STUDIES OF ZINC FINGER PROTEINS IN EPIGENETIC GENE REGULATION

NAN LIU

MÜNCHEN 2014

STUDIES OF ZINC FINGER PROTEINS IN EPIGENETIC GENE REGULATION

NAN LIU

Dissertation

an der Fakultät für Biologie

der Ludwig-Maximilians-Universität

München

vorgelegt von

Nan Liu

aus Yushu, China

München, den 2014

Erstgutachter: Prof. Dr. Heinrich Leonhardt

Zweitgutachter: PD Dr. Anna Friedl

Tag der mündlichen Prüfung: 04.12.2014

Contents

Content

Summary

Summary

Epigenetic regulation of gene expression is achieved through different mechanisms such as DNA methylation and histone modifications. Zinc finger proteins, among the most abundant proteins in eukaryotes, play a pivotal role in setting these epigenetic marks. In this study, we investigated the roles of the zinc finger proteins DNMT1, UHRF1 and TET3 in epigenetic gene regulation.

DNA methylation has long been regarded as a stable mark mediating gene repression, but the discovery of TET enzymes that oxidize methylcytosine has kindled the idea that DNA modifications are more diverse. To date, little is known about how TET proteins are targeted to specific genomic loci. In this work, we identified and characterized an alternative mouse TET3 isoform ($TET3^{CXXC}$) harboring a CXXC type zinc finger domain and also found associations between TET enzymes and the zinc finger protein CXXC4. Relative transcript levels suggest that distinct ratios of $TET3^{CXXC}$ and the TET3-CXXC4 complex may be present in different adult tissues. We propose that variable associations with CXXC modules may contribute to context specific functions of TET proteins.

DNMT1 is a DNA methyltransferase which harbors a CXXC type zinc finger and is responsible for maintaining DNA methylation patterns after DNA replication. Its interaction partner UHRF1 is essential for the propagation of DNA methylation by recruiting DNMT1 to the hemimethylated DNA substrate. However, the targeting mechanism of DNMT1 by UHRF1 is still not fully understood. In this study, we describe that UHRF1 ubiquitinates histone H3 depending on the PHD and RING zinc finger motifs, which provides a docking site for DNMT1 binding and thereby maintains DNA methylation. Therefore, DNMT1 not only copies existing DNA methylation patterns but rather integrates multiple signals from different epigenetic pathways.

To gain further insights into the cellular targets of the E3 ligase UHRF1, we developed a ubiquitination substrate assay and identified UHRF1-dependent targets related to different regulatory pathways. We show that UHRF1 ubiquitinates the heterochromatin protein CBX1 for proteasomal degradation, a process which is counteracted by USP7-mediated deubiquitination. This dynamic modification of CBX1 likely contributes to heterochromatin formation, thus providing a novel mechanism for epigenetic regulation.

In summary, zinc finger proteins fundamentally contribute to multiple layers of epigenetic regulation, and function as essential factors in epigenetic networks.

1 Introduction

1 Introduction

1.1 Epigenetic regulation

All cells of a multicellular organism are genetically homogeneous, but they show structural and functional heterogeneity due to the differential expression of genes (Jaenisch and Bird 2003). Gene expression patterns are set and maintained over numerous subsequent cell generations, and epigenetic regulation contributes to the establishment of cell-type specific gene expression leading to a broad range of functional and morphological diversity.

The term “epigenetics” was introduced by Waddington as a portmanteau of “epigenesis” and “genetics” in 1940s (Waddington 1942, 2012). In the original sense of its definition, epigenetics referred to “the branch of biology which studies the causal interactions between genes and their products, which brings the phenotype into being” (Waddington 1968). Over the years, with the observation of numerous biological phenomena related to epigenetics and the rapid growth of genetics research, epigenetics has evolved to a defined field of study. Nowadays, epigenetics is generally accepted as the study of changes in gene expression or cellular phenotype which occur without changes in the underlying DNA sequence (Goldberg, Allis et al. 2007). Notably, epigenetics contributes both the heritable changes in gene activity and also stable, long-term alterations in the transcriptional potential of a cell that are not necessarily heritable (Dupont, Armant et al. 2009).

Epigenetic mechanisms are thought to regulate gene expression by altering chromatin condensation and accessibility. Epigenetic modulators establish transcriptionally active or silent chromatin states at different layers, including DNA and histone modifications, non-coding RNA, nucleosome positioning as well as histone variants, building well-controlled transcriptional regulation networks (Fig. 1). Among these epigenetic marks, DNA methylation and histone modifications are the most extensively characterized.

1.1.1 DNA methylation

There are many ways controlling gene expression in eukaryotes, and DNA methylation is the most commonly studied chromatin mark that regulates gene activity in the

mammalian genome. In vertebrates, DNA methylation refers to the addition of a methyl group to the carbon 5 position of cytosine residues to form 5 methylcytosine (5mC). DNA methylation level differs strongly between species, for example, cytosine methylation is about 34% in *Arabidopsis*, 2.3% in *Escherichia coli (E. coli)*, 0.03% in *Drosophila*, and virtually none (< 0.0002%) in yeast species (Cokus, Feng et al. 2008, Capuano, Mulleder et al. 2014). DNA methylation levels of mammals are intermediate. In mammalian somatic cells, 5mC accounts for about 4% of cytosines and approximately 70%-80% of CpG dinucleotides throughout the genome are methylated (Ehrlich, Gama-Sosa et al. 1982). While cytosine methylation occurs almost exclusively at symmetric CpG dinucleotides in somatic tissues, non-CpG methylation is quite prevalent in embryonic stem cells (ESCs) accounting for 15%-20% of total cytosine methylation but its function is still unclear (Ramsahoye, Biniszkiewicz et al. 2000, Bird 2002, Lister, Pelizzola et al. 2009). The unmethylated CpG dinucleotides are mostly found at gene promoters in dense clusters termed “CpG islands”, maintaining transcription of active genes (Suzuki and Bird 2008, Cedar and Bergman 2009, Law and Jacobsen 2010).

Methylation of cytosines may repress gene expression by directly preventing the binding of transcription factors to their recognition targets (Becker, Ruppert et al. 1987, Bednarik, Duckett et al. 1991, Prendergast and Ziff 1991, Di Fiore, Palena et al. 1999). Many transcription factors show reduced binding preference to methylated target promoter sequences compared to unmethylated ones (Zhang and Pradhan 2014). For instance, YY1 is a ubiquitously distributed transcription factor which is involved in repressing and activating a diverse number of promoters. YY1 binding site in *peg3* gene locus is a conserved sequence element located in the first intron, which is involved in transcription and imprinting control. This binding site contains one CpG site and the methylation of this site is sufficient to abolish the binding activity of YY1 *in vitro* (Kim, Kollhoff et al. 2003).

Another mechanism involving DNA methylation in gene silencing is by the recruitment of binding proteins that preferentially recognize methylated DNA, which either block gene activation directly or indirectly via further recruitment of repressive protein complexes (Boyes and Bird 1991, Jones, Veenstra et al. 1998, Nan, Ng et al. 1998, Newell-Price, Clark et al. 2000). One such family are the methyl-CpG binding domain proteins (MBDs) that

mainly bind methylated CpG dinucleotides either acting as insulator for transcription factor binding or recruiting histone deacetylases, lysine methyltransferases or chromatin remodeling complexes which induce the formation of silent chromatin (Bakker, Lin et al. 2002, Fujita, Watanabe et al. 2003).

The biological functions of DNA methylation are fundamentally different in prokaryotes and eukaryotes. In bacteria, DNA methylation is closely involved in restriction modification system and acts as primitive immune system, allowing hosts to protect themselves from infection by foreign DNAs (Colot and Rossignol 1999, Bird 2002, Chen and Li 2004). Unlike bacteria, high DNA methylation in plants leads to the transcriptional silencing of transposable elements and other repetitive DNA sequences to maintain genome stability (Zhang, Yazaki et al. 2006, He, Chen et al. 2011).

In mammals, DNA methylation is a relative stable epigenetic mark and is critical for a variety of biological processes including early embryogenesis (Li, Bestor et al. 1992, Watanabe, Suetake et al. 2002, Haaf 2006), cell differentiation (Latham, Gilbert et al. 2008), gene and transposon silencing (Chen, Pettersson et al. 1998), genomic imprinting (Reik and Walter 2001, Ideraabdullah, Vigneau et al. 2008), and X chromosome inactivation (Heard and Disteche 2006, Senner and Brockdorff 2009). Furthermore, DNA methylation is also involved in regulating neuronal development and tumorigenesis (Feinberg, Ohlsson et al. 2006, Dulac 2010).

1.1.2 Histone modifications

Besides DNA methylation, histones modifications have also been implicated in the epigenetic gene regulation. In eukaryotic cells, there are two forms of chromatin that reflect the level of transcription activity of the cell: euchromatin and heterochromatin. Euchromatin is open chromatin with less density so that the DNA is accessible to the transcription machinery and can be actively transcribed. Conversely, heterochromatin is packaged into a highly condensed form which is not accessible to gene regulatory molecules and thus is silenced. The basic unit of chromatin are nucleosomes, each consisting of 146 base pairs of DNA wrapped in almost two turns around a histone

octamer made up of two H2A-H2B dimers and a H3-H4 tetramer (Davey, Sargent et al. 2002). Adjacent nucleosomes are connected by linker DNA and linker histones (Thoma et al. 1979; Luger et al. 1997). These core histones are relatively similar in structure and are highly conserved through evolution. Each of the histone proteins consists of a structured core and a unstructured tail domain (Biswas, Voltz et al. 2011). Histone tails are the common sites of posttranslational modifications (PTMs), including methylation and acetylation of lysines and arginines, phosphorylation of serines and threonines, ubiquitinylation and SUMOylation of lysines.

In comparison to DNA methylation, posttranslational modifications on histone tails are highly dynamic, which play a key role in the regulation of chromatin accessibility in eukaryotes (Cosgrove, Boeke et al. 2004, Kouzarides 2007, Biswas, Voltz et al. 2011). Some PTMs may directly influence the mobility and stability of nucleosome, therefore affect chromatin structure. For instance, acetylation on lysine residues neutralizes the positive charge of histones and reduces electrostatic attraction between histones and negatively charged DNA, therefore loosening the chromatin structure (Workman and Kingston 1998). Acetylation of H3 at K56 has been found to directly affect histone-DNA contacts (Masumoto, Hawke et al. 2005, Ozdemir, Spicuglia et al. 2005, Xu, Zhang et al. 2005). K56Q mutation designed to mimic H3K56 acetylation has been observed to alter nucleosome mobility and the wrapping of DNA around nucleosomes, which support the hypothesis that modifications can directly alter chromatin structure and dynamics (Masumoto, Hawke et al. 2005, Ferreira, Somers et al. 2007). Moreover, another mechanism by which PTMs of histones regulate gene expression is as a signal platform to recruit effector modules to local chromatin, so that functional outcome of PTMs is mostly determined by its readers (Ferreira, Somers et al. 2007, Yun, Wu et al. 2011). For example, heterochromatin protein 1 (HP1) is a family of heterochromatic adaptor molecules and is implicated in gene silencing (Wallrath 1998, Jones, Cowell et al. 2000). The chromodomain of HP1 specifically interacts with methylated H3 at lysine 9 (H3K9) and targets HP1 to heterochromatic regions, which leads to the repression of gene transcription (Lachner, O'Carroll et al. 2001, Nakayama, Rice et al. 2001). Thus, in the best characterized modifications of H3 and H4, generally acetylation modifications are

associated with transcriptional activation (Allfrey, Faulkner et al. 1964), while methylation of lysine residues present two opposite transcriptional activity: activation or repression, depending on which amino acids or to what extent the residue is modified (Norton, Imai et al. 1989).

Besides DNA methylation and histone modifications, noncoding RNAs are also reported as another layer of epigenetic regulation. Some noncoding RNAs have been demonstrated to interact with chromatin to modulate large-scale gene expression programs (Vance and Ponting 2014). In addition, nucleosome positioning and histone variants are also emerging as regulators of epigenetic gene expression. Nucleosomes are a barrier to transcription and the precise position of nucleosome, particularly around the transcription start sites, influences the initiation of transcription, thereby regulating gene expression (Schones, Cui et al. 2008, Cairns 2009). Histone variants differ in sequence and expression timing from the canonical counterparts and have been shown to have distinct functions, ranging from DNA repair and centromere determination to the regulation of gene expression (Wiedemann, Mildner et al. 2010). In summary, the cooperation and interplay between these different epigenetic layers compose a strict and dynamic epigenetic regulation system (Fig. 1), carrying out indispensible function in multiple biological processes.

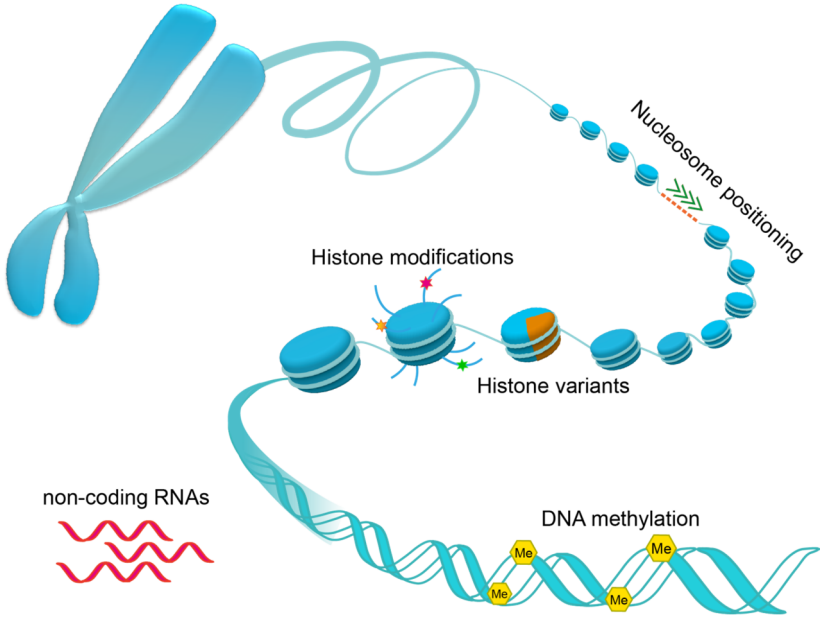

Figure 1. Schematic representation of fundamental mechanisms of epigenetic gene regulation. DNA methylation, histone modifications, nucleosome positioning and histone variants as well as non-coding RNAs are different epigenetic layers composing a strict and dynamic regulatory system (modified from Horsager. 2014 and episona website).

1.2 Zinc finger proteins (ZFPs) and zinc finger motifs

As mentioned above, changes in DNA methylation and histone modifications can alter the chromatin binding properties of transcription factors. To sense the DNA and histone modifications, a common feature of transcription regulators is the presence of epigenetic code binding motifs that direct the binding of these regulators to their targeting sites. Zinc finger motifs are the most abundant DNA binding modules in eukaryotic cells and are defined as small protein structures characterized by the coordination of zinc ions contributing to its structural stability. The first zinc binding domain was identified in the structure studies of transcription factor IIIA (TFIIIA) which is required for 5S RNA transcription in *Xenopus* oocytes (Miller, McLachlan et al. 1985). TFIIIA contains nine repeated zinc finger domains, each of which is stabilized by a zinc ion coordinated by a pair of cysteines and a pair of histidines (Klug 2010). The discovery of zinc finger domains identified a novel protein fold for nucleic acids recognition (Klug 2010). Since then, more

and more ZFPs have been found, which are predicted to account for at least 3% of all genes of the human genome (Bateman, Birney et al. 2002, Klug 2010). Multiple superfamilies of ZFPs are also identified in different kinds of living organisms including yeast, *Drosophila*, *Xenopus*, mouse and human. Due to their DNA binding properties, most ZFPs are transcriptional regulators involved in a broad range of functions in various cellular processes including cell proliferation, differentiation and survival. In addition to their well-known role in mediating DNA binding, zinc fingers are also molecular scaffolds for the recognition of RNA and proteins.

