellular functions, though the extent of functional overlap, if any, is unclear. In short, there are several important open avenues of investigation available, pursuit of which will expand mechanistic understanding of ZF MBP cellular roles in normal and disease states.

Acknowledgments

Financial support from grants RSG-14-185-01-DMC from the American Cancer Society and MCB-1715370 from the National Science Foundation is gratefully acknowledged.

Received 16 August 2019;
Received in revised form 13 September 2019;
Accepted 16 September 2019
Available online 16 October 2019

Keywords:
Zinc finger methyl-CpG binding proteins;
DNA methylation;
Epigenetics;
Protein–DNA interactions

Abbreviations used:
PTMs, posttranslational modifications; mCpG, methyl-CpG; CGIs, CpG islands; TF, transcription factor; MBPs, methyl-CpG binding proteins; C_2H_2, Cys_2His_2; ZF, zinc finger; ZBTB, zinc finger and BTB domain-containing protein; BLAST, Basic Local Alignment Search Tool; BTB/POZ, broad complex, tramtrack, bric-á-brac, pox zinc finger; ZFP57, zinc finger protein 57; KLF4, Krüppel-like factor 4; WT1, Wilms' tumor protein 1; EGR1, growth response protein 1; CTCF, CCCTC-binding factor; NCoR, nuclear receptor corepressor; HDAC, histone deacetylase; Sp, specificity protein; CtBP, C-terminal binding protein; HMT, histone methyltransferase; HDM, histone demethylase; KRAB, Krüpple-associated box; KAP1, KRAB-associated protein 1; DMR, differentially methylated region; HP1, heterochromatin protein 1; TAD, transactivation domain; RD, repressor domain; ChIP, chromatin immunoprecipitation; BS, bisulfite; TET, ten-eleven translocation.

References

[1] S.U. Kass, D. Pruss, A.P. Wolffe, How does DNA methylation repress transcription? Trends Genet. 13 (1997) 444–449.
[2] S. Sharma, T.K. Kelly, P.A. Jones, Epigenetics in cancer, Carcinogenesis 31 (2009) 27–36.
[3] A. Bird, DNA methylation patterns and epigenetic memory, Genes Dev. 16 (2002) 6–21.
[4] S. Augui, E.P. Nora, E. Heard, Regulation of X-chromosome inactivation by the X-inactivation centre, Nat. Rev. Genet. 12 (2011) 429–442.
[5] M.S. Bartolomei, A.C. Fergurson-Smith, Mammalian genomic imprinting, Cold Spring Harb. Perspect. Biol. 3 (2011) a002592.
[6] K.D. Robertson, A.P. Wolffe, DNA methylation in health and disease, Nat. Rev. Genet. 1 (2000) 11–19.
[7] S.B. Baylin, P.A. Jones, A decade of exploring the cancer epigenome – biological and translational implications, Nat. Rev. Cancer 11 (2011) 726–734.
[8] Z. Jin, Y. Liu, DNA methylation in human diseases, Genes Dev. 5 (2018) 1–8.
[9] M. Kulis, M. Esteller, DNA methylation and cancer, Adv. Genet. 70 (2010) 27–56.
[10] Y. Xiong, Y. Wei, Y. Gu, S. Zhang, J. Lyu, B. Zhang, et al., DiseaseMeth version 2.0 a major expansion and update of the human disease methylation database, Nucleic Acids Res. 45 (2017) D888–D895.
[11] A. Sandelin, P. Carninci, B. Lenhard, J. Ponjavic, Y. Hayashizaki, D.A. Hume, Mammalian RNA polymerase II core promoters: insights from genome-wide studies, Nat. Rev. Genet. 8 (2007) 424–436.
[12] P.A. Jones, P.W. Laird, Cancer epigentics comes of age, Nature Genet. 21 (1999) 163–167.
[13] S.B. Baylin, J.G. Herman, DNA hypermethylation in tumorigenesis: epigenetics joins genetics, Trends Genet. 16 (2000) 168–174.
[14] A. Razin, H. Cedar, DNA methylation and gene expression, Microbiol. Rev. 55 (1991) 451–458.
[15] A.C. Dantas Machado, T. Zhou, S. Rao, P. Goel, C. Rastogi, A. Lazarovici, et al., Evolving insights on how cytosine methylation affects protein-DNA binding, Brief Funct. Genom. 14 (2014) 61–73.
[16] S. Rao, T.-P. Chiu, J.F. Kribelbauer, R.S. Mann, H.J. Bussemaker, R. Rohs, Systematic prediction of DNA shape changes due to CpG methylation explains epigenetic effects on protein-DNA binding, Epigenetics Chromatin 11 (2018) 6.
[17] K. Gaston, M. Fried, 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. 23 (1995) 901–909.
[18] S.M.M. Iguchi-Ariaga, W. Schaffner, CpG methylation of the cAMP-responsive enhancer/promoter sequence TGACGT-CA abolishes specific factor binding as well as transcriptional activation, Genes Dev. 3 (1989) 612–619.
[19] G.C. Pendergast, E.B. Ziff, Methylation-sensitive sequence-specific DNA binding by the c-Myc basic region, Science 251 (1991) 186–189.
[20] A. Bird, D. Macleod, Reading the DNA methylation signal, Cold Spring Harb. Symp. Quant. Biol. 69 (2004) 113–118.
[21] A. Fournier, N. Sasai, M. Nakao, P.-A. Defossez, The role of methyl-binding proteins in chromatin organization and epigenome maintenance, Brief Funct. Genom. 11 (2011) 251–264.
[22] M. Joulie, B. Miotto, P.-A. Defossez, Mammalian methyl-binding proteins: what might they do? Bioessays 32 (2010) 1025–1032.
[23] L. Lopez-Serra, M. Estellar, Proteins that bind methylated DNA and human cancer: reading the wrong words, Br. J Cancer 98 (2008) 1881–1885.
[24] T. Shimbo, P.A. Wade, Proteins that read DNA methylation, Adv. Exp. Med. Biol. 945 (2016) 303–320.
[25] J.M. Cramer, D. Pohlmann, F. Gomez, L. Mark, B. Kornegay, C. Hall, et al., Methylation specific targeting of a chromatin remodeling complex from sponges to humans, Sci. Rep. 7 (2017) 40674.
[26] P.L. Jones, G.J. Veenstra, P.A. Wade, D. Vermaak, S.U. Kass, N. Landsberger, et al., Methylated DNA and MeCP2 recruit histone deactylase to repress transcription, Nat. Genet. 19 (1998) 187–191.
[27] N. Sasai, E. Matsuda, E. Sarashina, Y. Ishada, M. Kawaichi, Identification of a novel BTB-zinc finger transcriptional repressor, CIBZ, that interacts with CtBP corepressor, Genes Cells 10 (2005) 871–885.
[28] A. Weber, J. Marquardt, D. Elzi, N. Forster, S. Starke, A. Glaum, et al., Zbtb4 represses transcription of P21CIP1 and controls the cellular responses to p53 activation, EMBO J. 27 (2008) 1563–1574.
[29] H.G. Yoon, D.W. Chan, A.B. Reynolds, J. Qin, J. Wong, N-CoR mediates DNA methylation-dependent repression through a methyl CpG binding protein Kaiso, Mol. Cell 12 (2003) 723–734.
[30] M.P. Ball, J.B. Li, Y. Gao, J.H. Lee, E.M. LeProust, I.H. Park, B. Xie, G.Q. Daley, G.M. Church, Targeted and genome-scale strategies reveal gene-body methylation