As research continues, more zinc finger motifs have been characterized, each with a unique three dimensional structure. It is desirable to classify these motifs in a proper way, which would help to better understand and predict the function of ZFPs by assigning it to a particular group. Normally, proteins are classified according to the structure similarity, but the classification of small protein structure becomes exceedingly difficult with conventional methods or existing databases due to the short length of the protein chain. A lot of attempts have been made to classify these zinc finger motifs. Recently, a more systematic method has been used to classify zinc fingers. All available zinc finger motifs were classified into eight groups based on the protein backbone similarity around the zinc ligands (Krishna, Majumdar et al. 2003). Among these eight folded groups, C2H2 type, CXXC type, treble clef fingers such as PHD and RING type motifs are the most prevalent zinc finger motifs in epigenetic regulatory systems.

1.2.1 C2H2 type zinc fingers

The C2H2 type zinc finger is the best characterized class of zinc finger motifs and probably represents the largest family of regulatory proteins in mammals. It is present in many mammalian transcription factors and other DNA-binding proteins, and participates in a variety of cellular activities such as development, differentiation and tumor suppression.

The C2H2 finger family is often defined by the consensus sequence $C\text{-}X_{2\text{-}4}\text{-}C\text{-}X_{12}\text{-}H\text{-}X_{3}\text{-}H$, in which C and H represent the zinc ligands and the number of X shows the intervals between the zinc binding residues (Fig. 2A) (Narayan, Kriwacki et al. 1997). Two zinc

ligands come from a zinc knuckle at the end of the β-hairpin and the other two ligands are contributed by the C-terminal end of the α-helix (Iuchi 2001, Krishna, Majumdar et al. 2003). Zinc knuckle is a unique turn with the consensus sequence CPXCG (Wang, Jones et al. 1998, Grishin 2001). The primary role of C2H2 type fingers is to bind DNA recruiting chromatin effectors for regulation of gene expression (Iuchi 2001). Early growth response protein 1 (EGR1), also known as ZIF268 or NGFI-A, is one of the most well-studied triple C2H2 ZFPs. EGR1 is a transcription factor working together with its partners to activate or repress gene expression by binding to the DNA at enhancers or repressors (Gashler, Swaminathan et al. 1993). Its DNA binding domain contains three zinc fingers and each finger recognizes approximately three nucleotides by binding to the major groove of target DNA (Fig. 2B) (Pavletich and Pabo 1991, Elrod-Erickson, Rould et al. 1996). This recognition pattern also laid the foundation for designing engineered ZFPs both for research and therapeutic applications.

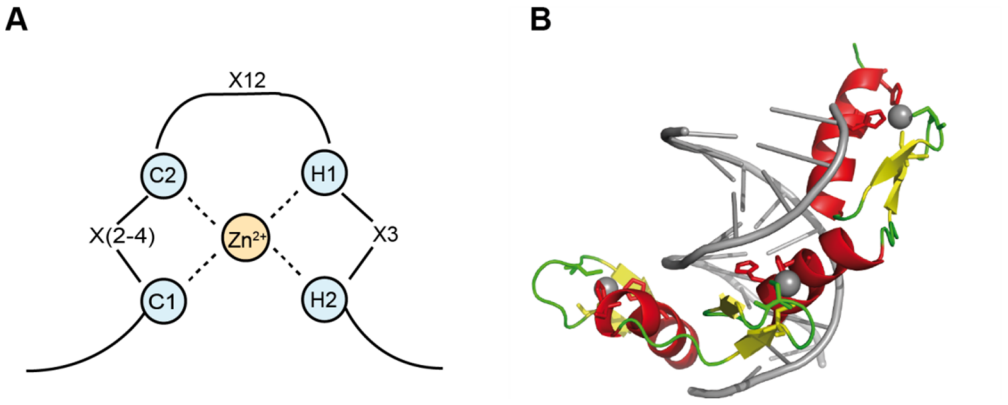

Figure 2. Structure of C2H2 type zinc fingers with DNA complex. (A) Schematic outline of the C2H2 type zinc finger domain. All C2H2 fingers show conserved structural feature, in which C represents cysteine and H represents histidine. The number of X shows the intervals between the zinc binding residues. Each Zn^{2+} ion (light brown) coordinates two histidines (light blue) and two cysteines (light blue). The color scheme is the same for Figure 2A, 3B, 4A and 5A. (B) Crystal structure of three zinc finger domains of EGR1 in complex with DNA. Three zinc fingers wrap around the DNA with α helices fitting into the major groove of DNA (PDB: 1ZAA). In all cases, the complex is shown in cartoon (helix in red, loop in green, sheet in yellow and double strand DNA in gray), zinc binding residues are shown in sticks and zinc ions are shown in gray spheres. All ribbon diagrams for zinc finger motifs in this study were visualized with Pymol software based on the corresponding PDB structure data.

Although the most famous role of C2H2 zinc finger is to bind DNA, there are several studies suggesting a role in recognition of RNA or in mediating protein-protein interactions (Brayer and Segal 2008). The zinc finger protein dsRBP-ZFa was the first C2H2 zinc finger protein described showing binding activity to double stranded RNA (dsRNA) and RNA-DNA hybrids in a sequence independent manner (Sun, Liu et al. 1996, Finerty and Bass 1997, Yang, May et al. 1999). Moreover, C2H2 zinc fingers also mediate protein-protein interactions via forming homo- or hetero-dimers (Iuchi 2001). For example, the last two C-terminal fingers of the DNA-binding protein IKAROS (IKZF1) bind each other and form a homo-dimer which in turn strengthens the DNA binding (Sun, Liu et al. 1996).

1.2.2 CXXC type zinc fingers

A CXXC type zinc finger was first identified in the transcriptional activator CpG-binding protein (CGBP) that exhibits specific binding preference for unmethylated CpG motifs, which was later renamed as CXXC1 (Voo, Carlone et al. 2000). Afterwards, CXXC domains were found in a variety of proteins with functions related to different chromatin remodelers such as methyl-CpG binding domain protein 1 (MBD1), mixed lineage leukemia protein 1 (MLL1), lysine-specific demethylase 2 (KDM2), DNA methyltransferase 1 (DNMT1) and ten-eleven translocation methylcytosine dioxygenase 1 (TET1) (Frauer, Rottach et al. 2011). However, the classification of CXXC domain is still uncertain. Some researchers regarded it as a fully new type zinc finger; while others classified CXXC to the category group “zinc binding loops” (Cross, Meehan et al. 1997, Fujita, Takebayashi et al. 1999, Krishna, Majumdar et al. 2003).

CXXC domain is relatively small, normally no more than 50 amino acids and is characterized by two cysteine-rich clusters and coordinates two zinc ions. The two cysteine-rich clusters are composed of C-X-X-C-X-X-C-$X_{4/5}$-C-G-X-C-X-X-C and C-X-X-R-X-C motifs (Fig. 3A and 3B) (Long, Blackledge et al. 2013). Each of the three zinc ligands are contributed from the cysteine cluster located in the N-terminal short helical segment, while each of the fourth one is from the C-terminal cysteine cluster that relatively far apart in the primary amino acids sequences (Allen, Grummitt et al. 2006, Song, Rechkoblit

et al. 2011). CXXC domains interrogate both the major and minor grooves of the DNA. A DNA binding loop segment from the CXXC domain penetrates into the major groove, while the region flanking the CXXC reaches around to the opposite DNA face and interacts with the minor groove (Fig. 3C) (Song, Rechkoblit et al. 2011, Long, Blackledge et al. 2013).

The general structure of CXXC domains is quite similar except the linker between the two cysteine-rich clusters; therefore, CXXC domains are further subdivided into three types based on the sequence similarity of the linker (Long, Blackledge et al. 2013). The sequence variation between these three subtypes is listed in Figure 3A. The DNA binding properties of the three types of CXXC domain exhibit divergence due to the different primary sequence of the linker regions. DNMT1, MLL1, MLL2, KDM2A, KDM2B and CFP1 are type I CXXC that mainly recognize unmethylated cytosine in a CpG context (Ohki, Shimotake et al. 2001, Allen, Grummitt et al. 2006, Cierpicki, Risner et al. 2010, Song, Rechkoblit et al. 2011, Xu, Bian et al. 2011). The type III CXXC in TET1 protein shows a more flexible DNA binding mode that permits recognition of CpG sequences regardless of modification states (Zhang, Zhang et al. 2010, Xu, Wu et al. 2011). Interestingly, not all CXXCs exhibit DNA-binding properties. MBD1 family contains more than ten isoforms and the main difference between the isoforms is the presence of two or three CXXCs. However, CXXC1 (MBD1-1) and CXXC2 (MBD1-2), both of which belong to type II CXXC, do not have DNA binding capacity, only CXXC3 (MBD1-3) shows DNA binding activity to unmethylated DNA and leads to repression of unmethylated reporter genes, which suggests an important role of the linker in the recognition of target sites (Nakao, Matsui et al. 2001, Jorgensen, Ben-Porath et al. 2004).

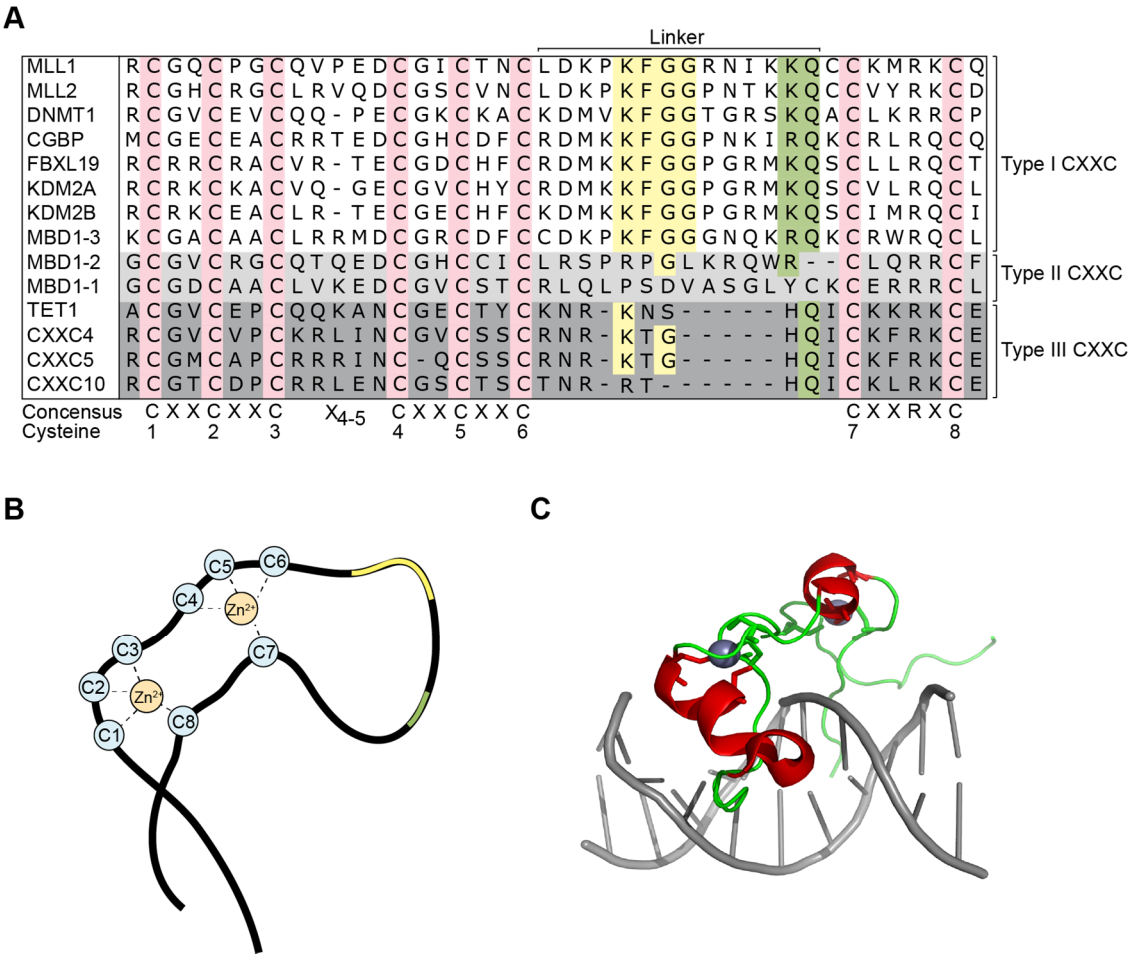

Figure 3. Sequence alignment and domain structure of CXXC type zinc fingers. (A) Alignment of CXXC domains of different proteins. CXXC domains can be divided into three subtypes depending on the similarity of their linker regions including KFGG (yellow) motif and KQ or RQ motifs (green). Eight cysteines are fully conserved across all of the CXXC domains and highlighted in pink. Accession numbers (for GenBank unless otherwise stated): MLL1, NP_001074518; MLL2, O08550 (SwissProt); DNMT1, NP_034196; CGBP, NP_083144; FBXL19, NP_766336; KDM2A, NP_001001984; KDM2B, NP_001003953; MBD1, NP_038622; CXXC4, NP_001004367; CXXC5, NP_598448; CXXC10, AGB05430 (modified from Frauer et al. 2011b and Long et al. 2013). (B) Schematic diagram of CXXC finger domain. Eight cysteines are distributed in two clusters (cysteine 1-6 and 7-8) which coordinate two zinc ions. KFGG and KQ or RQ linker regions are highlighted in yellow and green (modified from Long et al. 2013). (C) Crystal structure of DNMT1 CXXC domain in complex with DNA. DNA binding loop of CXXC domain interrogates the CpG site via the major groove, and the N- and C-terminal parts of CXXC domain interact with the minor groove of DNA (PDB: 3PTA).

1.2.3 Treble clef type zinc fingers

The name of treble clef motif comes from the projection of the domain C_α trace on one of the planes resembles the treble clef sigh (Grishin 2001). As one of the most versatile modules, the treble clef finger is present in a variety of protein families with diverse

functional properties (Grishin 2001). The treble clef motif is quite small in size with a core of 25 residues which consists of a zinc knuckle followed by a loop, a β-hairpin and an α-helix (Grishin 2001). Generally, the ligands for zinc binding site are cysteines provided by a CXXC pair of zinc knuckle and the first turn of an α-helix. However, histidine also works as zinc binding ligand, contributed by the C-terminal subsite (Grishin 2001, Krishna, Majumdar et al. 2003). The two halves of zinc binding sites exhibit distinct patterns of sequence conservation. N-terminal zinc binding subsite is characterized by the classical zinc knuckle CPXCG, and glycine normally precedes the second cysteine. As for the C-terminal subsite, the second cysteine is followed by a small residue, and usually a pair of uncharged residues is present before the first cysteine. There was no sequence similarity observed after the C-terminal subsite (Grishin 2001). Thus, in comparison with other zinc finger families, the lack of structure and sequence similarity between the two zinc half-binding sites has been regarded as one distinguishing feature of treble clef motifs. Besides, another feature of treble clef fingers is its ability to accommodate a variety of metal ion sites mainly placed between the β-hairpin and α-helix (Grishin 2001). Since the treble clef finger is very short, additional secondary structure elements are easily incorporated into treble clef domains, which further increased its variability. Thus, treble clef fingers are regarded as one of the most functionally diverse zinc binding motifs and has been divided into ten subgroups (Murzin, Brenner et al. 1995, Lo Conte, Ailey et al. 2000, Krishna, Majumdar et al. 2003). Really Interesting New Gene (RING) type fingers and Plant Homeo Domain (PHD) type fingers are two subtypes of treble cleft fingers, which are widely present in many epigenetic modifiers.

1.2.3.1 RING type fingers

Typically, RING type fingers harbor a C3HC4 amino acid motif and have been defined by the consensus sequence C-X_2-C-X_{9-39}-C-X_{1-3}-H-X_{2-3}-C-X_2-C-X_{4-48} -C-X_2-C which is shown in Figure 4A (Borden and Freemont 1996). In comparison with C2H2 fingers, RING type fingers are featured by the presence of the second zinc binding site and the third β-strand that forms hydrogen bonds with the principal β-hairpin (Grishin 2001). While the RING type finger is conserved in evolution, a small number of RING finger variants are classified

as RING-H2 (C3H2C3) family which has a histidine residue in the C4 position (Borden and Freemont 1996).

The well-known role of RING type fingers is in the ubiquitin-proteasome pathway tagging their targets for degradation by their E3 ubiquitin ligase activity. For example, the P53-induced protein with a RING-H2 domain (PIRH2) was found to regulate the turnover and function of a group of key factors involved in cell proliferation, differentiation and cell death (Jung, Qian et al. 2012, Halaby, Hakem et al. 2013). It harbors a highly conserved C3H2C3 type RING domain which ubiquitinates P53, CHK2 and P73, thus acts as a DNA damage response regulator (Leng, Lin et al. 2003, Corcoran, Montalbano et al. 2009, Jung, Qian et al. 2011, Bohgaki, Hakem et al. 2013). The structure diagram of RING finger domain of PIRH2 is shown in Figure 4B.

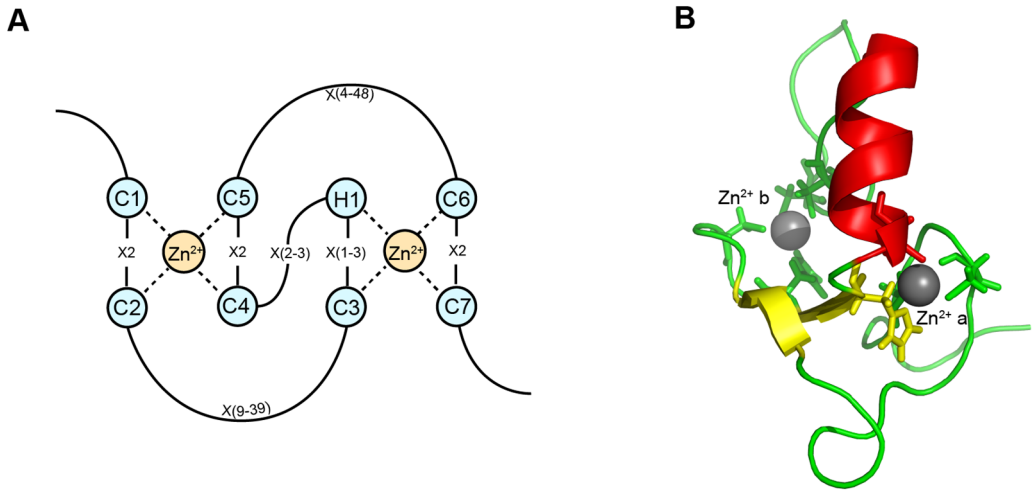

Figure 4. Structure of C3HC4 type RING fingers. (A) Schematic outline of C3HC4 type RING finger domains. RING fingers show a "cross-brace" motif. Each Zn^{2+} is ligated by either four cysteines or three cysteines and a histidine (modified from Borden et al. 1996). (B) Crystal structure of RING domain of PIRH2. The four binding ligands of Zn^{2+} a are provided by an N-terminal zinc knuckle and the loop between β sheet and α-helix, whereas the ligands for Zn^{2+} b are contributed by β-hairpin and a C-terminal zinc knuckle (PDB: 2ECM).

In addition to its important role in ubiquitination, RING fingers are also involved in macromolecular assembly (Borden and Freemont 1996, Saurin, Borden et al. 1996). The RING finger of the PML protein forms multiprotein complexes, also known as PML nuclear bodies, and mutations in its RING finger designed to abrogate zinc binding property lead to the disruption of nuclear bodies formation (Borden, Boddy et al. 1995). KRAB-

associated protein 1 (KAP1) is known as heterochromatin spreading regulator via its ability to influence epigenetic patterns and chromatin compaction. The RING finger within the RING-B box-coiled-coil (RBCC) domain of KAP1 is required for interaction with the KRAB domain of KOX-1, and mutations in RING domain abolished the association between these two domains (Peng, Begg et al. 2000, Peng, Begg et al. 2000). More and more new findings have pointed out novel functions of RING fingers in cell differentiation, cell cycle control and apoptosis (Borden 2000), suggesting possible additional functions in different cellular processes.

1.2.3.2 PHD type zinc fingers

The first PHD finger domain was discovered as a novel DNA binding motif in the homeodomain protein HAT3 in *Arabidopsis thaliana* (Schindler, Beckmann et al. 1993). It spans approximately 50-80 amino acids and consists of two strands of antiparallel β sheets and a C-terminal α-helix, which coordinates two zinc ions (Kwan, Gell et al. 2003, Li, Ilin et al. 2006). Its structure is characterized as C4HC3 motif and the consensus sequence is C-X_2-C-X_{8-21}-C-X_{2-4}-C-X_{4-5}-H-X_2-C-X_{12-46}-C-X_2-C (Fig. 5A), which is similar to that of the RING finger motif (Borden and Freemont 1996, Sanchez and Zhou 2011). Compared with RING type fingers, PHD fingers show greater conservation of sequence and spacing as well as additional conserved positions between individual metal ligands (Aasland, Gibson et al. 1995, Borden and Freemont 1996).

Although the PHD finger was described first with DNA binding property, proteins containing PHD finger play crucial roles in epigenetic regulation as chromatin modulator, mostly in the recognition of histone modifications. According to different ligand recognition properties, two major groups of PHD fingers are characterized. The first group of PHD fingers mediates the recognition of trimethylammonium group of lysine by a full or half “aromatic cage” which is formed by 2-4 aromatic and hydrophobic residues. The PHD finger of the bromodomain PHD finger transcription factor (BPTF) belongs to this class and its aromatic cage for binding trimethylated lysine 4 of histone H3 (H3K4me3) is composed of one tryptophan residue (W32) and three tyrosine residues (Y10, Y17 and Y23) (Fig. 5B and 5C) (Li, Ilin et al. 2006).

Another group of PHD fingers reads unmodified H3K4 (H3K4me0) by a combination of acidic and hydrophobic residues (Sanchez and Zhou 2011). These fingers are contained in many epigenetic modifiers such as autoimmune regulator (AIRE), tripartite motif-containing protein 24 (TRIM24), and DNA methyltransferase 3-like protein (DNMT3L) (Ooi, Qiu et al. 2007, Org, Chignola et al. 2008, Chakravarty, Zeng et al. 2009, Chignola, Gaetani et al. 2009, Tsai, Wang et al. 2010). Recently, several more PHD finger subsets were identified as readers of different modifications of histone tails. The PHD finger of ubiquitin-like PHD and RING finger domain-containing protein 1 (UHRF1) was found to recognize histone H3 unmodified arginine 2 (H3R2), which was required by UHRF1 to repress its direct target gene expression (Rajakumara, Wang et al. 2011), while the second PHD finger of chromodomain-helicase-DNA-binding protein 4 (CHD4) was reported to read methylated H3K9 (Musselman, Mansfield et al. 2009, Mansfield, Musselman et al. 2011). These different binding properties make PHD fingers versatile readers that exert complex and sophisticated functions in the epigenetic regulatory system.

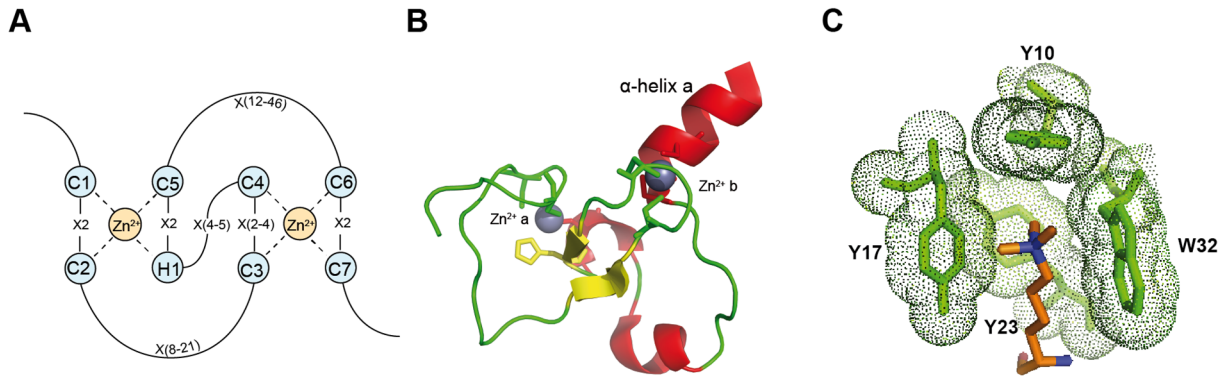

Figure 5. Structure of PHD fingers and the recognition sites with H3K4me3 containing peptide. (A) Schematic outline of PHD finger domains. PHD fingers show a C4HC3 structure motif of which each zinc atom is coordinated by either four cysteines or three cysteines and a histidine. (B) Crystal structure of the PHD finger domain of BPTF. An N-terminal loop and a short helical segment contribute two zinc ligands each for Zn^{2+} a, and the zinc knuckle and α-helix a each donates two zinc ligands for Zn^{2+} b (PDB: 2F6J). (C) The binding pocket in BPTF PHD finger recognizing H3K4me3. The binding pocket is composed of one tryptophan residue (W) and three tyrosine residues (Y) which are shown in green sticks and dots. Trimethylated lysine residue is indicated in sticks (PDB: 2F6J).

1.3 ZFPs act as modifiers in the epigenetic regulatory system

In human cells, the functional definition of genome goes far beyond its linear DNA sequence of 6 billion base pairs largely because of the well-controlled epigenetic regulation (Jakovcevski and Akbarian 2012). Deeply involved in the regulation of gene expression, many ZFPs are found to be epigenetic writers, readers and erasers (Fig. 6). For example, DNA methyltransferase 1 (DNMT1), the principal DNA methylation writer in mammalian cells, contains a CXXC type finger domain which is involved in maintaining the fidelity of its enzymatic activity (Song, Rechkoblit et al. 2011, Song, Teplova et al. 2012). After setting up of DNA methylation patterns, the readers MBDs recognize the marks to activate or repress gene expression. Among these epigenetic regulators, some carry distinct types of zinc finger domains and play multiple functions in epigenetic regulation networks. UHRF1 for instance, harboring chromatin binding domains and a RING type finger domain, acts as both DNA methylation reader and histone ubiquitination writer, connecting different epigenetic regulation layers (Hu, Li et al. 2011, Rajakumara, Wang et al. 2011, Wang, Shen et al. 2011, Xie, Jakoncic et al. 2012, Nishiyama, Yamaguchi et al. 2013). While the epigenetic modifications are highly dynamic, DNA methylation was believed to be the most stable epigenetic mark in the past decades. However, the CXXC type zinc finger protein TET1, recently emerged as a novel factor dynamically regulating DNA methylation. Thus, ZFPs fundamentally contribute to each layer and step of epigenetic regulation, and function as essential factors in the whole epigenetic information stream.

1.3.1 Writers of DNA modifications

DNA methylation is the most prevalent and extensively studied DNA modification in mammalian genomes. Three DNA methyltransferases, DNMT1, DNMT3A and DNMT3B, are responsible for the establishment and maintenance of DNA methylation patterns in mammals. DNA methylation patterns are set up during early embryonic development through a highly orchestrated process that involves genome-wide DNA methylation and demethylation. During early embryonic development, DNMT3A and DNMT3B are believed

to be DNA methylation writers and responsible for establishing *de novo* DNA methylation patterns (Okano, Bell et al. 1999, Kaneda, Okano et al. 2004, Zhang and Pradhan 2014). The third member of the DNMT3 family is DNMT3-like (DNMT3L) which is unable to catalyze the methyl group transfer due to the mutation of key catalytic residues, but still functions as a regulator of DNMT3A and DNMT3B (Chedin, Lieber et al. 2002, Suetake, Shinozaki et al. 2004). Another methyltransferase, DNMT2, has been reported as a tRNA methyltransferase rather than a DNA methyltransferase and has been renamed as tRNA aspartic acid methyltransferase 1 (TRDMT1) (Goll, Kirpekar et al. 2006, He, Chen et al. 2011).

DNMT1 is the principal DNA methyltransferase in mammals and is responsible for maintenance of methylation patterns after DNA replication (Reik, Dean et al. 2001, Li 2002, Jaenisch and Bird 2003). DNMT1 is a large protein comprising an N-terminal regulatory region (NTR) and C-terminal catalytic domain. Its NTR comprises two thirds of the molecule and contains several functional domains: a proliferating cell nuclear antigen (PCNA) binding domain (PBD), a heterochromatin targeting sequence (TS), a CXXC domain and two bromo-adjacent homology domains (BAH1/BAH2). NTR is connected to the C-terminal catalytic domain by seven glycine-lysine repeats $(KG)_7$ (Goll and Bestor 2005, Spada, Rothbauer et al. 2006, Rottach, Frauer et al. 2010, Qin, Leonhardt et al. 2011, Qin, Leonhardt et al. 2011). The PBD domain mediates the interaction of DNMT1 with PCNA at replication sites leading to the association of DNMT1 with the replication machinery (Leonhardt, Page et al. 1992, Chuang, Ian et al. 1997). The TS domain was reported to mediate association with heterochromatin and also dimerization of DNMT1 (Leonhardt, Page et al. 1992, Fellinger, Rothbauer et al. 2009). A recent crystal structure of DNMT1 showed that the TS domain is deeply inserted into the DNA-binding pocket of DNMT1 and masks the catalytic center completely together with the linker between the TS and CXXC domains, serving as an autoinhibitory mechanism of DNMT1 (Syeda, Fagan et al. 2011, Takeshita, Suetake et al. 2011). The CXXC domain was found to specifically bind to unmethylated CpG dinucleotides and keep unmethylated duplex CpG containing DNA away from the catalytic center by positioning the CXXC-BAH1 linker between DNA and active site of DNMT1, thus preventing its *de novo* methylation activity (Song, Rechkoblit

et al. 2011). The enzymatic activity of the C-terminal catalytic domain is under strict control of the N-terminal regulatory domain, which ensures that DNMT1 preferentially methylates hemimethylated DNA.

The activity of DNMT1 is not only regulated by its own N-terminal regulatory domain, but also regulated by a variety of interacting partners ranging from histone methyltransferases and deacetylases, and from DNA binding proteins to chromatin remodelers. PCNA was the first interacting partner that was shown to interact with DNMT1 and target DNMT1 to DNA replication sites in S phase during DNA replication (Leonhardt, Page et al. 1992, Chuang, Ian et al. 1997). Thus, direct interaction with PCNA ensures that the methylation patterns are faithfully preserved in newly synthesized DNA. However, compared with highly processive DNA replication, purified recombinant DNMT1 showed rather low catalytic activity (Jackson and Pombo 1998, Pradhan, Bacolla et al. 1999). Furthermore, a truncated DNMT1 protein lacking the PBD domain displayed reduced postreplicative DNA methylation maintenance activity *in vivo* (Schermelleh, Haemmer et al. 2007, Spada, Haemmer et al. 2007). These data indicate that the association with PCNA might be able to enhance the DNA methylation efficiency of DNMT1 *in vivo*. Besides PCNA, UHRF1 has recently emerged as an essential co-factor of DNMT1 for DNA methylation maintenance. Its genetic ablation in ESCs leads to genomic hypomethylation similar to *dnmt1*$^{-/-}$ ESCs (Bostick, Kim et al. 2007, Sharif, Muto et al. 2007). The mechanism how UHRF1 regulates the enzymatic activity of DNMT1 still remains to be elucidated, which is also a subject of this work and will be discussed in detail in later chapters. Furthermore, DNMT1 has also been reported to interact with HP1 protein, histone deacetylases HDAC1/2, histone lysine methyltransferases G9A and SUV39H1 (Fuks, Burgers et al. 2000, Robertson, Ait-Si-Ali et al. 2000, Rountree, Bachman et al. 2000, Tachibana, Sugimoto et al. 2002, Peters, Kubicek et al. 2003, Kim, Esteve et al. 2009, Qin, Leonhardt et al. 2011). In addition to interacting partners, the activity of DNMT1 is also regulated by posttranslational modifications such as phosphorylation, methylation, acetylation and ubiquitination. Phosphorylation of DNMT1 at serine 515 was shown to be important for the interaction between regulatory and catalytic domain (Esteve, Chang et al. 2011). In addition, phosphorylation of serine 143 by AKT1 kinase was

reported to play a role in the maintenance of DNMT1 stability (Esteve, Chang et al. 2011). All these regulatory mechanisms together form a strict complex network controlling the stability and activity of DNMT1.