signatures in human cells, Nat. Biotechnol. 27 (2009) 361–368.
[31] A.K. Maunakea, R.P. Nagarajan, M. Bilenky, T.J. Ballinger, C. D'Souza, S.D. Fouse, et al., Conserved role of intragenic DNA methylation in regulating alternative promoters, Nature 466 (2010) 253–257.
[32] R. Lister, M. Pelizzola, R.H. Dowen, R.D. Hawkins, G. Hon, J. Tonti-Filippini, et al., Human DNA methylomes at base resolution show widespread epigenomic differences, Nature 462 (2009) 315–322.
[33] D. Jjingo, A.B. Conley, S.V. Yi, V.V. Lunyak, I.K. Jordan, On the presence and role of human gene-body DNA methylation, Oncotarget 3 (2012) 462–474.
[34] P.A. Jones, Functions of DNA methylation: islands, start sites, gene bodies and beyond, Nat. Rev. Genet. 13 (2012) 484–492.
[35] G.L. Maor, A. Yearim, G. Ast, The alternative role of DNA methylation in splicing regulation, Trends Genet. 31 (2015) 274–280.
[36] S. Sarda, S. Hannenhalli, Orphan CpG islands as alternative promoters, Transcription 9 (2018) 171–176.
[37] C.T. Nguyen, F.A. Gonzales, P.A. Jones, Altered chromatin structure associated with methylation-induced gene silencing in cancer cells: correlation of accessibility, methylation, MeCP2 binding and acetylation, Nucleic Acids Res. 29 (2001) 4598–4606.
[38] T. Baubec, R. Ivanek, F. Lienert, D. Schubeler, Methylation-dependent and -independent genomic targeting principles of the MBD protein family, Cell 153 (2013) 480–492.
[39] G.D. Ginder, D.C.J. Williams, Readers of DNA methylation, the MBD family as potential therapeutic targets, Pharmacol. Ther. 184 (2018) 98–111.
[40] N. Mahmood, S.A. Rabbani, DNA methylation readers and cancer: mechanistic and therapeutic applications, Front. Oncol. 9 (2019) 489.
[41] O.J. Sansom, K. Maddison, A.R. Clarke, Mechanism of disease: methyl-binding domain proteins as potential therapeutic targets in cancer, Nat. Clin. Pract. Oncol. 4 (2007) 305–315.
[42] N. Sasai, P.-A. Defossez, Many paths to one goal? The proteins that recognize methylated DNA in eukaryotes, Int. J. Dev. Biol. 53 (2009) 323–334.
[43] A. Prokhortchouk, B. Hendrich, A. Jorgensen HR, M. Wilm, G. Georgiev, A. Bird, et al., The p120 catenin partner Kaiso is a DNA methylation-dependent transricptional repressor, Genes Dev. 15 (2001) 1613–1618.
[44] A.V. Prokhortchouk, D.S. Aitkhozhina, A.A. Sablina, A.S. Ruzov, E.B. Prokhortchouk, Kaiso, a new protein of the BTB/POZ family, specifcally binds to methylated DNA sequences, Russ. J. Genet. 37 (2001) 603–609.
[45] G.J.P. Filion, S. Zhenilo, S. Salozhin, D. Yamada, E. Prokhortchouk, P.-A. Defossez, A family of human zinc finger proteins that bind methylated DNA and repress transcription, Mol. Cell. Biol. 26 (2006) 169–181.
[46] N.O. Hudson, B.A. Buck-Koehntop, Zinc finger readers of methylated DNA, Molecules 23 (2018) 2555.
[47] R. Ren, J.R. Horton, X. Zhang, R.M. Blumenthal, X. Cheng, Detecting and interpreting DNA methylation marks, Curr. Opin. Struct. Biol. 53 (2018) 88–99.
[48] J.M. Daniel, A.B. Reynolds, The catenin $p120^{ctn}$ interacts with Kaiso, a novel BTB/POZ domain zinc finger transcription factor, Mol. Cell. Biol. 19 (1999) 3614–3623.
[49] J.M. Daniel, Dancing in and out of the nucleus: $p120^{ctn}$ and the transcription factor Kaiso, Biochim. Biophys. Acta 1773 (2007) 59–68.
[50] L. Zhang, M. Gallup, L. Zlock, Y.T. Feeling Chen, W.E. Finkbeiner, N.A. McNamara, Cigarette smoke mediates nuclear to cytoplasmic trafficking of transcriptional inhibitor Kaiso through MUC1 and p120-catenin, Am. J. Pathol. 186 (2016) 3146–3159.
[51] P.-X. Zhang, Y. Wang, Y. Liu, G.-Y. Jiang, Q.-C. Li, E.-H. Wang, p120-catenin isoform 3 regulates subcellular localization of Kaiso and promotes invasion in lung cancer cells via a phosphorylation-dependent mechanism, Int. J. Oncol. 38 (2011) 1625–1635.
[52] C.C. Pierre, S.M. Hercules, C. Yates, J.M. Daniel, Dancing from bottoms up – roles of the POZ-ZF transcription factor Kaiso in cancer, Biochim. Biophys. Acta Rev. Cancer 1871 (2019) 64–74.
[53] J.M. Daniel, C.M. Spring, H.C. Crawford, A.B. Reynolds, A. Baig, 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 (2002) 2911–2919.
[54] H. Kiefer, F. Chatail-Hermitte, P. Ravassard, E. Bayard, I. Brunet, J. Mallet, ZENON, a novel POZ Kruppel-like DNA binding protein associated with differentiation and/or survivial of late postmitotic neurons, Mol. Cell. Biol. 25 (2005) 1713–1729.
[55] Y. Oikawa, R. Omon, T. Nishii, Y. Ishida, M. Kawaich, E. Matsuda, The methyl-CpG-binding protein CIBZ suppresses myogenic differentiation by directly inhibiting myogenin expression, Cell Res. 21 (2011) 1578–1590.
[56] N. Sasai, M. Nakao, P.-A. Defossez, Sequence-specific recognition of methylated DNA by human zinc-finger proteins, Nucleic Acids Res. 38 (2010) 5015–5022.
[57] B.A. Buck-Koehntop, R.L. Stanfield, D.C. Ekiert, M.A. Martinez-Yamout, H.J. Dyson, I.A. Wilson, et al., Molecular basis for recognition of methylated and specific DNA sequences by the zinc finger protein Kaiso, Proc. Natl. Acad. Sci. U. S. A. 109 (2012) 15229–15234.
[58] E.N. Nikolova, R.L. Stanfield, H.J. Dyson, P.E. Wright, CH···O hydrogen bonds mediate highly specific recognition of methylated CpG sites by the zinc finger protein Kaiso, Biochemistry 57 (2018) 2109–2120.
[59] F. Bohne, D. Langer, U. Martiné, C.S. Eider, R. Cencic, M. Begemann, et al., Kaiso mediates human ICR1 methylation maintenance and H19 transcriptional fine regulation, Clin. Epigenetics 8 (2016) 47.
[60] N.S. Donaldson, C.C. Pierre, M.I. Antsey, S.C. Robinson, S.M. Weerawardane, J.M. Daniel, Kaiso represses the cell cycle gene cyclin D1 via sequence-specific and methyl-CpG-dependent mechanisms, PLoS One 7 (2012) e50398.
[61] E.C. Lopes, E. Valls, M.E. Figueroa, A. Mazur, F.-G. Meng, G. Chiosis, et al., Kaiso contributes to DNA methylation-dependent silencing of tumor suppressor genes in colon cancer cell lines, Cancer Res. 68 (2008) 7258–7263.
[62] C.C. Pierre, J. Longo, B.I. Bassey-Archibong, R.M. Hallett, S. Milosavljevic, L. Beatty, et al., Methylation-dependent regulation of hypoxia inducible factor-1 alpha gene expression by the transcription factor Kaiso, Biochim. Biophys. Acta 1849 (2015) 1432–1441.
[63] A. Pozner, T.W. Terooatea, B.A. Buck-Koehntop, Cell specific Kaiso (ZBTB33) regulation of cell cycle through cyclin D1 and cyclin E1, J. Biol. Chem. 291 (2016) 24538–24550.
[64] S. Qin, B. Zhang, W. Tian, L. Gu, Z. Lu, D. Deng, Kaiso mainly locates in the nucleus in vivo and binds to

methylated, but not hydroxymethylated DNA, Chin. J. Cancer Res. 27 (2015) 148–155.
[65] M. Rodova, K.F. Kelly, M. VanSuan, J.M. Daniel, M.J. Werle, Regulation of the rapsyn promoter by Kaiso and d-catenin, Mol. Cell. Biol. 24 (2004) 7188–7196.
[66] S. Zhenilo, I. Deyev, E. Litvinova, N. Zhigalova, D. Kaplun, A. Sokolov, et al., DeSUMOylation switches Kaiso from activator to repressor upon hyperosmotic stress, Cell Death Differ. 25 (2018) 1938–1951.
[67] J. Rush, A. Moritz, K.A. Lee, A. Guo, V.L. Goss, E.J. Spek, H. Zhang, R.D. Poakiewicz, M.J. Comb, Immunoaffinity profiling of tyrosine phosphorylation in cancer cells, Nat. Biotechnol. 23 (2005) 94–101.
[68] S.K. Raghav, S.M. Waszak, I. Krier, C. Gubelmann, A. Isakova, T.S. Mikkelsen, et al., Integrative genomics identifies the corepressor SMRT as a gatekeeper of adipogenesis through the transcription factors C/EBPbeta and KAISO, Mol. Cell 46 (2012) 335–350.
[69] P.-A. Defossez, K.F. Kelly, G.J.P. Filion, R. Perez-Torrado, F. Magdinier, H. Menoni, et al., The human enhancer blocker CTC-binding factor interacts with the transcription factor Kaiso, J. Biol. Chem. 280 (2005) 43017–43023.
[70] I.A. De La Rosa-Velazquez, H. Rincon-Arano, L. Benitez-Bribiesca, F. Recillas-Targa, Epigenetic regulation of the human retinoblastoma tumor suppressor gene promoter by CTCF, Cancer Res. 67 (2007) 2577–2585.
[71] N.S. Donaldson, C.L. Nordgaard, C.C. Pierre, F. KK, S.C. Robinson, L. Swystun, et al., Kaiso regulates Znf131-mediated transcriptional activiation, Exp. Cell Res. 316 (2010) 1692–1705.
[72] S.C. Robinson, N.S. Donaldson-Kabwe, A. Dvorkin-Gheva, J. Longo, L. He, J.M. Daniel, The POZ-ZF transcription factor Znf131 is implicated as a regulator of Kaiso-mediated biological processes, Biochem. Biophys. Res. Commun. 493 (2017) 416–421.
[73] C.W. Barrett, J.J. Smith, L.C. Lu, N. Markham, K.R. Stengel, S.P. Short, et al., Kaiso directs the transcriptional corepressor MTG16 to the Kaiso binding site in target promoters, PLoS One 7 (2012) e51205.
[74] B.I. Bassey-Archibong, S.M. Hercules, L.G.A. Rayner, D.H.A. Skeete, S.P. Smith Connell, I. Brain, et al., Kaiso is highly expressed in TNBC tissues of women of African ancestry compared to Caucasian women, Cancer Causes Control 28 (2017) 1295–1304.
[75] J. Jones, H. Wang, B. Karanam, S. Theodore, W. Dean-Colomb, D.R. Welch, et al., Nuclear localization of Kaiso promotes the poorly differentiated phenotype and EMT in infiltrating ductal carcinomas, Clin. Exp. Metastasis 31 (2014) 497–510.
[76] J.F. Vermeulen, R.A.H. van de Ven, C. Ercan, P. van der Groep, E. van der Wall, P. Bult, et al., Nuclear Kaiso expression is associated with high grade and triple-negative invasive breast cancer, PLoS One 7 (2012) e37864.
[77] J. Cofre, J.R. Menezes, L. Pizzatti, E. Abdelhay, Knock-down of Kaiso induces proliferation and blocks granulocytic differentiation in blast crisis of chronic myeloid leukemia, Cancer Cell Int. 12 (2012) 28.
[78] S.P. Short, C.W. Barrett, K.R. Stengel, F.L. Revetta, Y.A. Choksi, L.A. Coburn, et al., Kaiso is required for MTG16-dependent effects on colitis-associated carcinoma, Oncogene 38 (2019) 5091–5106.
[79] L. Wang, J. Ma, X. Wang, F. Peng, X. Chen, B. Zheng, et al., Kaiso (ZBTB33) downregulation by mirna-181a inhibits cell proliferation, invasion, and the epithelial–mesenchymal transition in glioma cells, Cell. Physiol. Biochem. 48 (2018) 947–958.
[80] C.C. Pierre, J. Longo, M. Mavor, S.B. Milosavljevic, R. Chaudhary, E. Gilberth, et al., Kaiso overexpresison promotes intestinal inflammation and potentiates intestinal tumorigenesis in $Apc^{Min/+}$ mice, Biochim. Biophys. Acta 1852 (2015) 1846–1855.
[81] A.V. Prokhortchouk, O. Sansom, J. Selfridge, I.M. Caballero, S. Salozhin, D. Aithozhina, et al., Kaiso-deficient mice show resistance to intestinal cancer, Mol. Cell. Biol. 26 (2006) 199–208.
[82] M.A. Young, S. May, A. Damo, Y.S. Yoon, M.W. Hur, W. Swat, et al., Epigenetic regulation of Dlg1, via Kaiso, alters mitotic spindle polarity and promotes intestinal tumorigenesis, Mol. Cancer Res. 17 (2019) 686–696.
[83] S.-D. Dai, Y. Wang, Y. Miao, Y. Zhao, Y. Zhang, G.-Y. Jiang, et al., Cytoplasmic Kaiso is associated with poor prognosis in non-small cell lung cancer, BMC Cancer 9 (2009) 178.
[84] G. Jiang, Y. Wang, S. Dai, Y. Liu, M. Stoecker, E. Wang, P120-catenin isoforms 1 and 3 regulate proliferation and cell cycle of lung cancer cells via beta-catenin and Kaiso respectively, PLoS One 7 (2012) e30303.
[85] J. Jones, A. Mukherjee, B. Karanam, M. Davis, J. Jaynes, R.R. Reams, et al., African Americans with pancreatic ductal adenocarcinoma exhibit gender differences in Kaiso expression, Cancer Lett. 380 (2016) 513–522.
[86] A. Abisoye-Ogunniyan, H. Lin, A. Ghebremedhin, A.B. Salam, B. Karanam, S. Theodore, et al., Transcriptional repressor Kaiso promotes epithelial to mesenchymal transition and metastasis in prostate cancer through direct regulation of miR-200c, Cancer Lett. 431 (2018) 1–10.
[87] J. Jones, H. Wang, J. Zhou, S. Hardy, T. Turner, D. Austin, et al., Nuclear Kaiso indicates aggressive prostate cancers and promotes migration and invasiveness of prostate cancer cells, Am. J. Pathol. 181 (2012) 1836–1846.
[88] H. Wang, W. Liu, S. Black, O. Turner, J.M. Daniel, W. Dean-Colomb, et al., Kaiso, a transcriptional repressor, promotes cell migration and invasion of prostate cancer cells through regulation of miR-31 expression, Oncotarget 7 (2016) 5677–5689.
[89] B.I. Bassey-Archibong, L.G.A. Rayner, S.M. Hercules, C.W. Aarts, A. Dvorkin-Gheva, J.L. Bramson, et al., Kaiso depletion attenuates the growth and survival of triple negative breast cancer cells, Cell Death Dis. 8 (2017) e2689.
[90] D.-I. Koh, D. Han, H. Ryu, W.-I. Choi, B.-N. Jeon, M.-K. Kim, et al., KAISO, a critical regulator of p53-mediated transcription of CDKN1A and apoptotic genes, Proc. Natl. Acad. Sci. U. S. A. 111 (2014) 15078–15083.
[91] J.M. Kwiecien, B.I. Bassey-Archibong, W. Dabrowski, L.G.A. Rayner, A.R. Lucas, J.M. Daniel, Loss of Kaiso expression in breast cancer cells prevents intra-vascular invasion in the lung and secondary metastasis, PLoS One 12 (2017) e0183883.
[92] J. Feng, Upregulation of microRNA-4262 targets Kaiso (ZBTB33) to inhibit the proliferation and EMT of cervical cancer cells, Oncol. Res. 26 (2018) 1215–1225.
[93] R. Chaudhary, C.C. Pierre, K. Nanan, D. Wojtal, S. Morone, C. Pinelle, G.A. Wood, S. Robine, J.M. Daniel, The POZ-ZF transcription factor Kaiso (ZBTB33) induces inflamation and progenitor cell differentiation in the murine intestine, PLoS One 8 (2013) e74160.
[94] S.C. Robinson, R. Chaudhary, R. Jiménez-Saiz, L.G.A. Rayner, L. Bayer, M. Jordana, et al., Kaiso-induced