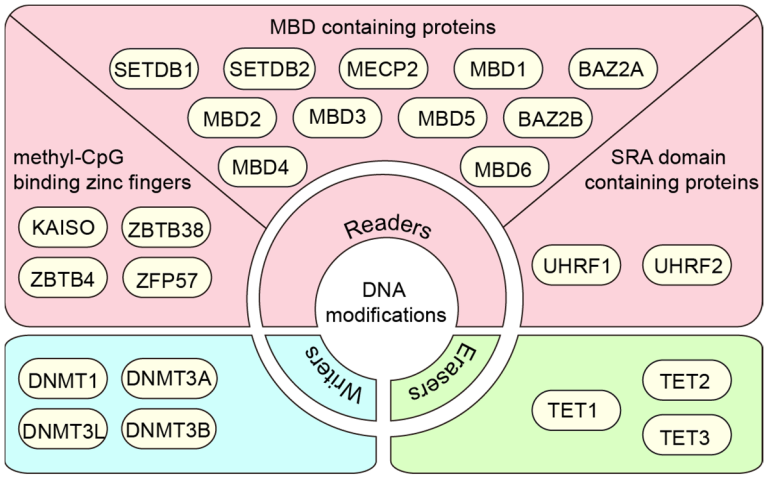

Figure 6. Overview of proteins as DNA modifiers. DNA modification writers and erasers are shown in blue and green respectively, and readers of DNA methylation are in pink and are divided into three groups: MBD containing proteins, methyl-CpG binding zinc fingers and SRA domain containing proteins. Notice, rather than directly erase DNA modifications, TET family proteins oxidize methyl group and may thus initiate DNA demethylation processes.

1.3.2 Readers of DNA modifications

DNA modification readers are also called effector proteins which recognize different modification states on DNA. Readers showing specific affinity for the most prevalent modification of DNA methylation are collectively named as methyl-CpG binding proteins (MBP) and play a pivotal role in the interpretation of DNA methylation (Filion, Zhenilo et al. 2006). There are three branches of the MBP family: MBD containing proteins, methyl-CpG binding zinc fingers and SRA domain containing proteins (Fig. 6) (Parry and Clarke 2011).

1.3.2.1 MBD containing proteins

The methyl-CpG binding domain (MBD) is present in a family of proteins conserved throughout the eukaryotic lineage (Hashimoto, Vertino et al. 2010). The MBD is about 85

amino acids in length, and in most cases, MBD confers the ability to bind methylated CpGs (Nan, Meehan et al. 1993). Currently there are 11 members of this group containing the MBD, but they do not all directly interact with methylated DNA (Parry and Clarke 2011). According to the presence of other domains, this family can be further divided into three groups: histone methyltransferases (HMT-MBD), histone acetyltransferases (HAT-MBD) and MECP2-MBD proteins.

SETDB1 and SETDB2 are two members of HMT-MBD group, both carrying MBD domain and SET domain. Compared to their functions as specific histone methyltransferase for H3K9, the DNA binding affinity of HT-MBD group is still unclear. The HAT-MBD group comprising of two proteins, BAZ2A and BAZ2B, differ at specific residues in the MBD domain from other members and cannot recognize methylated DNA (Parry and Clarke 2011). The main identified role of these two proteins is in the epigenetic silencing of ribosomal DNA within the nucleolus (Strohner, Nemeth et al. 2001, Santoro, Li et al. 2002, Zhou, Santoro et al. 2002, Strohner, Nemeth et al. 2004).

The MECP2-MBD group is the most extensively studied group of MBD containing proteins and has at least seven members including MECP2 and MBD1-6. MECP2 preferentially binds to methylated CpG sites by its MBD and recruits corepressor complexes, such as the SIN3A complex and the nuclear receptor corepressor (N-CoR) to silence gene expression (Lewis, Meehan et al. 1992, Nan, Meehan et al. 1993). As the largest member of this family, MBD1 has more than ten isoforms and confers different DNA binding preference according to the presence of the additional unmethylated DNA binding motif CXXC3. The biological significance of the dual DNA binding affinity of MBD1 is still unknown, but according to the research regarding the role of MBD1 in gene transcription repression, an intact MBD domain, rather than CXXC3 is indispensable for targeting MBD1 to pericentric heterochromatin and repressing gene expression, while CXXC3 is required for targeting MBD1 to unmethylated CpG-rich regions and acts as a repressor of unmethylated CpG island promoters (Jorgensen, Ben-Porath et al. 2004). Similar to MBD1, other members of this family also play roles in regulation of gene expression via association with distinct repression complexes, while they show different binding ability to DNA substrates (Feng and Zhang 2001, Hendrich, Guy et al. 2001, Feng, Cao et al. 2002, Saito and Ishikawa 2002,

Le Guezennec, Vermeulen et al. 2006, Parry and Clarke 2011). Notably, the C-terminal catalytic domain of MBD4 is homologous to bacterial DNA damage-specific glycosylases that mediate repair of hypermutable CpG by removing thymine or uracil from mismatched CpG sites (Hendrich, Hardeland et al. 1999). MECP2 additionally shows binding affinity to 5 hydroxymethylcytosine (5hmC) enriched within active genes and accessible chromatin in the brain (Mellen, Ayata et al. 2012). These new findings suggested that MBD containing proteins may also function as an important factor in the DNA demethylation pathway (Millar, Guy et al. 2002, Wong, Yang et al. 2002, Sjolund, Senejani et al. 2013).

1.3.2.2 Methyl-CpG binding zinc fingers

KAISO, ZFP57, ZBTB4 and ZBTB38 are the four members identified belonging to the group of methyl-CpG binding zinc fingers. KAISO is capable of binding a pair of methylated CpG dinucleotides in the consensus sequence 5'-C^{m}GC^{m}G-3' with its C-terminal C2H2 type zinc fingers (Prokhortchouk, Hendrich et al. 2001). These zinc finger motifs also target KAISO to unmethylated DNA with the consensus sequence 5'-CTGCNA-3', named as KAISO binding sequence (KBS), but it is still unclear whether KBS binding and methyl-CpG binding are related or separate activities of KAISO (Daniel, Spring et al. 2002). The specific DNA binding by KAISO can recruit N-CoR complex to methylated and unmethylated promoters to silence gene transcription (Yoon, Chan et al. 2003).

Similar to KAISO, transcription factor ZFP57 also shows binding affinity to its substrates in TGCC^{m}GC-dependent manner, which is mediated by its second and third C2H2 zinc fingers (Quenneville, Verde et al. 2011, Liu, Toh et al. 2012). This selective DNA binding property contributes to the maintenance of both maternal and paternal imprints (Loh, Zhang et al. 2007, Li, Ito et al. 2008). ZBTB4 and ZBTB38 have been reported to recognize DNA sequences containing one single methylated CpG site *in vitro* and *in vivo* (Filion, Zhenilo et al. 2006). The biological roles of these two proteins are still unclear.

1.3.2.3 SRA domain containing proteins

There are two members in this family, UHRF1 and ubiquitin-like PHD and RING finger domain-containing protein 2 (UHRF2); both proteins contain five distinctic functional domains (Fig. 7): a ubiquitin-like domain (Ubl), a tandem tudor domain (TTD), a PHD, a SET and RING associated (SRA) domain and a RING domain.

UHRF1 acts as a DNMT1-interacting partner which is essential for targeting DNMT1 to replication foci and maintains DNA methylation patterns during DNA replication (Bostick, Kim et al. 2007, Sharif, Muto et al. 2007). Generally, Ubl domains involve in ubiquitin-mediated intra-cellular proteolysis (Mori, Ikeda et al. 2012), however, it has not been structurally clarified how the Ubl domain in UHRF1 functions. The TTD domain has been shown to specifically bind H3K9me2/3 and is involved in associating UHRF1 to pericentric heterochromatin (Nady, Lemak et al. 2011, Xie, Jakoncic et al. 2012). The binding properties of the PHD to histone H3 tail remain unclear. Several studies demonstrated that PHD finger specifically recognizes unmodified histone H3 arginine 2 (H3R2) and the interaction is inhibited by methylation of H3R2, which links UHRF1 to the regulation of euchromatic gene expression (Rajakumara, Wang et al. 2011, Wang, Shen et al. 2011). When the PHD works together with the TTD, they were shown to bind H3K9me2/3 and overall binding activity was largely enhanced by the PHD (Xie, Jakoncic et al. 2012, Cheng, Yang et al. 2013). UHRF1 also harbors an SRA domain which recognizes hemi-methylated CpG site via a base flipping mechanism and targets DNMT1 to DNA substrates (Sharif, Muto et al. 2007, Arita, Ariyoshi et al. 2008, Avvakumov, Walker et al. 2008, Delagoutte, Lallous et al. 2008, Hashimoto, Horton et al. 2008, Qian, Li et al. 2008). This SRA domain mediated DNA binding is required for recruiting DNMT1 to its hemimethylated DNA substrates. In addition, UHRF1 contains a RING domain possessing E3 ubiquitin ligase activity and was shown to ubiquitinate DNMT1 and regulate its stability with other deubiquitinases (Du, Song et al. 2010, Felle, Joppien et al. 2011, Qin, Leonhardt et al. 2011). UHRF1 does also target histones for ubiquitination *in vitro* and *in vivo*, with a preference for histone H3. Recently, it was reported that UHRF1-dependent H3K23 ubiquitination acts as a platform for the recruitment of DNMT1 to DNA replication sites (Citterio, Papait et al. 2004, Nishiyama, Yamaguchi et al. 2013). Taken together, TTD, SRA,

PHD and RING domain work in concert and contribute differently to the targeting of DNMT1 to replication sites thus make UHRF1 act as an essential cofactor for DNA methylation maintenance.

As an epigenetic code reader, UHRF2 also shows the binding affinity to H3K9me2/3 and hemimethylated DNA similar to its paralog UHRF1. The interplay between TTD and PHD domains induces the preference for hemimethylated DNA and enhances its binding activity to methylated histone tail (Pichler, Wolf et al. 2011). However, the increased binding affinity for chromatin is unable to recruit DNMT1 to DNA replication foci for supporting maintenance DNA methylation (Pichler et al. 2011; Zhang et al. 2011), suggesting different functions for UHRF2 and UHRF1. In searching of potential functions in epigenetic regulation, UHRF2 was identified as a specific reader of 5hmC with a quantitative mass spectrometric assay in neuronal progenitors (Spruijt, Gnerlich et al. 2013). *In vitro* DNA binding assay and X-ray crystallography analysis of UHRF2-SRA in complex with 5hmC containing DNA further demonstrated that SRA specifically recognizes 5hmC, which shed new light on the biological functions of UHRF2 in active DNA demethylation pathway (Zhou, Xiong et al. 2014).

Similar to UHRF1, UHRF2 also possesses ubiquitin ligase activity and ubiquitinates cyclins D1 and E1 inducing G1 arrest when overexpression (Mori, Ikeda et al. 2011). The ubiquitination of nuclear polyglutamine aggregates by UHRF2 is able to rescue cells from polyglutamine induced cytotoxicity (Iwata, Nagashima et al. 2009). UHRF2 also acts as a small ubiquitin like modifier (SUMO) E3 ligase independent of its RING domain. UHRF2 effectively enhances zinc finger protein 131 (ZNF131) SUMOylation but does not enhance ZNF131 ubiquitination suggesting that UHRF2 has independent functional domains and different regulatory mechanisms for its ubiquitination and SUMOylation activity (Oh and Chung 2013).

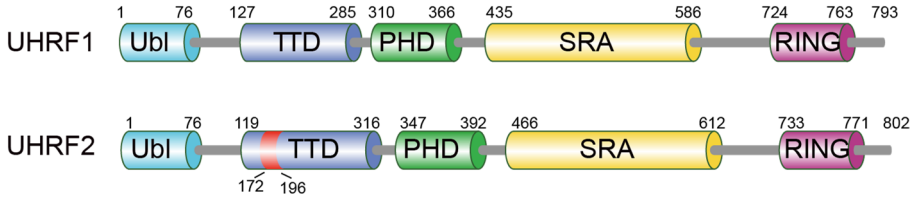

Figure 7. Domain organization of human UHRF1 and UHRF2. They both contain five functional domains: a Ubl domain, a TTD, a PHD, a SRA domain and a RING domain. In addition, UHRF2 harbors a stretch (red) in the TTD domain which is not present in UHRF1.

1.3.3 Erasers of DNA modifications

Two waves of genome-wide loss of 5mC during pre-implantation embryos and developing primordial germ cells (PGCs) suggest that DNA methylation is a stable but reversible epigenetic mark (Mayer, Niveleau et al. 2000, Oswald, Engemann et al. 2000, Hajkova, Erhardt et al. 2002, Saitou, Kagiwada et al. 2012). Thus, the identification of potential enzymes catalyzing DNA demethylation has been of interest for researchers over the past two decades. Though a direct DNA demethylase is not discovered yet, mounting evidence supported that the DNA repair machinery is deeply involved in restoration of unmodified cytosine, including thymine-DNA glycosylase (TDG), activation induced cytidine deaminase (AID) and apolipoprotein B mRNA editing enzyme catalytic polypepetide (APOBEC) mediated base excision repair (BER). Consistently, the oxidation products, 5hmC, 5 formylcytosine (5fC) and 5 carboxycytosine (5caC), converted from 5mC by TET family proteins were reported and regarded as intermediates in DNA demethylation pathways. Therefore, DNA demethylation is mainly achieved through a cyclic enzymatic cascade consisting of methylation of cytosine, iterative oxidation of 5mC to 5hmC, 5fC and 5caC by TETs and final replacement by unmodified cytosine via either replication dependent dilution or DNA glycosylase initiated BER (Fig. 8) (Wu and Zhang 2014).

TET family proteins were first discovered by systematic bioinformatic analysis of thymidine hydroxylase base J-binding protein 1 (JBP1) in *Trypanosoma brucei* (Yu, Genest et al. 2007, Cliffe, Kieft et al. 2009, Tahiliani, Koh et al. 2009). TET1 was initially found as a fusion partner of histone H3K4 methyltransferase MLL1 in certain acute myeloid and lymphocytic leukemia (Ono, Taki et al. 2002, Lorsbach, Moore et al. 2003). Alignment

analysis of protein databases identified the other two proteins with significant homology to TET1, which are named as TET2 and TET3 (Lorsbach, Moore et al. 2003). All three TET proteins possess enzymatic activity capable of hydrolyzing 5mC to 5hmC and further progressively oxidizing to 5fC and 5caC, and the enzymatic activity is carried out by the conserved C-terminal catalytic domain consisting of a cysteine-rich domain and a double-stranded β-helix (DSBH) domain which is a typical Fe^{2+} and 2-oxoglutarate (2OG)-dependent oxygenase domain. In addition to their catalytic domains, a type III CXXC is identified at the TET1 N-terminus but not in TET2 or TET3. It has been shown that TET1 CXXC domain binds to CpG sequences regardless of the modification states (Zhang, Zhang et al. 2010, Xu, Wu et al. 2011). Thus, the TET1 CXXC and their potential interacting partners may mediate the recognition of DNA substrates to initiate site-specific DNA demethylation.