intestinal inflammation is preceded by diminished E-cadherin expression and intestinal integrity, PLoS One 14 (2019) e0217220.
[95] A.V. Kulikov, V.S. Korostina, E.A. Kulikova, D.V. Fursenko, A.E. Akulov, M.P. Moshkin, et al., Knockout Zbtb33 gene results in an increased locomotion, exploration and pre-pulse inhibition in mice, Behav. Brain Res. 297 (2016) 76–83.
[96] E.A. Kulikova, A.V. Kulikov, Kaiso protein in the regulation of brain and behavior, Curr. Protein Pept. Sci. 19 (2018) 692–698.
[97] P.J. Stogios, G.S. Downs, Jauhal JJS, S.K. Nandra, G.G. Privé, Sequence and structural analysis of BTB domain proteins, Genome Biol. 6 (2005) R82.
[98] A. Roussel-Gervais, I. Naciri, O. Kirsh, L. Kasprzyk, G. Velasco, G. Grillo, et al., Loss of the methyl-CpG-binding protein ZBTB4 alters mitotic checkpoint, inreases aneuploidy and promotes tumorigenesis, Cancer Res. 77 (2017) 62–73.
[99] R. Bomben, S. Gobessi, M. Dal Bo, S. Volinia, D. Marconi, E. Tissino, et al., The miR-17~92 family regulates the response to Toll-like receptor 9 triggering of CLL cells with unmutated IGHV genes, Leukemia 26 (2012) 1584–1593.
[100] W. Bu, Y. Wang, X. Min, MicroRNA-106b promotes the proliferation, migration and invasion of retinoblastoma cells by inhibiting the expression of ZBTB4 protein, Exp. Ther. Med. 16 (2018) 4537–4545.
[101] K. Kim, G. Chadalapaka, S.-O. Lee, D. Yamada, X. Sastre-Garau, P.-A. Defossez, et al., Identification of oncogenic microRNA-17-92/ZBTB4/specificty protein axis in breast cancer, Oncogene 31 (2011) 1034–1044.
[102] K.H. Kim, G. Chadalapaka, S.S. Pathi, U.-H. Jin, J.-S. Lee, Y.-Y. Park, et al., Induction of the transcriptional repressor ZBTB4 in prostate cancer cells by drug-induced targeting of microRNA-17-92/106b-25 clusters, Mol. Cancer Ther. 11 (2012) 1852–1862.
[103] M. Zhao, Z. Zhao, Concordance of copy number loss and down-regulation of tumor suppressor genes: a pan cancer study, BMC Genom. 17 (2016) 532.
[104] D. Yamada, R. Perez-Torrado, G. Filion, M. Caly, B. Jammart, V. Devignot, et al., The human protein kinase HIPK2 phosphorylates and downregulates the methyl-binding transcription factor ZBTB4, Oncogene 28 (2009) 2535–2544.
[105] G. Chadalapaka, I. Jutooru, S. Safe, Celastrol decreases specificity proteins (Sp) and fibroblast growth factor receptor-3 (FGFR3) in bladder cancer cells, Carcinogenesis 33 (2012) 886–894.
[106] W.W. Yang, G. Chadalapaka, S.-G. Cho, S. Lee, U.-H. Jin, I. Jutooru, et al., The transcriptional repressor ZBTB4 regulates EZH2 through microRNA-ZBTB4-specificity protein signaling axis, Neoplasia 16 (2014) 1059–1069.
[107] Y. Yu, R. Shang, Y. Chen, J. Li, Z. Liang, J. Hu, et al., Tumor suppressive ZBTB4 inhibits cell growth by regulating cell cycle progression and apoptosis in Ewing sarcoma, Biomed. Pharmacother. 100 (2018) 108–115.
[108] K. Karki, E. Hedrick, R. Kasiappan, U.-H. Jin, S. Safe, Piperlongumine induces reactive oxygen species (ROS)-dependent downregulation of specificity protein transcription factors, Cancer Prev. Res. 10 (2017) 467–477.
[109] R. Kasiappan, I. Jutooru, K. Mohankumar, K. Karki, A. Lacey, S. Safe, Reactive oxygen species (ROS)-inducing triterpenoid inhibits rhabdomyosarcoma cell and tumor growth through targeting Sp transcription factors, Mol. Cancer Res. 17 (2019) 794–805.
[110] S. Safe, MicroRNA-specificity protein (Sp) transcription factor interactions and significance in carcinogenesis, Curr. Pharmacol. Rep. 1 (2015) 73–78.
[111] A. Pozner, N.O. Hudson, J. Trewhella, T.W. Terooatea, S.A. Miller, B.A. Buck-Koehntop, The C-terminal zinc fingers of ZBTB38 are novel selective readers of DNA methylation, J. Mol. Biol. 430 (2018) 258–271.
[112] N.O. Hudson, F.G. Whitby, B.A. Buck-Koehntop, Structural insights into methylated DNA recognition by the C-terminal zinc fingers of the DNA reader protein ZBTB38, J. Biol. Chem. 293 (2018) 19835–19843.
[113] J. Kang, A.M. Kang, A system of linear equations for the identification of DNA binding affinity of zinc fingers, J. Biol. Chem. 294 (2019) 65.
[114] B. Miotto, C. Marchal, G. Adelmant, N. Guinot, P. Xie, J.A. Marto, et al., Stabilization of the methyl-CpG binding protein ZBTB38 by the deubiquitinase USP9X limits the occurrence and toxicity of oxidative stress in human cells, Nucleic Acids Res. 46 (2018) 4392–4404.
[115] Y. Shi, J. Sawada, G. Sui, B. Affar el, J.R. Whetstine, F. Lan, et al., Coordinated histone modifications mediated by a CtBP co-repressor complex, Nature 422 (2003) 735–738.
[116] C. Marchal, M. de Dieuleveult, C. Saint-Ruf, N. Guinot, L. Ferry, S.T. Olalla Saad, et al., Depletion of ZBTB38 potentiates the effects of DNA demethylating agents in cancer cells via CDKN1C mRNA up-regulation, Oncogenesis 7 (2018) 82.
[117] B. Miotto, M. Chibi, P. Xie, S. Koundrioukoff, H. Moolman-Smook, D. Pugh, et al., The RBBP6/ZBTB38/MCM10 axis regulates DNA replication and common fragile site stability, Cell Rep. 7 (2014) 575–587.
[118] E. Sloan, M.H. Tatham, M. Groslambert, M. Glass, A. Orr, R.T. Hay, et al., Analysis of the SUMO2 proteome during HSV-1 infection, PLoS Pathog. 11 (2015) e1005059.
[119] Z. Kote-Jarai, A.A. Olama, G.G. Giles, G. Severi, J. Schleutker, M. Weischer, et al., Seven prostate cancer susceptibility loci identified by a multi-stage genome-wide association study, Nat. Genet. 43 (2011) 785–791.
[120] D.F. Gudbjartsson, G.B. Walters, G. Thorleifsson, H. Stefansson, B.V. Halldorsson, P. Zusmaovich, et al., Many sequence variants affecting diversity of adult human height, Nat. Genet. 40 (2008) 609–615.
[121] T. Nishii, Y. Oikawa, Y. Ishida, M. Kawaichi, E. Matsuda, CtBP-interacting BTB zinc finger protein (CIBZ) promotes proliferation and G1/S transition in embryonic stem cells via Nanog, J. Biol. Chem. 287 (2012) 12417–12424.
[122] T. Kotoku, K. Kosaka, M. Nishio, Y. Ishida, M. Kawaichi, E. Matsuda, CIBZ regulates mesodermal and cardiac differentiation of by suppressing T and Mesp1 expression in mouse embryonic stem cells, Sci. Rep. 6 (2016) 34188.
[123] Y. Cai, J. Li, S. Yang, P. Li, X. Zhang, H. Liu, CIBZ, a novel BTB domain-containing protein, is involved in mouse spinal cord injury via mitochondrial pathway independent of p53 gene, PLoS One 7 (2012) e33156.
[124] Y. Oikawa, E. Matsuda, T. Nishil, Y. Ishida, M. Kawaichi, Down-regulation of CIBZ, a novel substrate of caspase-3, induces apoptosis, J. Biol. Chem. 283 (2008) 14242–14247.
[125] T. Ocskó, D.M. Tóth, G. Hoffmann, V. Tubak, T.T. Glant, T.A. Rauch, Transcription factor Zbtb38 downregulates the

expression of anti-inflammatory IL1r2 in mouse model of rheumatoid arthritis, Biochim. Biophys. Acta Gene Regul. Mech. 1861 (2018) 1040–1047.
[126] J. Jing, J. Liu, Y. Wang, M. Zhang, L. Yang, F. Shi, et al., The role of ZBTB38 in promoting migration and invasive growth of bladder cancer cells, Oncol. Rep. 41 (2019) 1980–1990.
[127] Y. Cai, J. Li, Z. Zhang, J. Chen, Y. Zhu, R. Li, et al., Zbtb38 is a novel target for spinal cord injury, Oncotarget 8 (2017) 45356–45366.
[128] J. Chen, L. Yan, H. Wang, Z. Zhang, D. Yu, C. Xing, et al., ZBTB38, a novel regulator of autophagy initiation targeted by RB1CC1/FIP200 in spinal cord injury, Gene 678 (2018) 8–16.
[129] M.B. Alonso, G. Zoidl, C. Taveggia, F. Bosse, C. Zoidl, M. Rahman, et al., Identification and characterization of ZFP-57, a novel zinc finger transcription factor in the mammalian peripheral nervous system, J. Biol. Chem. 279 (2004) 25653–25664.
[130] G. Ecco, M. Imbeault, D. Trono, KRAB zinc finger proteins, Development 144 (2017) 2719–2729.
[131] S. Quenneville, G. Verde, A. Corsinotti, A. Kapopoulou, J. Jakobsson, S. Offner, et al., Embryonic stem cells, ZFP57/KAP1 recognize a methylated hexanucleotide to affect chromatin and DNA methylation of imprinting control regions, Mol. Cell 44 (2011) 361–372.
[132] Y. Liu, H. Toh, H. Sasaki, X. Zhang, X. Cheng, An atomic model of Zfp57 recognition of CpG methylation within a specific DNA sequence, Genes Dev. 26 (2012) 2374–2379.
[133] R. Strogantsev, F. Krueger, K. Yamazawa, H. Shi, P. Gould, M. Goldman-Roberts, et al., Allele-specific binding of ZFP57 in the epigenetic regulation of imprinted and non-imprinted monoallelic expression, Genome Biol. 16 (2015) 112.
[134] H.P.J. Voon, R.J. Gibbons, Maintaining memory of silencing at imprinted differentially metylated regions, Cell. Mol. Life Sci. 73 (2016) 1871–1879.
[135] X. Li, P. Leder, Identifying genes preferentially expressed in undifferentiated embryonic stem cells, BMC Cell Biol. 8 (2007) 37.
[136] X. Li, M. Ito, F. Zhou, N. Youngson, X. Zuo, P. Leder, et al., A maternal-zygotic effect gene, Zfp57, maintains both maternal and paternal imprints, Dev. Cell 15 (2008) 547–557.
[137] Z. Anvar, M. Cammisa, V. Riso, I. Baglivo, H. Kukreja, A. Sparago, et al., ZFP57 recognizes multiple and closely spaced sequence motif variants to maintain repressive epigenetic marks in mouse embryonic stem cells, Nucleic Acids Res. 44 (2016) 1118–1132.
[138] X. Zuo, J. Sheng, H.-T. Lau, C.M. McDonald, M. Andrade, D.E. Cullen, 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. 287 (2012) 2107–2118.
[139] S.E. Boonen, D.J. Mackay, J.M. Hahnemann, L. Docherty, K. Grønskov, A. Lehmann, et al., Transient neonatal diabetes, ZFP57, and hypomethylation of multiple imprinted loci: a detailed follow-up, Diabetes Care 36 (2013) 505–512.
[140] D.J. Mackay, J.L. Callaway, S.M. Marks, H.E. White, C.L. Acerini, S.E. Boonen, et al., Hypomethylation of multiple imprinted loci in individuals with transient neonatal diabetes is associated with mutations in ZFP57, Nat. Genet. 40 (2008) 949–951.
[141] D. Leung, T. Du, U. Wagner, W. Xie, A.Y. Lee, P. Goyal, et al., Regulation of DNA methylaltion turnover at LTR retrotransposons and imprinted loci by the histone methyltransferase Setdb1, Proc. Natl. Acad. Sci. U. S. A. 111 (2014) 6690–6695.
[142] D.C. Schultz, K. Ayyanathan, D. Negorev, G.G. Maul, F. Rauscher Jr., SETDB1: a novel KAP-1-associated histone H3 lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatin genes by KRAB zinc-finger proteins, Genes Dev. 16 (2002) 919–932.
[143] S. Quenneville, P. Turelli, K. Bojkowska, C. Raclot, S. Offner, A. Kapopoulou, et al., The KRAB-ZFP/KAP1 system contributes to the eraly embryonic establishment of site-specific DNA methylation patterns maintained during development, Cell Rep. 2 (2012) 766–773.
[144] F. Fuks, P.J. Hurd, R. Depus, T. Kouzarides, The DNA methyltransferases associate with HP1 and the SUV39H1 histone methyltransferase, Nucleic Acids Res. 31 (2003) 2305–2312.
[145] B. Lehnertz, Y. Ueda, A.A. Derijck, U. Braunschweig, L. Perez-Burgos, S. Kubicek, et al., Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin, Curr. Biol. 13 (2003) 1192–1200.
[146] Y. Tada, Y. Yamaguchi, T. Kinjo, X. Song, T. Akagi, H. Takamura, et al., The stem cell transcription factor ZFP57 induces IGF2 expression to promote anchorage-independent growth in cancer cells, Oncogene 34 (2015) 752–760.
[147] Y. Shoji, H. Takamura, I. Ninomiva, S. Fushida, Y. Tada, T. Yokota, et al., The embryonic stem cell-specific transcription factor ZFP57 promotes liver matstasis of colorectal cancer, J. Surg. Res. 237 (2019) 22–29.
[148] L. Chen, X. Wu, H. Xie, N. Yao, Y. Xia, G. Ma, et al., ZFP57 suppress proliferation of breast cancer cells through down-regulation of MEST-mediated Wnt/β-catenin signalling pathway, Cell Death Dis. 10 (2019) 169.
[149] J.M. Shields, R.J. Christy, V.W. Yang, Identification and characterization of a gene encoding a gut-enriched Kruppel-like factor expressed during growth arrest, J. Biol. Chem. 271 (1996) 20009–20017.
[150] M.D. Ilsley, K.R. Gillinder, G.W. Magor, S. Huang, T.L. Bailey, M. Crossley, et al., Krüppel-like factors compete for promoters and enhancers to fine-tune transcription, Nucleic Acids Res. 45 (2017) 6572–6588.
[151] S.-F. Yet, M.M. McA'Nulty, S.C. Folta, H.-W. Yen, M. Yoshizumi, C.-M. Hseih, et al., Human EZF, a Krüppel-like zinc finger protein, is expressed in vascular endothelial cells and contains transcriptional activation and repression domains, J. Biol. Chem. 273 (1998) 1026–1031.
[152] D.E. Geiman, H. Ton-That, J.M. Johnson, V.W. Yang, Transactivation and growth suppression by the gut-enriched Krüppel-like factor (Krüppel-like factor 4) are dependent on acidic amino acid residues and protein-protein interaction, Nucleic Acids Res. 28 (2000) 1106–1113.
[153] A.M. Ghaleb, V.W. Yang, Krüppel-like factor 4 (KLF4): what we currently know, Gene 611 (2017) 27–37.
[154] X. Chen, H. Xu, P. Yuan, F. Fang, M. Huss, V.B. Vega, et al., Integration of external signaling pathways with the core