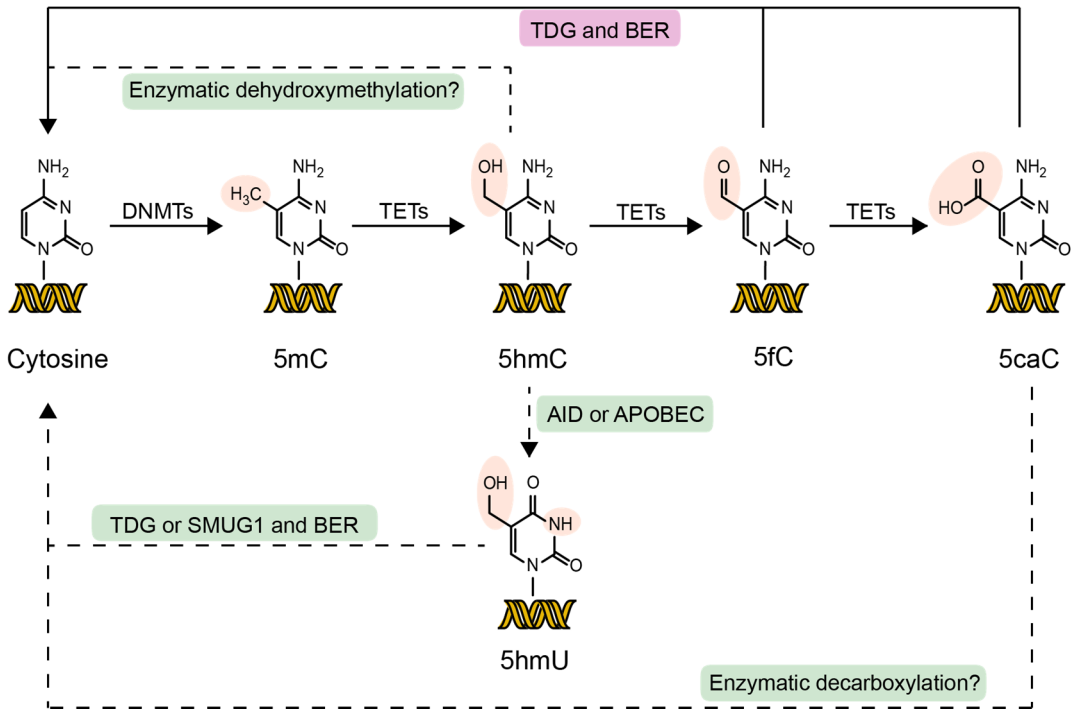

Figure 8. Overview of mechanisms of TET-mediated DNA demethylation. Cytosine is methylated by DNMTs to 5mC and sequentially oxidized by TETs to 5hmC, 5fC and 5caC. 5fC and 5caC can be excised by TDG and replaced by cytosine via BER. This pathway has been supported by multiple studies and is highlighted in pink and solid line. Other proposed mechanisms including dehydroxymethylation of 5hmC, decarboxylation of 5caC and deamination of 5hmC to 5hmU by AID or APOBEC, with subsequent removal by TDG or SMUG1 and finally replacement by cytosine via BER, are still hypothetical and therefore highlighted in green and dashed line (modified from Pastor et al. 2013).

Following the oxidation of 5mC by TET proteins, there are two ways to achieve the DNA demethylation. One is the DNA replication dependent manner through regulation of DNA methyltransferase association with its DNA substrates. Since UHRF1 is a crucial cofactor of DNMT1 and was identified as a reader of 5mC, its binding affinity to DNA containing 5hmC is of interest. While there are conflicting results about the binding preference of UHRF1 to hemi 5hmC (^{h}CG: CG) and hemi 5mC (^{m}CG: CG) substrates (Frauer, Hoffmann et al. 2011, Hashimoto, Liu et al. 2012, Mellen, Ayata et al. 2012), DNMT1 is significantly less efficient in methylating ^{h}CG: CG (by a factor of 60) than ^{m}CG: CG. Thus, the oxidized 5mC might block the binding sites of UHRF1 and affect maintenance methylation that leads to subsequent replication dependent passive dilution of 5mC (Pastor, Aravind et al. 2013). Notably, this mode is distinct from simple passive dilution of 5mC, as it may be effective even in the presence of a functional methylation maintenance machinery (Wu and Zhang 2014).

In addition to the passive DNA demethylation mechanism, several replication independent active demethylation mechanisms have been proposed. These mechanisms include dehydroxymethylation by DNMT enzymes (Liutkeviciute, Lukinavicius et al. 2009, Chen, Wang et al. 2012), enzymatic decarboxylation of 5caC (Schiesser, Hackner et al. 2012) and DNA glycosylase mediated excision of different oxidized 5mC bases (Guo, Su et al. 2011, He, Li et al. 2011, Maiti and Drohat 2011). Among these proposed active demethylation pathways, TDG-mediated excision of 5fC/5caC has received experimental support (He, Li et al. 2011, Maiti and Drohat 2011, Zhang, Lu et al. 2012). TDG is a member of the uracil DNA glycosylase (UDG) superfamily and is known to remove the pyrimidine base from T:G or U:G mismatches to initiate BER (Stivers and Jiang 2003). Especially after the discovery of additional oxidized cytosine derivatives, many studies have been carried out to revisit the role of TDG in DNA demethylation (Wu and Zhang 2014). In fact, TDG exhibits robust *in vitro* excision activity to 5fC:G and 5caC:G mismatch in duplex DNA (He, Li et al. 2011, Maiti and Drohat 2011). Structural analysis of TDG with 5caC complex further demonstrated that TDG has a high binding affinity to 5caC:G compared with its conventional substrate T:G mismatch (Zhang, Lu et al. 2012).

Consistently, TDG depletion in mouse ESCs caused 2-10 folds increase in levels of 5fC and 5caC which mainly accumulate at a large number of gene regulatory elements, supporting that TDG plays a role in active DNA demethylation pathways (Shen, Wu et al. 2013, Song, Szulwach et al. 2013). Cleavage of the glycosidic bound of 5fC and 5caC by TDG produces abasic sites that are further repaired by BER to restore unmethylated cytosine. Besides TDG, NEIL glycosylases have also been reported to contribute to DNA demethylation via excision of 5fC and 5caC and cause gene reactivation (Mueller, Bauer et al. 2014).

The second DNA repair based DNA demethylation mechanism involves AID and APOBEC. As the first step, 5hmC is deaminated to 5-hydroxyuracil (5hmU) by AID or APOBEC, then removed by single strand selective monofunctional uracil DNA glycosylase 1 (SMUG1) or TDG, finally replaced by cytosine via BER. However, some studies argue that AID primarily acts on single stranded DNA and recombinant AID and APOBEC has no detectable activity for 5hmC (Bransteitter, Pham et al. 2003, Nabel, Jia et al. 2012, Rangam, Schmitz et al. 2012). Therefore, further studies are needed to verify the feasibility of this mechanism.

1.3.4 Histone modifiers

In comparison to DNA methylation, histone modifications are more diverse and have more profound effects on chromatin structure. Distinct modifiers binding to targeting sites on histone tails establish or erase distinct posttranslational modifications such as acetylation, methylation, ubiquitination and phosphorylation. Notably, ZFPs play an important role in the dynamic regulation of histone modification states.

Many ZFPs act as histone modification writers to establish distinct histone marks which can be recognized by other chromatin modifiers to further set proper epigenetic codes. Many RING type zinc finger containing ZFPs have been reported as lysine ubiquitination writers working in such way. For example, polycomb group protein (PcG) RING2 is a member of human PRC1-like complex, which is necessary for maintaining the transcriptionally repressive state of its target genes throughout development. RING2 contains an N-terminal RING domain that mediates the monoubiquitination of histone H2A at lysine 119 (Wang, Wang et al. 2004). H2Aub creates a binding site for PRC2

complex and promotes H3K27 trimethylation on H2Aub nucleosomes, which forms a positive feedback loop establishing H3K27me3 chromatin regions (Kalb, Latwiel et al. 2014).

Besides, some ZFPs contain several types of zinc finger modules, and this multivalent recognition has emerged as a prevalent way for recognizing distinct chromatin states. As mentioned before, PHD type fingers are commonly found to read histone modifications especially methylation and acetylation at lysine residues (Li, Ilin et al. 2006, Shi, Hong et al. 2006, Taverna, Ilin et al. 2006, Wysocka, Swigut et al. 2006, Matthews, Kuo et al. 2007, Nakamura, Umehara et al. 2007, Lange, Kaynak et al. 2008, Zeng, Zhang et al. 2010). While the primary role of CXXC type zinc fingers is to recongnize specific DNA modification. These two types of zinc fingers could work together to read the composite marks at specific genomic loci. For example, KDM2A responsible for histone H3K36 demethylation contains both PHD and CXXC fingers that are required for chromatin association. In comparison to PHD fingers, the CXXC domain shows strong evidence for targeting KDM2A to unmethylated CpG islands (Blackledge, Zhou et al. 2010), and this chromatin recruitment of KDM2A triggers site-specific H3K36 demethylation to regulate gene expression.

In summary, the inter- and intra-molecular interactions mediated by zinc finger motifs set up dynamic interplays between different histone posttranslational modifications, and also link the two major epigenetic pathways - DNA methylation and histone modifications, building a well-controlled epigenetic regulation network.

1.4 Applications of ZFPs

The DNA recognition mode of zinc finger motifs is quite unlike other DNA binding proteins, in which one amino acid may mediate the recognition of more than one nucleotide base. As for zinc finger recognition, it principally is a one-to-one recognition between individual amino acid from a zinc finger motif to an individual DNA base, especially for C2H2 type zinc fingers three amino acids at the α-helix bind three successive bases on one stand of DNA (Miller, McLachlan et al. 1985, Pavletich and Pabo 1991, Wolfe, Nekludova et al. 2000, Klug 2010). Furthermore, fingers recognize their target as independent module, different fingers could be fused together to specifically recognize longer DNA sequences (Klug 2010). Therefore, zinc finger motifs are ideal natural tools for *de novo* design of proteins recognizing any given DNA sequences (Klug 2010).

1.4.1 Genome modification with engineered zinc finger nucleases

Genome editing is commonly used to insert, replace, mutate or remove a gene from a genome in order to understand the function of corresponding gene product in development and disease. This technology facilitates the genetic modifications by the creation of a double-strand break (DSB) in the DNA sequence of interest and followed by the generation of desired modifications during subsequent DNA break repair (Urnov, Rebar et al. 2010). DSB induced in genome editing is frequently generated by using endonuclease. An appropriate endonuclease possesses two basic qualities: specific recognition of target DNA and cleavage activity at the targeted loci. To achieve this goal, zinc finger nucleases (ZFNs) are designed and generated by fusing assembled zinc finger DNA binding motifs to a nuclease domain (Kim, Cha et al. 1996, Bibikova 2001). The assembled DNA binding domain contains a varying number of C2H2 type zinc fingers, each recognizing three bases of DNA (Miller, McLachlan et al. 1985, Wolfe, Nekludova et al. 2000); and the nuclease domain is the catalytic domain of the FokI restriction enzyme indicated in Figure 9. FokI is a bipartite restriction endonuclease and functions as a dimer meaning it requires two adjacent and independent binding events, thus ensuring the high fidelity of the cleavage specificity (Vanamee, Santagata et al. 2001). Additionally,

development of FokI variants which only cleave as a heterodimer further improved the specificity of targeted editing (Miller 2007, Szczepek 2007). With these improvements, ZFNs-mediated gene editing has been applied in various organisms and used to introduce different modifications into endogenous loci: gene disruption, gene correction and gene addition (Fig. 9). While ZFNs have been widely used in genome modification, a major limitation of this approach is that they prefer GNN triplets in comparison of AT-rich target sequences (Bibikova, Golic et al. 2002, Lloyd, Plaisier et al. 2005, Morton, Davis et al. 2006). Interestingly, a second designed nuclease based on transcription activator-like effectors (TALEs) was reported in the past several years (Boch, Scholze et al. 2009, Moscou and Bogdanove 2009). The combination of ZFNs with TALE-based nucleases (TALENs) can greatly increase specificity and improve efficiency in genome editing (Beumer, Trautman et al. 2013, Yan, Smith et al. 2013).

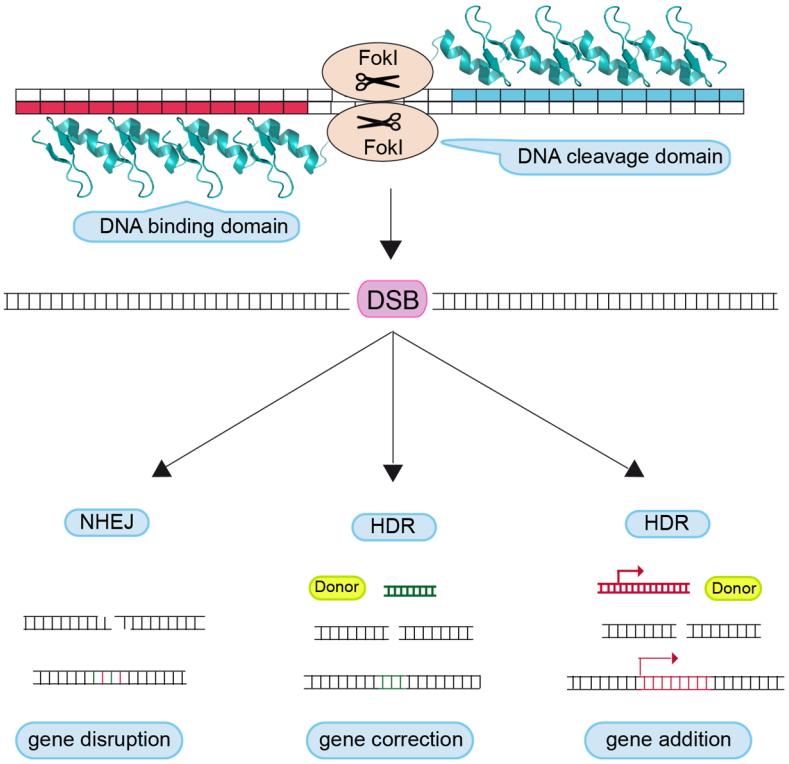

Figure 9. Structure and design of ZFNs binding to target DNA and potential outcomes of DSBs generated by a pair of ZFNs. Each ZFN contains a nuclease domain of FokI and an array of C2H2 type zinc finger motifs engineered to specifically recognize a sequence of interest. ZFNs treatment frequently results in DSB, and cells will repair DSBs via HDR or NHEJ pathways dependent on the presence of homologous donor DNA.

These pathways can be used for defined genome editing including gene disruption, gene correction and gene addition (modified from Urnov et al. 2010 and Sangamo BioSciences).

1.4.2 Manipulating gene expression with engineered zinc finger transcription factors

Engineered zinc finger transcription factors (ZFP-TFs) are designed transcription factors that imitate natural gene regulators. By attaching gene activation or repression domain to the engineered ZFP, one can generate ZFP-TFs that are capable of manipulating gene expression. This novel technology has found increasing applications in research and medicine.

In 1994, a three-finger peptide was created to recognize a unique nine base pair DNA region of transforming oncogene p190 *bcr-abl* cDNA and caused the repression of oncogenic expression. This application led to future studies of potential applications in gene regulation for research purpose and therapeutic correction (Choo, Sanchez-Garcia et al. 1994). From then on, researchers have been working on regulating gene expression with different combinations of regulatory domains and ZFPs. Kruppel-associated box (KRAB) is one of the most powerful transcriptional repression domain and is commonly used to inhibit gene expression. Transcriptional repression mediated by KRAB domain requires interaction with many heterochromatin associating proteins such as KAP1, HP1 and chromatin remodeling complex NuRD (Fig. 10A) (Kim, Chen et al. 1996, Lechner, Begg et al. 2000, Peng, Begg et al. 2000, Schultz, Friedman et al. 2001, Schultz, Ayyanathan et al. 2002, Sripathy, Stevens et al. 2006). The Choo group successfully applied KRAB fused ZFP to inhibit human immunodeficiency virus (HIV) expression (Reynolds, Ullman et al. 2003). A subsequent study applied a similar method to investigate the effect of designed ZFP-TF in regulating the replicative cycle of herpes simplex virus (HSV) (Papworth, Moore et al. 2003).

Similar to repression of gene expression by fusing assembled ZFPs to KRAB domain, gene transcription activation domains can be used to activate gene expression. VP16 is the most popular transcription activator which is widely fused to host transcription factors to amplify their activity (Hirai, Tani et al. 2010). The transcription activation domain (TAD) of VP16 interacts with basal transcription factors such as TFIIA, TFIIB, TFIID and TFIIH as well

as histone acetyltransferase P300 and PCAF, thereby leads to chromatin decondensation and activates gene transcription (Fig. 10B) (Lin, Ha et al. 1991, Xiao, Pearson et al. 1994, Zhu, Joliot et al. 1994, Kobayashi, Boyer et al. 1995, Tumbar, Sudlow et al. 1999, Kundu, Palhan et al. 2000). The Wolffe group fused designed ZFP targeting vascular endothelial growth factor A (VEGF-A) to VP16-TAD and observed the upregulated expression of VEGF-A both at mRNA and protein level (Liu, Rebar et al. 2001). Later experiments successfully applied ZFP-TF to activate the expression of VEGF-A in a mouse model which evoked a potentially therapeutic biophysiologic effect (Rebar, Huang et al. 2002). Recently, engineered ZFN-TFs have been designed to activate endogenous gene expression. The Kreader group fused p65, a strong activating subunit of nuclear factor kappa B complex, to over 300 designed ZFPs targeting different promoter regions and they identified several ZFP-p65s that are able to activate OCT4, SOX2, KLF4 and c-MYC which are the key factors for induced pluripotent stem cell reprogramming (Ji, Fischer et al. 2014). However, they also identified several inactive ZFP-p65s binding as efficient as the most active one, suggesting that efficient binding might be not sufficient for activation of gene expression. These results inspired researchers to reconsider the design principles of ZFP-TFs.

While assembled ZFPs are powerful tool in manipulating gene expression, combining it with certain inducible system, including organic molecules or blue light dependent systems (Beerli, Schopfer et al. 2000, Blackledge, Zhou et al. 2010), can make this tool more adjustable, reversible, and repeatable. Thus, improved engineered ZFPs are still widely used as a powerful tool in biological research.

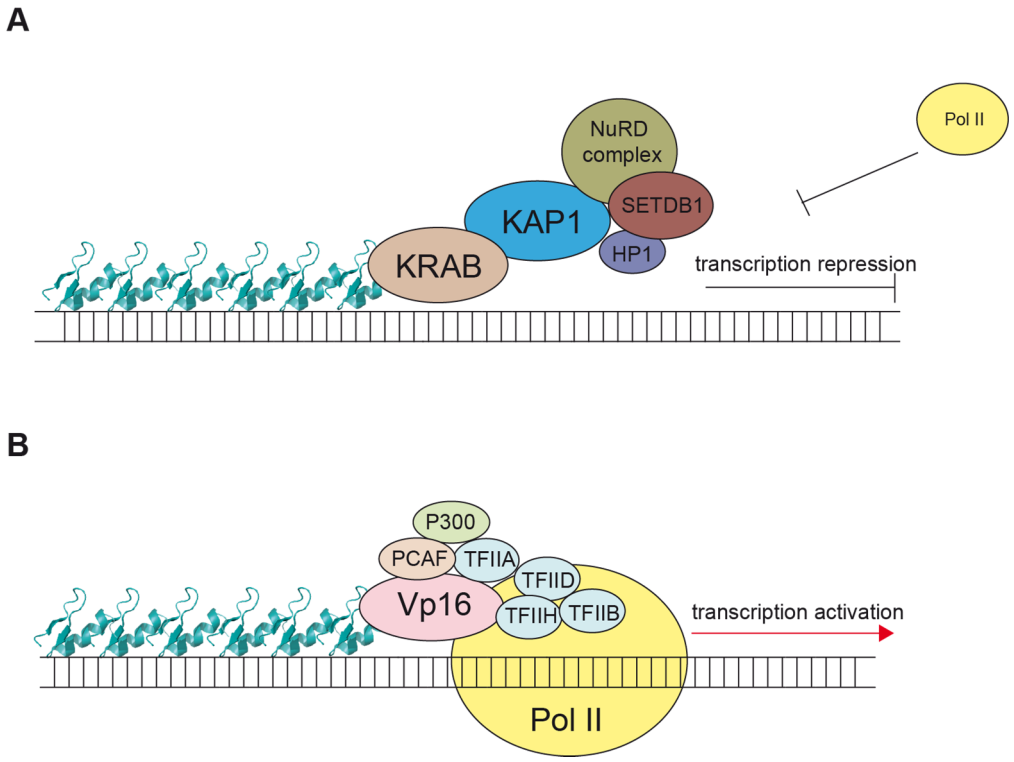

Figure 10. Schematic representation of ZFP-TFs mediated gene regulation. (A) Engineered ZFP fused to a KRAB domain leads to gene transcription repression. The KRAB domain interacts with and recruits heterochromatin co-regulators such as heterochromatin associates and chromatin remodeling complexes and results in gene silencing. (B) Engineered ZFPs fused to a Vp16-TAD domain leads to transcription activation. VP16 interacts with numerous proteins involved in gene activation including basal transcription factors and chromatin decondensation proteins.

1.4.3 Therapeutic applications of engineered ZFPs

Based on research findings with engineered ZFPs, the first therapeutic application of ZFN technology is on the treatment of HIV/AIDS. It was shown that a genetic 32-bp deletion of *Ccr5*, a major HIV-1 co-receptor, confers resistance to HIV-1 infection (Samson, Libert et al. 1996, Novembre, Galvani et al. 2005). The researchers engineered ZFNs to disrupt the *Ccr5* in human $CD34^+$ hematopoietic stem/progenitor cells (HSPCs) and produced both mono- and bi-allelically disrupted cells at a mean frequency of 17% of total alleles in a population. Mice transplanted with $Ccr5^{-/-}$ HSPCs had significantly lower HIV-1 level and preserved the transplanted human cells throughout their tissues. This study showed that the transient ZFN treatment can efficiently disrupt *Ccr5* of HSPCs and yield cells that

remain competent and support hematopoiesis (Perez, Wang et al. 2008, Holt, Wang et al. 2010). Successful application of assembled ZFPs was also shown in the treatment of glioblastoma multiforme (GBM), a progressive and usually fatal brain tumor. In addition, the Sangamo company is also working on ZFP therapeutics in neuroregenerative programs in spinal cord injury, traumatic brain injury and stroke.

With the powerful function in controlling gene expression, certain ZFP-TFs are under evaluation for the therapeutic treatment for some human diseases. In 2008, Sangamo company initiated phase II clinical trial of treatment for amyotrophic lateral sclerosis (ALS) characterized by muscle spasticity, rapidly progressive weakness due to muscle atrophy, and difficulty in speaking, swallowing and breathing. Manipulating the expression of vascular endothelial growth factor A (VEGF-A) with engineered ZFP-TF changes motor nerve function and muscle composition in both human cell line and animal model (Liu, Figley et al. 2010, Pawson, Duran-Jimenez et al. 2010). In addition, Sangamo company is also working on ZFP-TFs therapeutics in Parkinson's disease (PD) which is a chronic, progressive neurological disorder with increasing incidence in the aging population. Activation of the endogenous glial cell line-derived neurotrophic factor (GDNF) expression with assembled ZFP-TF promotes the survival of dopaminergic neurons. These assembled ZFPs provide a unique approach to modify genetic information and show their competitive advantages over chemical drugs. As research going on and technology progressing, ZFPs-based techniques will have wider applications both in scientific research and clinical fields.

1.5 Aims of this work

TET family enzymes have been extensively studied since the initial discovery of their enzymatic activity in 2009. Plenty of evidence suggests that TET proteins play roles in diverse biological processes, including epigenetic control of gene transcription, cell differentiation, embryonic development and tumorgenesis, but how TET proteins are targeted to specific genomic loci is still poorly defined. Interestingly, in human and mouse the *Tet2* and *Tet3* genes are adjacent to *Cxxc4* and *Cxxc10-1*, respectively. The CXXC domains encoded by these loci, together with those in TET1 and CXXC5, identify a distinct homology group within the CXXC domain family. Based on these observations, one of the main objectives of this thesis was to study the connections between TET proteins and zinc finger modules by characterizing their binding properties as well as by analyzing their regulatory roles in the function of TET proteins.

UHRF1 is an important epigenetic regulator which is required for maintenance of DNA methylation through recruiting DNMT1 to DNA replication forks. We were interested in the mechanism of UHRF1 targeting DNMT1 to DNA substrates. In particular, we aimed to elucidate the functional role of the PHD and RING domains of UHRF1 in controlling the maintenance of DNA methylation. Besides, UHRF1 is also suggested to be involved in dynamic changes of chromatin. We tried to analyze the functional link between UHRF1 and chromatin dynamics by identifying and analyzing the cellular targets of its ubiquitin E3 ligase activity. It is likely that the dynamic modification of chromatin-associated proteins by UHRF1 contributes to heterochromatin dynamics, thus providing a novel link to epigenetic regulation.

2 Materials and Methods

2 Materials and Methods

2.1 Materials

2.1.1 Technical devices

Devices	Type	Supplier
Agarose gel system	Mupid-Ex	Advance co
Bacterial incubator	UL 40	Memmert GmbH
Bacterial shaker	Certomat H+R	B.Braun
Cell culture microscope	EVOS xl	AMG
Centrifuge	Avanti J30l	Beckman Coulter GmbH
Clean bench	Herasafe KS, Class II	Fisher Scientific GmbH
CO2 incubator	Binder CB150	BINDER Inc.
Epiflourescence microscope	Axiophot 2	Carl Zeiss MicroImaging GmbH
FACS Aria II	Sorp	Becton Dickinson
Fixed angle rotor	fixed angle 1720	Hettich Zentrifugen
Fixed angle rotor	JA-14	Beckman Coulter GmbH
Freezer (-20°C)	Comfort	neoLab Migge Laborbedarf
Freezer (-80°C)	MDF-594	SANYO GmbH
Fridge	Premium	Liebherr
Gel documentation system	UV System	INTAS
High content imaging system	Operetta	PerkinElmer
Laser scanning confocal microscope	SP5	Leica microsystems
PCR maschine	Mastercycler	Eppendorf AG
Photometer	NanoVue	GE Healthcare

Pipettor	Eppendorf Research	Eppendorf AG
Pipettor	PIPETBOY	INTEGRA Biosciences GmbH
Power Supply unit	Bio-Rad PowerPac 300	Bio-Rad Laboratories GmbH
Real-time PCR System	7500 Fast	Applied Biosystems
Roller mixer	RM5	CAT
SDS PAGE system	Mini-Protean Tetra	Bio-Rad Laboratories GmbH
Shaker	DOS-10L	NeoLab Migge Laborbedarf
Sonifier	Branson Digital Sonifier 450D	G. Heinemann Ultraschall- und Labortechnick
Spinning disc confocal microscope	Ultraview VOX	PerkinElmer
Table centrifuge	Mikro 22R	Hettich Zentrifugen
Table top centrifuge	Centrifuge 5454	Eppendorf AG
Table top centrifuge	Rotina 38R	Hettich Zentrifugen
Microplate reader	Infinite M1000	TECAN
Thermo shaker	Thermomixer comfort	Eppendorf AG
Vortex mixer	Neolab 7-2020	NeoLab Migge laborbedarf GmbH
Water bath	Type 1013	GFL
Waving platform shaker	Polymax 1040	Heidolph Instruments GmbH&Co
Western blot scanning system	Typhoon Trio Variable Mode Imager	GE Healthcare

2.1.2 Consumables

Consumables	Manufacturer
CryoTube™	Nunc GmbH & Co. KG
Erlenmeyer flask (1 L)	SCHOTT
FACS tube (5 ml Polystrene Round-Bottom Tube with Cell-Strainer Cap)	Becton Dickinson
Falcon™ Tubes (15 ml, 50 ml)	Becton Dickinson
Fine forceps	Neolab
Immersol™ 518F (immersion oil)	Carl Zeiss
Laboratory bottle (1 L)	SCHOTT
Laboratory vacuum manifold Vac-Man®	Promega
Latex exam gloves „Satin PLUS“	Kimberly-Clark
Microscope coverslips (Ø 20 mm)	Menzel GmbH + Co KG
Nail polish (transparent)	Lacura Beauty
Nitrile laboratory gloves	SLG Süd-Laborbedarf
Parafilm®M sealing film	neoLab Migge Laborbedarf-Vertriebs
Pipette tips (10 µl / 200 µl / 1000 µl)	Brand Tech Scientific
pipettes (single channel)	Eppendorf
PureYield™ Binding and Clearing Columns	Promega
Reaction tubes (1,5 ml, 2 ml)	Eppendorf
Serological pipettes	Carl Roth GmbH + Co. KG
Soft wipes (KIMTECH Science)	Kimberly-Clark
Whatman® filter paper	Whatman
Cell culture plates & flasks Falcon	Becton Dickinson
Laboratory bottle (100 ml, 250 ml, 500 ml, 1 L)	SCHOTT
Pipette tips with filter (10 µl, 200 µl, 100 µl)	SLG Süd Laborbedarf

2.1.3 Reagents and kits

Reagents and kits	Supplier
4',6-Diamidino-2-phenylindol (DAPI)	Roche Diagnostics
2i = PD (PD 0325901) + CHIR (CT 99021 - CHIR 99021)	Axon Medchem
5-Aza-2'-deoxycytidine	Sigma-Aldrich
5-hydroxymethyl-dCTP	Bioline
5-methyl-dCTP	Jena Bioscience
Accutase	PAA Laboratories
Agar	AppliChem
Agarose	Sigma-Aldrich
Ammonium peroxodisulfate (APS)	Carl Roth GmbH + Co. KG
Ampicilline	AppliChem
Arginine/lysine isotopes	Silantes
B27 (50x)	Invitrogen GmbH
Betaine	Sigma-Aldrich
Blasticidin	Invitrogen GmbH
Bromphenol blue sodium salt	AppliChem
BSA	Sigma-Aldrich
Chloroform	Roth GmbH
Cycloheximide	Sigma-Aldrich
Dimethylsulfoxide (DMSO)	AppliChem
Disodiumhydrogenphosphate (Na_2HPO_4)	Carl Roth GmbH + Co. KG
Dithiothreitol (DTT)	Carl Roth GmbH + Co. KG
DMEM, high glucose with L-glutamine	PAA Laboratories GmbH
DMEM/F12	Invitrogen GmbH
Dnase I	AppliChem
dNTPs	PeqLab

Dodecylsulfate-sodium-salt	SERVA GmbH
Dulbecco's PBS (1x), without Ca^{2+} and Mg^{2+}	Sigma-Aldrich
ECL Plus reagent	Thermo Scientific
EDTA-dihydrate	AppliChem
EGF	Peprotech
Ethanol (98%)	AppliChem
Ethanol, absolute	AppliChem
Ethanol, technical grade	AppliChem
Ethidium bromide	AppliChem
EZ DNA Methylation-Gold Kit™	Zymo
FastDigest® restriction enzymes + Buffer	Fermentas
Fetal bovine serum (FBS)	PAA Laboratories
Fetal bovine serum "Gold"	PAA Laboratories
FGF2	Peprotech
Formaldehyde	Sigma-Aldrich
Gelatine	Sigma-Aldrich
Gentamicin (50 mg/ml)	PAA Laboratories
GFP-Trap®	ChromoTek
GlutaMax I (200 mM)	Invitrogen
Glycerol	Carl Roth GmbH + Co. KG
Glycin	Carl Roth GmbH + Co. KG
HEPES	PAA Laboratories
High-Capacity cDNA Reverse Transcription Kit	Applied Biosystems
HisTrap FF	GE Healthcare
HotStarTaq®Plus DNA Polymerase Kit	QIAGEN
Hydrochloric acid (HCl)	Carl Roth GmbH + Co. KG
Imidazol	AppliChem

Isopropanol (2-Propanol)	Carl Roth GmbH + Co. KG
Isopropyl β-d-thiogalactopyranoside (IPTG)	AppliChem
Kanamycin	SERVA
LB-Medium (Luria/Miller)	Carl Roth GmbH + Co. KG
L-Glutamine	PAA Laboratories
LIF (ESGRO)	Millipore
Lipofectamine 2000	Invitrogen
L-Mimosine	Sigma-Aldrich
Lysozyme	Sigma-Aldrich
Magnesium chloride ($MgCl_2$)	Sigma-Aldrich
MEM Non-essential Amnio Acid Solution	PAA Laboratories
MG-132	Santa Cruz Biotechnology
Milk powder	Carl Roth GmbH + Co. KG
N2 (100x)	Invitrogen
N-Ethylmaleimide	Sigma-Aldrich
Neurobasal medium	Invitrogen
Nonyl phenoxypolyethoxylethanol (NP-40)	Sigma-Aldrich
NucleoSpin Triprep Kit	Macherey-Nagel
NucleoSpin® Gel and PCR Clean up	Macherey-Nagel
OptiMEM	Invitrogen
PageRuler™ Prestained Protein Ladder	Fermentas
PBS (phosphate buffer saline)	PAA Laboratories
Penincillin/Strepomycin	PAA Laboratories
pET28a vector	Novagen
Phusion® High Fidelity Polymerase + Kit	New England Biolabs
PMSF (Phenylmethylsulfonylflouride)	SERVA
Polyacrylamid	Carl Roth GmbH + Co. KG