transcriptional network in embryonic stem cells, Cell 133 (2007) 1106–1117.
[155] S. Hu, J. Wan, Y. Su, Q. Song, Y. Zeng, H.N. Nguyen, et al., DNA methylation presents distinct binding sites for human transcription factors, eLife 2 (2013) e00726.
[156] C.G. Spruijt, F. Gnerlich, A.H. Smits, T. Pfaffeneder, P.W. Jansen, C. Bauer, et al., Dynamic readers for 5-(hydroxy)methylcytosine and its oxidized derivatives, Cell 152 (2013) 1146–1159.
[157] Y. Yin, E. Morgunova, A. Jolma, E. Kaasinen, B. Sahu, S. Khund-Sayeed, et al., Impact of cytosine methylation on DNA binding specificities of human transcription factors, Science 356 (2017) eaaj2239.
[158] H. Hashimoto, D. Wang, A.N. Steves, P. Jin, R.M. Blumenthal, X. Zhang, et al., Distinctive Klf4 mutants determine preference for DNA methylation status, Nucleic Acids Res. 44 (2016) 10177–10185.
[159] Y. Liu, Y.O. Olanrewaju, Y. Zheng, H. Hashimoto, R.M. Blumenthal, X. Zhangg, et al., Structural basis for Klf4 recognition of methylated DNA, Nucleic Acids Res. 42 (2014) 4859–4867.
[160] J. Wan, Y. Su, Q. Song, B. Tung, O. Oyinlade, S. Liu, et al., Methylated cis-regulatory elements mediate KLF4-dependent gene transactivation and cell migration, Elife 6 (2017) e20068.
[161] D.T. Dang, J. Pevsner, V.W. Yang, The biology of the mammalian Krüppel-like family of transcription factors, Int. J. Biochem. Cell Biol. 32 (2000) 1103–1121.
[162] P.M. Evans, C. Liu, Role of Krüppel-like factor 4 in normal homeostasis, cancer, and stem cells, Acta Biochim. Biophys. Sin. 40 (2008) 554–564.
[163] M.K. Farrugia, D.B. Vanderbilt, M.A. Salkeni, J.M. Ruppert, Kruppel-like pluripotency factors as Modulators of cancer cell therapeutic responses, Cancer Res. 76 (2016) 1677–1682.
[164] B.D. Rowland, D.S. Pepper, KLF4, p21 and context-dependent opposing forces in cancer, Nat. Rev. Cancer 6 (2006) 11–23.
[165] K. Takahashi, S. Yamanaka, Induction of pluripotentstem cells from mouse embryonic and adult fibroblastcultures by defined factors, Cell 126 (2006) 663–676.
[166] K. Takahashi, K. Tanabe, M. Ohnuki, M. Narita, T. Ichisaka, K. Tomoda, et al., Induction of pluripotent stem cells from adult human fibroblasts by defined factors, Cell 131 (2007) 861–872.
[167] E.A. El-Karim, E.G. Hagos, A.M. Ghaleb, B. Yu, V.W. Yang, Krüppel-like factor 4 regulates genetic stability in mouse embryonic fibroblasts, Mol. Cancer 12 (2013) 89.
[168] E.G. Hagos, A.M. Ghaleb, W.B. Dalton, A.B. Bialkowska, V.W. Yang, Mouse embryonic fibroblasts null for the Krüppel-like factor 4 gene are unstable, Oncogene 28 (2009) 1197–1205.
[169] X. Chen, E.M. Whitney, S.Y. Gao, V.W. Yang, Transcriptional profiling of Krüppel-like factor 4 reveals a function in cell cycle regulation and epitheleal differentiation, J. Mol. Biol. 326 (2003) 665–677.
[170] L. Shie J, Z.Y. Chen, M. Fu, R.G. Pestell, C.C. Tseng, Gut-enriched Krüppel-like factor represses cyclin D1 promoter activity through Sp1 motif, Nucleic Acids Res. 28 (2000) 2969–2976.
[171] H.S. Yoon, A.M. Ghaleb, M.O. Nandan, I.M. Hisamuddin, W.B. Dalton, V.W. Yang, Krüppel-like factor 4 prevents centrosome amplification following gamma-irradiation-induced DNA damage, Oncogene 24 (2005) 4017–4025.
[172] N. Ky, C.B. Lim, J. Li, J.P. Tam, M.S. Hamza, Y. Zhao, KLF4 suppresses HDACi induced caspase activation and the SAPK pathway by trageting p57(Kip2), Apoptosis 14 (2009) 1095–1107.
[173] C. Liu, E.P. DeRoo, C. Stecyk, M. Wolsey, M. Szuchnicki, E.G. Hagos, Impaired autophagy in mouse embryonic fibroblasts null for Krüppel-like factor 4 promotes DNA damage and increases apoptosis upon serum starvation, Mol. Cancer 14 (2015) 101.
[174] B. Wang, M.Z. Zhao, N.P. Cui, D.D. Lin, A.Y. Zhang, Y. Qin, et al., Krüppel-like factor 4 induces apoptosis and inhibits tumorigenic progression in SK-BR-3 breast cancer cells, FEBS Open Bio. 5 (2015) 147–154.
[175] N.C. Whitloock, J.H. Bahn, S.H. Lee, T.E. Eling, S.J. Baek, Resveratrol-induced apoptosis is mediated by early growth response-1, Krüppel-like factor 4, and activating transcription factor 3, Cancer Prev. Res. 4 (2011) 116–127.
[176] B.D. Rowland, R. Bernards, D.S. Peeper, The KLF4 tumour suppressor is a transcriptional repressor of p53 that acts as a context-dependent oncogene, Nat. Cell Biol. 7 (2005) 1074–1082.
[177] B.D. Rowland, D.S. Peeper, KLF4, p21 and context-dependent opposing forces on cancer, Nat. Rev. Cancer 6 (2006) 11–23.
[178] W. Hu, W.L. Hofstetter, H. Li, Y. Zhou, Y. He, A. Pataer, et al., Putative tumor-suppressive function of Kruppel-like factor 4 in primary lung carcinoma, Clin. Cancer Res. 15 (2009) 5688–5695.
[179] D. Wei, W. Gong, M. Kanai, C. Schlunk, L. Wang, J.C. Yao, et al., Drastic down-regulation of Krüppel-like factor 4 expression is critical in human gastric cancer development and progression, Cancer Res. 65 (2005) 2746–2754.
[180] N. Zhang, J. Zhang, Z.W. Wang, L. Zha, Z. Huang, Altered expression of Krüppel-like factor 4 and beta-catenin in human gastric cancer, Oncol. Lett. 3 (2012) 1017–1022.
[181] W. Zhao, I.M. Hisamuddin, M.O. Nandan, B.A. Babbin, N.E. Lamb, V.W. Yang, Identification of Krüppel-like factor 4 as a potential tumor suppressor gene in colorectal cancer, Oncogene 23 (2004) 395–402.
[182] Y.J. Chen, C.Y. Wu, C.C. Chang, C.J. Ma, M.C. Li, C.M. Chen, Nuclear Krüppel-like factor 4 expression is associated with human skin squamous cell carcinoma progression and metastasis, Cancer Biol. Ther. 7 (2008) 777–782.
[183] S.K. Tai, M.H. Yang, S.Y. Chang, Y.C. Chang, W.Y. Li, T.L. Tsai, et al., Persistent Krüppel-like factor 4 expression predicts progression and poor prognosis of head and neck squamous cell carcinoma, Cancer Sci. 102 (2011) 895–902.
[184] F. Yu, J. Li, H. Chen, J. Fu, S. Ray, S. Huang, et al., Kruppel-like factor 4 (KLF4) is required for maintenance of breast cancer stem cells and for cell migration and invasion, Oncogene 30 (2011) 2161–2172.
[185] K.M. Call, T. Glaser, C.Y. Ito, A.J. Buckler, J. Pelletier, D.A. Haber, et al., Isolation and characterization of a zinc finger polypeptide gene at the human chromosome 11 Wilms' tumor locus, Cell 60 (1990) 509–520.
[186] D.A. Haber, A.J. Buckler, T. Glaser, K.M. Call, J. Pelletier, R.L. Sohn, et al., An internal deletion within an 11p13 zinc finger gene contributes to the development of Wilms' tumor, Cell 61 (1990) 1257–1269.
[187] S.G. Roberts, Transcriptional regulation by WT1 in development, Curr. Opin. Genet. Dev. 15 (2005) 542–547.