Polyethylenimine (PEI)	Sigma-Aldrich
Potassium chloride	Carl Roth GmbH + Co. KG
Potassium dihydrogen phosphate (KH_2PO_4)	Merck
Power SYBR Green PCR Master Mix	Applied Biosystems
Propidium Iodide	Sigma-Aldrich
Protease Inhibitor	SERVA
Protein assay kit	Pierce
Pure Yield™ Plasmid MidiPrep System	Promega
Puromycin	Invitrogen GmbH
QIAamp DNA Mini Kit	QIAGEN
Hot Start Polymerase	QIAGEN
Restriction enzymes, NEBuffer	New England Biolabs
RNAse A	AppliChem
RNase-free DNase I	Roche Diagnostics
RNeasy kit	QIAGEN
Rotiphorese Gel 30 (Acrylamide)	Carl Roth GmbH + Co. KG
Smart Ladder, Smart Ladder Small Fragment	Eurogentec
SOC Medium	New England Biolabs
Sodium chloride (NaCl)	Carl Roth GmbH + Co. KG
Sodium hydroxide (NaOH)	Carl Roth GmbH + Co. KG
Sodium sulfate (Na_2SO_4)	Sigma-Aldrich
β-Mercaptoethanol	Invitrogen
StrataClone™ PCR Cloning Kit	Agilent Technologies
StrataCone™ SoloPack® Competent Cells	Agilent Technologies
Streptomycin	PAA Laboratories
Superdex 75 preparative gel filtration column	GE Healthcare
T4 DNA Ligase	New England Biolabs

TaqMan probes	Applied Biosystems
TEMED	Merck
Tetramethylammonium-chloride (TMAC)	Sigma-Aldrich
Trichloroacetic acid (TCA)	Sigma-Aldrich
Tris	Carl Roth GmbH + Co. KG
Triton X-100	Carl Roth GmbH + Co. KG
Trypsin 2.5%	Invitrogen
Trypsin/EDTA 1x	PAA Laboratories
Tween 20	Carl Roth GmbH + Co. KG
Vectashield Mounting Medium	Alexis
Zero Blunt® PCR Cloning Kit	Invitrogen

2.1.4 Cell lines

Cell lines	Description	Type	Source
J1 wt	wild type	mESCs	(Lei, Oh et al. 1996)
dnmt1$^{-/-}$ J1	*dnmt1* c allele homozygous null, J1 background	mESCs	(Lei, Oh et al. 1996)
E14 wt	wild type	mESCs	(Sharif, Muto et al. 2007)
uhrf1$^{-/-}$ E14	homozygote *uhrf1* null, E14 background	mESCs	(Sharif, Muto et al. 2007)
ENC1	differentiated from mESCs	neural progenitor cell	
C2C12	mouse myoblasts	somatic cell	(Blau, Pavlath et al. 1985)
HEK293T	human embryonic kidney 293 cell line containing SV40 large T-antigen	somatic cell	(DuBridge, Tang et al. 1987)
BHK (*lac*O)	hamster kidney cell line, *lac*	somatic cell	(Tsukamoto, Hashiguchi

operator array integrated	et al. 2000)

2.1.5 Plasmid constructs

Constructs	Plasmid number	Promoter	Bacterial resistance	Description
GFP	592	CMV	kanamycin	Enhanced green fluorescent protein eukaryotic expression vector
RFP-DNMT1	1118	CAG	ampicillin	N-terminal RFP tagged full length DNMT1 eukaryotic expression vector
GFP-DNMT1	1626	CAG	ampicillin	N-terminal GFP tagged full length DNMT1 eukaryotic expression vector
GFP-DNMT1 Δ458-500	1312	CAG	ampicillin	Eukaryotic expression vector, N-terminal GFP tagged DNMT1 with TS domain deletion
Ch-USP7	2612	CAG	ampicillin	N-terminal Cherry tagged full length USP7 eukaryotic expression vector
Ch-USP7^{C224S}	2613	CAG	ampicillin	Eukaryotic expression vector, N-terminal Cherry tagged catalytically inactive full length USP7
HA-ubiquitin	1634	CMV	ampicillin	HA tagged ubiquitin eukaryotic expression vector
UHRF1-GFP	1976	CAG	ampicillin	C-terminal GFP tagged full length UHRF1 eukaryotic expression vector
Ch-UHRF1	1756	CAG	ampicillin	N-terminal Cherry tagged full length UHRF1 eukaryotic expression vector
LacI-GBP	1398	CMV	kanamycin	Bacterial Lac repressor (LacI) fused to GFP binding protein (GBP), used for F3H assay.
GFP-TET1	2271	CAG	ampicillin	N-terminal GFP tagged full length TET1 eukaryotic

				expression vector
$CXXC^{TET1}$ -GFP	2641	CAG	ampicillin	N-terminal GFP tagged CXXC domain of TET1, eukaryotic expression vector
GFP-Tudor	1936	CAG	ampicillin	N-terminal GFP tagged Tudor domain of UHRF1, eukaryotic expression vector
GFP-PHD	1937	CAG	ampicillin	N-terminal GFP tagged PHD domain of UHRF1, eukaryotic expression vector
GFP-SRA	1938	CAG	ampicillin	N-terminal GFP tagged SRA domain of UHRF1, eukaryotic expression vector
GFP-CBX1	2143	CAG	ampicillin	N-terminal GFP tagged full length CBX1, eukaryotic expression vector
GFP-CBX3	3029	CAG	ampicillin	N-terminal GFP tagged full length CBX3, eukaryotic expression vector
GFP-CBX5	3215	CAG	ampicillin	N-terminal GFP tagged full length CBX5, eukaryotic expression vector
GFP-Ubl	3061	CAG	ampicillin	N-terminal GFP tagged Ubl domain of UHRF1, eukaryotic expression vector
GFP-RING	3063	CAG	ampicillin	N-terminal GFP tagged RING domain of UHRF1, eukaryotic expression vector
GPF-CBX1 delc	3025	CAG	ampicillin	N-terminal GFP tagged CBX1 with deletion of C-terminal 6 amino acids, eukaryotic expression vector
pET-28a-$TET3^{CXXC}$	3224	T7	kanamycin	N-terminal 6xHis tagged antigen for generating antibody against $TET3^{CXXC}$, prokaryotic expression vector
GFP-CBX1 K180R	3216	CAG	ampicillin	N-terminal GFP tagged CBX1 K180R mutant, eukaryotic

				expression vector
GFP-CBX1 K181R	3217	CAG	ampicillin	N-terminal GFP tagged CBX1 K181R mutant, eukaryotic expression vector
GFP-CBX1 K184R	3218	CAG	ampicillin	N-terminal GFP tagged CBX1 K184R mutant, eukaryotic expression vector
GFP-H3 wt	3225	CAG	ampicillin	N-terminal GFP tagged full length histone H3.1 eukaryotic expression vector
GFP-H3 R2A	3219	CAG	ampicillin	N-terminal GFP tagged histone H3.1 R2A mutant eukaryotic expression vector
GFP-H3 K18A	3220	CAG	ampicillin	N-terminal GFP tagged histone H3.1 K18A mutant eukaryotic expression vector
GFP-H3 K23A	3221	CAG	ampicillin	N-terminal GFP tagged histone H3.1 K23A mutant eukaryotic expression vector
GFP-H3 K18A- K23A	3222	CAG	ampicillin	N-terminal GFP tagged histone H3.1 K18A-K23A double mutant eukaryotic expression vector
Ch-UHRF1 V429D-V431E	3223	CAG	ampicillin	N-terminal Cherry tagged UHRF1 V429D-V431E double mutant, eukaryotic expression vector
Cherry	1625	CAG	ampicillin	Fluorescent protein Cherry eukaryotic expression vector
GFP-CXXC4	2311	CAG	ampicillin	N-terminal GFP tagged full length CXXC4 eukaryotic expression vector
GFP-CXXC5	2713	CAG	ampicillin	N-terminal GFP tagged full length CXXC5 eukaryotic expression vector
CXXC4-GFP	2253	CAG	ampicillin	C-terminal GFP tagged full length CXXC4 eukaryotic expression vector

CXXC5-GFP	2246	CAG	ampicillin	C-terminal GFP tagged full length CXXC5 eukaryotic expression vector
Ch-CXXC4	2312	CAG	ampicillin	N-terminal Cherry tagged full length CXXC4 eukaryotic expression vector
GFP-TET3^{CXXC}L	2988	CAG	ampicillin	N-terminal GFP tagged full length TET3 isoform containing CXXC domain, eukaryotic expression vector
CXXC^{TET3}-GFP	2714	CAG	ampicillin	C-terminal GFP tagged CXXC domain of TET3, eukaryotic expression vector
GFP-TET3	2273	CAG	ampicillin	N-terminal GFP tagged full length TET3, eukaryotic expression vector

2.1.6 Primer sequences

2.1.6.1 Primer sequences for generating constructs (primers were produced by Eurofins Genomics).

Name	Sequence
CXXC ^{TET1} forward	5’-GGC GAT CGC ATG TCT ACG CCG CCA ATG-3’
CXXC ^{TET1} reverse	5’-CGC GGC CGC CTG GCT TCT TTT TGA GCA-3’
CXXC4 forward	5’-ATG CAC CAC CGG AAC GAC TCC CAG CG-3’
CXXC4 reverse	5’-TTA AAA GAA CCA TCG GAA CGC TTC AGC-3’
CXXC5 forward	5’-AAG CGA TCG CAT GTC GAG CCT CGG CGG TGG-3’
CXXC5 reverse	5’-GCG CGG CCG CTC ACT GAA ACC ACC GGA AGG-3’
CXXC ^{TET3} forward	5’-ATG CGA TCG CAT GCT GCG AGG GGG TGG AGA T-3’
CXXC ^{TET3} reverse	5’-ATG CGG CCG CCC GCT TTT TTC TTC AGC ACC TC-3’
TET3^{CXXC}L forward	5’-GGG CGA TCG CAT GAG CCA GTT TCA GGT GCC CTT GG-3’

TET3^{CXXC}L reverse	5’-GCG GCC GCC TAG ATC CAG CGG CTG TAG GGG CC-3’
CBX1 K180R forward	5’-GGG CGA TCG CAT GGG GAA AAA-3’
CBX1 K180R reverse	5’-CTG CGG CCG CCT AAT TCT TGT CGT CTT TAC-3’
CBX1 K181R forward	5’-GGG CGA TCG CAT GGG GAA AAA-3’
CBX1 K181R reverse	5’-TTG CGG CCG CCT AAT TCT TGT CGT CTC G-3’
CBX1 K184R forward	5’-GGG CGA TCG CAT GGG GAA AAA-3’
CBX1 K184R reverse	5’-CTG CGG CCG CCT AAT TCC GGT-3’
H3 WT forward	5’-ATG GGC GAT CGC ATG GCT CGT ACG AAG CAA-3’
H3 WT reverse	5’-AAC TGC GGC CGC TTA TGC CCT TTC CCC ACG GA-3’
H3 K9A forward	5’-CAA ACA GCT CGC GCG TCT ACC GGC GGC-3’
H3 K9A reverse	5’-GCC GCC GGT AGA CGC GCG AGC TGT TTG-3’
H3 K18A forward	5’-AAA GCT CCG CGC GCG CAG CTT GCT ACT-3’
H3 K18A reverse	5’-AGT AGC AAG CTG CGC GCG CGG AGC TTT-3’
H3 K23A forward	5’-CAG CTT GCT ACT GCA GCA GCC CGT AAG-3’
H3 K23A reverse	5’-CTT ACG GGC TGC TGC AGT AGC AAG CTG-3’
H3 R2A forward	5’-GAT CGC ATG GCT GCT ACG AAG CAA ACA-3’
H3 R2A reverse	5’-TGT TTG CTT CGT AGC AGC CAT GCG ATC-3’
UHRF1 V429D-V431E forward	5’-CCT GGT GAC CCT GAG GGC ACC ATG TGG CG-3’
UHRF1 V429D-V431E reverse	5’-GGT GCC CTC AGG GTC ACC AGG GAT GGG C-3’
CBX5 forward	5’-TGG GCG ATC GCA TGG GAA AGA AGA CCA AGA-3’
CBX5 reverse	5’-ACT GCG GCC GCT TAG CTC TTC GCG CTT TCT TTT-3’
CBX1 forward	5’-CCG CGA TCG CAT GGG GAA AAA GCA AAA CAA GAA-3’
CBX1 reverse	5’-CCG CGG CCG CCT AAT TCT TGT CGT CTT TTT TGT C-3’
CBX3 forward	5’-GGG CGA TCG CAT GGC CTC CAA TA-3’
CBX3 reverse	5’-CTG CGG CCG CTT ATT GTG CTT C-3’
CBX1 delc forward	5’-GGG CGA TCG CAT GGG GAA AAA-3’

CBX1 delc reverse	5'-CTG CGG CCG CTT AGT CAT CAT CC-3'

2.1.6.2 Primer sequences for relative quantification using qPCR (primers were produced by Metabion).

Name	Sequence
gapdh forward*	5'-CAT GGC CTT CCG TGT TCC TA-3'
gapdh reverse*	5'-CTT CAC CAC CTT CTT GAT GTC ATC-3'
tet1 forward*	5'-CCA GGA AGA GGC GAC TAC GTT-3'
tet1 reverse*	5'-TTA GTG TTG TGT GAA CCT GAT TTA TTG T-3'
tet2 forward*	5'-ACT TCT CTG CTC ATT CCC ACA GA-3'
tet2 reverse*	5'-TTA GCT CCG ACT TCT CGA TTG TC-3'
Total *tet3* forward*	5'-GAG CAC GCC AGA GAA GAT CAA-3'
Total *tet3* reverse*	5'-CAG GCT TTG CTG GGA CAA TC-3'
cxxc4 forward	5'-ACC TGG CAC TTC GCT AGA GAG A-3'
cxxc4 reverse	5'-TTG CCC TTC ATT CCC AAA TG-3'
cxxc5 forward	5'-CAG CAG TTG TAG GAA CCG AAA GA-3'
cxxc5 reverse	5'-TCC CGA CGG AAG CAT CAC-3'
cxxc10 forward	5'-GTG GAG ATG GGC GGA AGA A-3'
cxxc10 reverse	5'-GAT CTG GTG TGT GCG ACG AT-3'
tet3^{CXXC}*L* forward	5'-ATC GTC GCA CAC ACC AGA TC-3'
tet3^{CXXC}*L* reverse	5'-TCC TTC ACG AGC ATT TAT TTC CA-3'
tet3 forward	5'-GCG GCC GAT GCA GTA GTG-3'
tet3 reverse	5'-ATC AAC TGG GCT GAG CTC TGA-3'

* (Szwagierczak, Bultmann et al. 2010)

2.1.6.3 Primer sequences for 5’ RACE, conventional RT-PCR and generation of probes for northern blotting (primers were produced by Eurofins Genomics).