[188] Z.Y. Wang, Q.Q. Qiu, T.F. Deuel, The Wilms' tumor gene product WT1 activates or suppresses transcription through separate functional domains, J. Biol. Chem. 268 (1993) 9172–9175.
[189] D.A. Haber, R.L. Sohn, A.J. Buckler, J. Pelletier, K.M. Call, D.E. Housman, Alternative splicing and genomic structure of the Wilms' tumor gene WT1, Proc. Natl. Acad. Sci. U. S. A. 88 (1991) 9618–9622.
[190] H. Hashimoto, Y.O. Olanrewaju, Y. Zheng, G.G. Wilson, X. Zhang, X. Cheng, Wilms tumor protein recognizes 5-carboxylcytosine within a specific DNA sequence, Genes Dev. 28 (2014) 2304–2313.
[191] R. Stoll, B.M. Lee, E.W. Debler, J.H. Laity, I.A. Wilson, H.J. Dyson, et al., Structure of the Wilms tumor suppressor protein zinc finger domain bound to DNA, J. Mol. Biol. 372 (2007) 1227–1245.
[192] A. Caricasole, A. Duarte, S.H. Larsson, N.D. Hastie, M. Little, G. Holmes, et al., RNA binding by the Wilms tumor suppressor zinc finger proteins, Proc. Natl. Acad. Sci. U. S. A. 93 (1996) 7562–7566.
[193] T. Ullmark, G. Montano, U. Gullberg, DNA and RNA binding by the Wilms' tumour gene 1 (WT1) protein +KTS and KTS isoforms-from initial observations to recent global genomic analyses, Eur. J. Haematol. 100 (2018) 229–240.
[194] F.J. Rauscher 3rd, J.F. Morris, O.E. Tournay, D.M. Cook, T. Curran, Binding of the Wilms' tumor locus zinc finger protein to the EGR-1 consensus sequence, Science 250 (1990) 1259–1262.
[195] M. Kann, S. Ettou, Y.L. Jung, M.O. Lenz, M.E. Taglienti, P.J. Park, et al., Genome-wide analysis of Wilms' tumor 1-controlled gene expression in podocytes reveals key regulatory mechanisms, J. Am. Soc. Nephrol. 26 (2015) 2097–2104.
[196] F.J. Motamedi, D.A. Badro, M. Clarkson, M. Rita Lecca, S.T. Bradford, F.A. Buske, et al., WT1 controls antagonistic FGF and BMP-pSMAD pathways in early renal progenitors, Nat. Commun. 5 (2014) 4444.
[197] T. Ullmark, L. Jarvstråt, C. Sandén, G. Montano, H. Jernmark-Nilsson, H. Lilljebjorn, et al., Distinct global binding patterns of the Wilms tumor gene 1 (WT1) KTS and +KTS isoforms in leukemic cells, Haematologica 102 (2017) 336–345.
[198] H. Nakagama, G. Heinrich, J. Pelletier, D.E. Housman, Sequence and strcutural requirements for high-affinity DNA binding by the WT1 gene product, Mol. Cell. Biol. 15 (1995) 1489–1498.
[199] S. Ito, L. Shen, Q. Dai, S.C. Wu, L.B. Collins, J.A. Swenberg, et al., Tet proteins can convert 5-methylcytosine to 5-formylsytosine and 5-carboxylcytosine, Science 333 (2011) 1300–1303.
[200] M. Tahiliani, K.P. Koh, Y. Shen, W.A. Pastor, H. Bandukwala, Y. Brudno, et al., Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1, Science 324 (2009) 930–935.
[201] Y. Wang, M. Xiao, X. Chen, L. Chen, Y. Xu, L. Lv, et al., WT1 recruits TET2 to regulate its target gene expression and suppress leukemia cell proliferation, Mol. Cell 57 (2015) 662–673.
[202] S.B. Lee, D.A. Haber, Wilms tumor and the WT1 gene, Exp. Cell Res. 264 (2001) 74–99.
[203] P. Niaudet, M.C. Gubler, WT1 and glomerular diseases, Pediatr. Nephrol. 21 (2006) 1653–1660.
[204] K.D. Wagner, N. Wagner, A. Schedl, The complex life of WT1, J. Cell Sci. 116 (2003) 1653–1658.
[205] N.D. Hastie, Wilms' tumour 1 (WT1) in development, homeostasis and disease, Development 144 (2017) 2862–2872.
[206] P. Hohenstein, N.D. Hastie, The many facets of the Wilm's tumor gene, WT1, Hum. Mol. Genet. 15 (2006) R196–R201.
[207] X.W. Qi, F. Zhang, H. Wu, J.L. Liu, B.G. Zong, C. Xu, et al., Wilms' tumor 1 (WT1) expression and prognosis in solid cancer patients: a systematic review and meta-analysis, Sci. Rep. 5 (2015) 8924.
[208] N. Tatsumi, Y. Oji, N. Tsuji, A. Tsuda, M. Higashio, S. Aoyagi, et al., Wilms' tumor gene WT1-shRNA as a potent apoptosis-inducing agent for solid tumors, Int. J. Oncol. 32 (2008) 701–711.
[209] A.M. Beckmann, P.A. Wilce, Egr transcription factors in the nervous system, Neurochem. Int. 31 (1997) 477–510.
[210] F. Duclot, M. Kabbaj, The role of Early Growth Response 1 (EGR1) in brain plasticity and neuropsychiatric disorders, Front. Behav. Neurosci. 11 (2017) 35.
[211] A. Veyrac, A. Besnard, J. Caboche, S. Davis, S. Laroche, The transcription factor Zif268/Egr1, brain plasticity, and memory, Prog. Mol. Biol. Transl. Sci. 122 (2014) 89–129.
[212] J.I. Pagel, E. Deindl, Early growth response 1 – a transcription factor in the crossfire of signal transcduction cascades, Indian J. Biochem. Biophys. 48 (2011) 226–235.
[213] J. Milbrandt, A nerve growth factor-induced gene encodes a possible transcriptional regulatory factor, Science 238 (1987) 797–799.
[214] B.A. Christy, L.F. Lau, D. Nathans, A gene activated in mouse 3T3 cells by serum growth factors encodes a protein with "zinc finger" sequences, Proc. Natl. Acad. Sci. U. S. A. 85 (1988) 7857–7861.
[215] P. Lemaire, O. Revelant, R. Bravo, P. Charnay, Two mouse genes encoding potential transcription factors with identical DNA-binding domains are activated by growth factors in cultured cells, Proc. Natl. Acad. Sci. U. S. A. 85 (1988) 4691–4695.
[216] R.W. Lim, B.C. Varnum, H.R. Herschman, Cloning of tetradecanoyl phorbol ester-induced "primary response" sequences and their expression in density-arrested Swiss 3T3 cells and a TPA non-proliferative variant, Oncogene 1 (1987) 263–270.
[217] R.W. Lim, B.C. Varnum, T.G. O'Brien, H.R. Herschman, Induction of tumor promotor-inducible genes in murine 3T3 cell lines and tetradecanoyl phorbol acetate-nonproliferative 3T3 variants can occur through protein kinase C-dependent and -independent pathways, Mol. Cell. Biol. 9 (1989) 1790–1793.
[218] V.P. Sukhatme, X.M. Cao, L.C. Chang, C.H. Tsai-Morris, D. Stamenkovich, P.C. Ferreira, et al., A zinc finger-encoding gene coregulated with c-fos during growth and differentiation, and after cellular depolarization, Cell 53 (1988) 37–43.
[219] B. Bozon, S. Davis, S. Laroche, A requirement for the immediate early gene zif268 in reconsolidation of recognition memory after retrieval, Neuron 40 (2003) 695–701.
[220] L.M. Khachigian, V. Lindner, A.J. Williams, T. Collins, Egr-1-induced endothelial gene expression: a common theme in vascular injury, Science 271 (1996) 1427–1431.

[221] J.L. Lee, B.J. Everitt, K.L. Thomas, Independent cellular processes for hippocampal memory consolidation and reconsolidation, Science 304 (2004) 839–843.
[222] E. Sanchez-Guerrero, S.R. Jo, B.H. Chong, L.M. Khachigian, EGFR and the complexity of receptor crosstalk in the cardiovascular system, Curr. Mol. Med. 13 (2013) 3–12.
[223] L. Zandarashvili, M.A. White, A. Esadze, J. Iwahara, Structural impact of complete CpG methylation within target DNA on specific complex formation of the inducible transcription factoer Egr-1, FEBS Lett. 589 (2015) 1748–1753.
[224] B.A. Buck-Koehntop, P.-A. Defossez, On how mammalian transcription factors recognize methylated DNA, Epigenetics 8 (2013) 131–137.
[225] F. Lu, H.-T. Zhang, DNA methylation and nonsmall cell lung cancer, Anat. Rec. 294 (2011) 1787–1795.
[226] M.F. Ritchie, C. Yue, Y. Zhou, P.J. Houghton, J. Soboloff, Wilms tumor suppressor 1 (WT1) and early growth response 1 (EGR1) are regulators of STIM1 expression, J. Biol. Chem. 285 (2010) 10591–10596.
[227] V. Scharmhorst, A.L. Menke, J. Attema, J.K. Haneveld, N. Riteco, G.J. van Steenbrugge, et al., EGR-1 enhances tumor growth and modulates the effect of the Wilms' tumor 1 gene products on tumorigenicity, Oncogene 19 (2000) 791–800.
[228] A. Blattler, P.J. Farnham, Cross-talk between site-specific transcription factors and DNA methylation states, J. Biol. Chem. 288 (2013) 34287–34294.
[229] A. Kubosaki, Y. Tomaru, M. Tagami, E. Arner, H. Miura, T. Suzuki, et al., Genome-wide investigation of in vivo EGR-1 binding sites in monocytic differentiation, Genome Biol. 10 (2009) R41.
[230] C.A. Kemme, D. Nguyen, A. Chattopadhyay, J. Iwahara, Regulation of transcription factors via natural decoys in genomic DNA, Transcription 7 (2016) 115–120.
[231] C.A. Kemme, A. Esadze, J. Iwahara, Influence of quasi-specific sites on kinetics of target DNA search by a sequence-specific DNA-binding protein, Biochemistry 54 (2015) 6684–6691.
[232] C.A. Kemme, R. Marquez, R.H. Luu, J. Iwahara, 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. 45 (2017) 7751–7759.
[233] V.V. Lobanenkov, R.H. Nicolas, V.V. Adler, H. Paterson, E.M. Klenova, A.V. Polotskaja, et al., A novel sequence-specific DNA binding protein which interacts with three regularly spaced direct repeats of the CCCTC-motif in the 5′-flanking sequence of the chicken c-myc gene, Oncogene 5 (1990) 1743–1753.
[234] J. Ausió, P.T. Georgel, MeCP2 and CTCF: enhancing the cross-talk of silencers, Biochem. Cell Biol. 95 (2017) 593–608.
[235] L. Braccioli, E. de Wit, CTCF: a Swiss-army knife for genome organization and transcription regulation, Essays Biochem. 63 (2019) 157–165.
[236] S. Kim, N.-K. Yu, B.-K. Kaang, CTCF is a multifunctional protein in genome regularion and gene expression, Exp. Mol. Med. 47 (2015) e166.
[237] J. Zlatanova, P. Caiafa, CTCF and its protein partners: divide and rule? J. Cell Sci. 122 (2009) 1275–1284.
[238] J.E. Phillips, V.G. Corces, CTCF: master weaver of the genome, Cell 137 (2009) 1194–1211.
[239] H. Hashimoto, D. Wang, J.R. Horton, X. Zhang, V.G. Corces, X. Cheng, Structural basis for the versatile and methylation-dependent binding of CTCF to DNA, Mol. Cell 66 (2017) 711–720.
[240] H.S. Rhee, B.J. Pugh, Comprehensive genome-wide protein-DNA interactions Detected at single-nucleotide resolution, Cell 147 (2011) 1408–1419.
[241] H. Nakahashi, K.-R.K. Kwon, W. Resch, L. Vian, M. Dose, D. Stavreva, et al., A genome-wide map of CTCF multivalency redefines the CTCF code, Cell Rep. 3 (2013) 1678–1689.
[242] H. Wang, M.T. Maurano, H. Qu, K.E. Varley, J. Gertz, F. Pauli, et al., Widespread plasticity in CTCF occupancy linked to DNA methylation, Genome Res. 22 (2012) 1680–1688.
[243] A.C. Bell, G. Felsenfeld, Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene, Nature 405 (2000) 482–485.
[244] A.T. Hark, C.J. Schoenherr, D.J. Katz, R.S. Ingram, J.M. Levorse, S.M. Tilghman, CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus, Nature 405 (2000) 486–489.
[245] M. Rende, I. Baglivo, B. Burgess-Beusse, S. Esposito, R. Fattorusso, G. Felsenfeld, et al., Critical DNA binding interactions of the insulator protein CTCF: a small number of zinc fingers mediate strong binding, and a single finger-DNA interaction controls binding at imprinted loci, J. Biol. Chem. 282 (2007) 33336–33345.
[246] G.N. Filippova, C.P. Thienes, B.H. Penn, D.H. Cho, Y.J. Hu, J.M. Moore, et al., CTCF-binding sites flank CTG/CAG repeats and form a methylation-sensitive insulator at the DM1 locus, Nat. Genet. 28 (2001) 335–343.
[247] A. Klug, The discovery of zinc fingers and their applications in gene regulation and genome manipulation, Annu. Rev. Biochem. 79 (2010) 213–231.
[248] S.A. Wolfe, L. Nekludova, C.O. Pabo, DNA recognition by Cys_2His_2 zinc finger proteins, Annu. Rev. Biophys. Biomol. Struct. 3 (1999) 183–212.
[249] Y. Liu, R.M. Blumenthal, X. Cheng, A common mode of recognition for methylated CpG, Trends Biochem. Sci. 38 (2013) 177–183.
[250] H. Hashimoto, X. Zhang, Y. Zheng, G.G. Wilson, X. Cheng, Denys-Drash syndrome ssociated WT1 glutamine 369 have altered sequence-preferences and altered responses to epigenetic modifications, Nucleic Acids Res. 44 (2016) 10165–10176.
[251] Y. Liu, Y.O. Olanrewaju, Y. Zhang, X. Cheng, DNA recognition of 5-carboxylcytosine by a Zfp57 mutant at an atomic resolution of 0.97 Å, Biochemistry 52 (2013) 9310–9317.
[252] P.S. Ho, G.J. Quigley, R.F.J. Tilton, A. Rich, Hydration of methylated and nonmethylated B-DNA and Z-DNA, J. Phys. Chem. 92 (1988) 939–945.
[253] C. Mayer-Jung, D. Moras, Y. Timsit, Hydration and regulation of methylated CpG steps in DNA, EMBO J. 17 (1998) 2709–2718.
[254] K.L. Ho, I.W. McNae, L. Schmiedeberg, R.J. Klose, A.P. Bird, M.D. Walkinshaw, MeCP2 binding to DNA depends upon hydration of methyl-CpG, Mol. Cell 29 (2008) 525–531.
[255] J. Otani, K. Arita, T. Kato, M. Kinoshita, H. Kimura, I. Suetake, et al., Structural basis of the versatile DNA recognition ability of the methyl-CpG binding domain of

methyl-CpG binding domain protein 4, J. Biol. Chem. 288 (2013) 6351–6362.
[256] A. Lazarovici, T. Zhou, A. Shafer, A.C. Dantas Machado, T.R. Riley, R. Sandstrom, et al., Probing DNA shape and methylation state on a genomic scale with DNase I, Proc. Natl. Acad. Sci. U. S. A. 110 (2013) 6376–6381.
[257] R. Lavery, M. Moakher, J.H. Maddocks, D. Petkeviciute, K. Zakrzewska, Conformational analysis of nucleic acids revisited: curves+, Nucleic Acids Res. 37 (2009) 5917–5929.
[258] H. Zhu, G. Wang, J. Qian, Transcription factors as readers and effectors of DNA methylation, Nat. Rev. Genet. 17 (2016) 551–565.
[259] T.L. Bailey, M. Bodén, F.A. Buske, M. Frith, C.E. Grant, L. Clementi, et al., MEME SUITE: tools for motif discovery and searching, Nucleic Acids Res. 37 (2009) W202–W208.