Name	Sequence
GSP1	5’ -AGG TCC ATC AAC TGG GCT-3’
$(dT)_{17}$-adaptor	5’-GAC TCG AGT CGA CAT CGA $(T)_{17}$-3’
adaptor primer	5’-GAC TCG AGT CGA CAT CG-3’
GSP2	5’-AGC ACC TCA CAC TTG CG-3’
GSP3	5’-GCA GCT GGT ACA AGA CC-3’.
Primer a	5’- GCG ATC GCA TGA GCC AGT TTC AGG -3’
Primer c	5’- AAG CGG CCG CCA GTC GGG CTT CTG GTC TAC -3’
Primer b	5’- ATG GCT GGG AGT GAG AC -3’
Primer d	5’- ATC GCA GGT GCA GTT GGG TG -3’
cxxc10 probe forward	5’-CAC ACC CAT TGG CTC ACC T-3’
cxxc10 probe reverse	5’-GGG TCT CAC TCC CAG CCA-3’
tet3 probe forward	5’-GCT CTC AAC TAC CTG CTT CC-3’
tet3 probe reverse	5’-CAT TGA GGC CAC ATC TCC G-3’

2.1.6.4 Sequences of oligonucleotides used for preparation of double stranded DNA substrates (oligonucleotides were produced by Metabion).

M: 5-methylcytosine X: 5-hydroxymethylcytosine

Name	Sequence
CGup	5’-CTCAACAACTAACTACCATCCGGACCAGAAGAGTCATCATGG-3’
um647N	5’-ATTO647N-CCATGATGACTCTTCTGGTCCGGATGGTAGTTAGTTGTTGAG-3’
MGup	5’-CTCAACAACTAACTACCATCMGGACCAGAAGAGTCATCATGG-3’
mC700	5'-ATTO700-CCATGATGACTCTTCTGGTCMGGATGGTAGTTAGTTGTTGAG-3'
hmCGup	5'-CTCAACAACTAACTACCATCXGGACCAGAAGAGTCATCATGG-3'

hmC550	5'-ATTO550-CCATGATGACTCTTCTGGTCXGGATGGTAGTTAGTTGTTGAG-3'
um550	5'-ATTO550-CCATGATGACTCTTCTGGTCCGGATGGTAGTTAGTTGTTGAG-3'
um700	5'-ATTO700-CCATGATGACTCTTCTGGTCCGGATGGTAGTTAGTTGTTGAG-3'
um590	5'-ATTO590-CCATGATGACTCTTCTGGTCCGGATGGTAGTTAGTTGTTGAG-3'
noCGup	5'-CTCAACAACTAACTACCATCTGGACCAGAAGAGTCATCATGG-3'
noCG647N	5'- ATTO647N-CCATGATGACTCTTCTGGTCTGGATGGTAGTTAGTTGTTGAG-3'

2.2 Methods

2.2.1 Expression constructs

Throughout this study enhanced GFP, monomeric red fluorescent protein (RFP) and monomeric Cherry fusion constructs were used and are referred to as GFP, RFP and Cherry (Ch) fusions, respectively, for brevity. The generation of expression constructs for GFP, RFP-DNMT1, GFP-DNMT1, GFP-DNMT1 Δ458-500, Ch-USP7, Ch-USP7^{C224S}, HA-ubiquitin, UHRF1-GFP, Ch-UHRF1, LacI-GBP, GFP-TET1, GFP-CXXC^{TET1}, GFP-Tudor, GFP-PHD, GFP-SRA, GFP-CBX1 have been described previously (Schermelleh, Spada et al. 2005, Fellinger, Rothbauer et al. 2009, Meilinger, Fellinger et al. 2009, Frauer, Rottach et al. 2011, Qin, Leonhardt et al. 2011, Song, Rechkoblit et al. 2011, Pichler, Jack et al. 2012, Yang, Liu et al. 2013) (Table 2.1.5). To generate GFP-CBX3 and GFP-CBX5 constructs, the CBX3 and CBX5 coding sequences were amplified using cDNA from mouse E14 ESCs and subcloned into pCAG-GFP-IB vectors. The GFP-UHRF1 single domain constructs for UbI and RING domain and GFP-CBX1 delc construct were generated by PCR using the corresponding wt full length construct as template. The coding sequence for the antigenic peptide for generating TET3^{CXXC} antibody was amplified by PCR using the corresponding wt full length construct as template and inserted into pET-28 a (+) vector. GFP-CBX1 K180R, K181R, K184R, GFP-H3 R2A, K18R, K23R, K18R-K23R, and Ch-UHRF1 V429D-V431E were obtained by overlap extension PCR on the corresponding wt construct. The coding sequence of the Uba domain of RAD23A (amino acids 158 to 212) was amplified using cDNA from mouse E14 ESCs. To generate the GFP-2Uba and Ch-2Uba constructs, a duplicate Uba coding sequence was subcloned into both the pCAG-GFP-IB and the pCAG-Cherry-IB vector. Coding sequences for CXXC4, CXXC5, TET3^{CXXC}L, CXXC^{TET3} and TET3 were amplified using cDNA from mouse NSCs and subcloned into pCAG-GFP-IB vectors to generate N-terminal GFP fusions. Sequences coding for CXXC^{TET1}, CXXC^{TET3}, CXXC4 and CXXC5 were inserted into pCAG-Tev-GFP (derived from pCAG-GFP-IB) to generate C-terminal GFP fusions. CXXC4 coding sequences was also inserted into pCAG-Cherry-IB to generate N-terminal Cherry fusions. All constructs were verified by DNA sequencing.

2.2.2 Antibodies

The generation of anti-multi-red antibody has been previously described (Rottach, Kremmer et al. 2008). The anti-HA antibody 12CA5 was raised in mouse and used to detect ubiquitinated proteins. The anti-TET3^{CXXC} polyclonal antibody was generated in collaboration with the BioGenes GmbH. Endogenous UHRF1 was visualized by a rabbit anti-UHRF1 antibody (Citterio, Papait et al. 2004). For detection of GFP fusion proteins by western blot, a mouse anti-GFP (Roche) or a rat anti-GFP (ChromoTek) antibody was used. RFP or Ch fusion proteins were detected by the rat anti-red antibody 5F8 (Rottach, Kremmer et al. 2008). HA-ubiquitin was detected by the mouse monoclonal anti-HA antibody 12CA5. Equal loading of cell lysates was followed by a mouse anti-β-actin antibody (Sigma-Aldrich). For detection of histone H3, the rabbit anti-H3 antibody (Abcam) was used. Depending on the expected intensity of the signals, secondary antibodies either conjugated to horseradish peroxidase (anti-rabbit (Biorad), anti-rat and anti-mouse (Dianova)) or conjugated to fluorescent dyes (anti-mouse and anti-rat Alexa Fluor 647N, anti-mouse Alexa Fluor 594 as well as anti-rat and anti-mouse Alexa Fluor 488 (Invitrogen)) were applied. For detection of HRP-conjugated antibodies, an ECL Plus reagent (Thermo Scientific) was used.

2.2.3 Cell culture and transfection

Human embryonic kidney (HEK) 293T cells and baby hamster kidney (BHK) cells were cultured in DMEM supplemented with 10% fetal calf serum and 50 µg/ml gentamycine. ESCs including wt J1, *dnmt1*$^{-/-}$, E14 and *uhrf1*$^{-/-}$ were cultured without feeder cells in gelatinized flasks as described (Schermelleh, Spada et al. 2005). Culture medium was supplemented with 1000 U/ml recombinant leukemia inhibitory factor LIF (Millipore). The NSC line ENC1 used throughout this study was derived from E14 ESCs as described (Conti, Pollard et al. 2005) and was maintained in Knockout-DMEM/F12 containing 2 mM GlutaMAX-I 100 U/ml penicillin, 100 µg/ml streptomycin, and supplemented with 1% N2 and 20 ng/ml each FGF-2 and EGF. The *dnmt1*$^{-/-}$ ESCs used in this study are homozygous for the c allele (Lei, Oh et al. 1996). Mouse E14 wt and *uhrf1*$^{-/-}$ cells (Sharif, Muto et al.

2007) have been reported before (Meilinger, Fellinger et al. 2009). Mouse ESCs were transfected with Lipofectamine 2000 reagent (Invitrogen) according to the manufacturer's instructions. HEK293T cells and BHK cells were transfected using polyethylenimine as transfection reagent according to the manufacturer's instructions.

2.2.4 Generation of stable cell lines and DNA methylation analyses

48 h after expression of GFP tagged constructs in *dnmt1*$^{-/-}$ ESCs and *uhrf1*$^{-/-}$ ESCs, GFP positive mESCs were separated using a fluorescence activated cell sorting (FACS) Aria II instrument (Becton Dickinson). Stably expressing cells were expanded in selection medium containing 10 µg/ml blasticidin (GFP-DNMT1 wt and GFP-DNMT1 Δ458-500) or 500 ng/ml puromycin (UHRF1-GFP wt, H346G and H730A) and GFP positive cells were FACS sorted a second time. Furthermore, the UHRF1-GFP wt, H346G and H730A cell lines were single cell sorted. Single clones of GFP-DNMT1 Δ458-500 and corresponding wt (Qin, Leonhardt et al. 2011) were picked manually. For all cell lines, clones with low expression levels were chosen for further analyses. The expression level and the accuracy of the expressed GFP fusion constructs were checked by western blot analyses.

2.2.5 Cell inhibitor treatment

For the *in vivo* ubiquitination assay, transfected HEK293T cells were incubated with medium supplemented with 2 mM N-ethylmaleimide (NEM) for 30 min before harvesting. For cycloheximide treatment, transfected cells were incubated in medium containing 10 µg/ml cycloheximide and harvested at the indicated time points.

2.2.6 Coimmunoprecipitation (Co-IP) and western blot

For co-immunoprecipitation assays, GFP fusion pulldowns using the GFP-Trap (ChromoTek) were performed as described (Rothbauer, Zolghadr et al. 2008). The GFP and RFP, or Ch fusion constructs were co-expressed in HEK293T cells and protein extracts were equalized to the same GFP concentration prior to co-immunoprecipitation with the

GFP-Trap (ChromoTek). Bound fractions were firstly detected by fluorescence intensity measurements and secondly by western blot analyses.

For detection of ubiquitinated proteins, cells were lysed in buffer containing 150 mM KCl, 50 mM Tris-HCl (pH 7.4), 5 mM $MgCl_2$, 1% Triton X-100, 5% Glycerol, 2 mM phenylmethylsulphonyl fluoride and 2 mM 2-mercaptoethanol and 5 mM NEM. After brief sonication, cell lysates were cleared by centrifugation at 4°C for 10 min. Supernatants were incubated with the GFP-Trap beads for 2 h at 4°C under gentle rotation. The beads were then washed three times with lysis buffer and resuspended in SDS-PAGE sample buffer. The anti-HA mouse monoclonal antibody 12CA5 was used for detection of ubiquitinated proteins.

2.2.7 Western blot quantification

Western blots developed with the ECL western blotting substrate were scanned and digital images were analyzed using Image J. The density was measured on the digital negative and background density was subtracted.

2.2.8 Immunofluorescence staining

Cells were grown on coverslips overnight, fixed with 3.7% formaldehyde in PBS for 10 min and washed with PBS for three times, followed by permeabilization in PBS containing 0.5% Triton X-100 for 5 min and block with PBS containing 3% BSA for 1 h. Cells were then incubated with primary antibody for 1 h and Alexa Fluor labeled secondary antibody for 1 h at RT. The antibodies were diluted in PBS containing 0.01% Tween-20 and 3% BSA. Cells were counterstained with DAPI and mounted in Vectashield (Vector Laboratories). Images were acquired with a TCS SP5 confocal microscope (Leica microsystems).

2.2.9 Protein production and purification

Protein production and purification were performed according to the handbook for high-level expression and purification of 6x His-tagged proteins (QIAGEN). Briefly, BL21(DE3) *E. coli* cells (New England Biolabs) carrying the expression construct were grown at 37°C

until OD_{600nm}~0.8 followed by expression induction with 1mM Isopropyl-β-D-thiogalactopyranosid (IPTG) at 37°C for 4 hours. After harvesting, bacterial cells were resuspended in 10 ml lysis buffer (300 mM NaCl, 50 mM Na_2HPO4, pH 8.0, 10 mM imidazole, 1 mM β-mercaptoethanol) and treated with with DNaseI (1 mg/ml), PMSF (2 mM) and lysozyme (10 mg/ml) for 1 h, followed by sonication. After clearing lysates by centrifugation and filtration, soluble extracts were loaded on a pre-equilibrated 1 ml HiTrap-column (GE Healthcare). Elution was performed with Ni-NTA elution buffer containing linear gradient ranging from 20 mM to 250 mM imidazol. Different elution fractions were further analyzed by SDS-PAGE assay. Fractions containing antigen peptide were pooled and further purified by gel filtration on a Superdex 75 column (GE Healthcare).

2.2.10 RNA isolation and Real-time PCR

Total RNA was isolated from cells and tissues using the nucleospin triprep kit from Macherey-Nagel. 500 ng of total RNA was reverse transcribed with a high-capacity cDNA reverse transcription kit (Applied Biosystems) according to the manufacturer's instructions. Real-time PCR was conducted using power SYBR green PCR master mix (Applied Biosystems) on a 7500 fast real-time PCR system (Applied Biosystems). PCR efficiency and primer pair specificity was examined using a standard curve of serially diluted cDNA and melting curve, respectively. After normalizing to the transcript level of glyceraldehyde phosphate dehydrogenase (*gapdh*), data were analyzed based on $2^{-\Delta\Delta CT}$ method (Livak and Schmittgen 2001).

2.2.11 Northern blot

Northern blotting was performed according to the DIG application manual for filter hybridization (Roche). Probes were generated and labeled by PCR using DIG-dUTP. Ten micrograms each of total RNA from ESCs and NSCs were separated on formaldehyde-agarose gels, transferred to Hybond-N^+ nylon membranes (GE healthcare) and immobilized by UV crosslinking. Blots were pre-hybridized with DIG Easy hyb (Roche) at

50°C for 30 min followed by overnight hybridization at 50°C. Probes were applied at a final concentration of 100 ng/ml in DIG Easy hyb. After washing, the blots were incubated with blocking solution (Roche) for 30 min, followed by incubation with alkaline phosphatase conjugated anti-digoxygenin antibody (Roche) for 30 min at room temperature. The membrane was washed twice, equilibrated with detection buffer (0.1 M Tris-HCl, 0.1 M NaCl, pH 9.5) and chemiluminescence with CDP-Star substrate (Roche) was used to detect the bound antibody.

2.2.12 5' RACE

5' RACE was performed according to Nature Methods (2005). 100 ng of total RNA from ENC1 NSCs were reverse transcribed as described above, but using the gene-specific primer 1 (GSP1). To remove excess primer, the reaction was purified with a silica mini-column (Nucleospin Gel and PCR Clean-up; Macherey-Nagel). After tailing with terminal deoxynucleotide transferase and dATP the tailed cDNA was subjected to nested PCR reactions with Phusion High-Fidelity DNA Polymerase (New England Biolabs). In the first reaction the upstream primers were $(dT)_{17}$-adaptor primer and adaptor primer, while the downstream primer was gene-specific primer 2 (GSP2). Cycling parameters were as follows: one cycle of 98°C for 30 s, 94°C for 5 min, 50°C for 5 min, and 72°C for 40 min, followed by 30 cycles of 94°C for 40 s, 54°C for 1 min, and 72°C for 3 min, with a final cycle of 94°C for 40 s, 54°C for 1 min, and 72°C for 15 min. In the second reaction the upstream primer was adaptor primer and the downstream primer was gene-specific primer 3 (GSP3). Cycling parameters were as follows: 98°C for 30 s, (98°C for 15 s, 55°C for 20 s, and 72°C for 30 s) 30 cycles, 72°C for 10 min. PCR products were purified by gel electrophoresis followed by silica column purification, cloned into pCR-Blunt with Zero Blunt PCR Cloning Kit (Invitrogen) and analyzed by DNA sequencing.

2.2.13 F3H assay

F3H assay was performed as described (Dambacher, Deng et al. 2012, Eskat, Deng et al. 2012, Herce, Deng et al. 2013). Briefly, BHK cells containing a *lac* operator repeats array

were seeded on coverslips in 6-well plates for microscopy. Cells were co-transfected with fluorescent fusion protein expression vectors and a LacI-GBP fusion construct using polyethylenimine and fixed about 16 h after transfection with 3.7% formaldehyde in PBS for 10 min, washed with PBST (PBS with 0.02% Tween), stained with 200 ng/ml DAPI and mounted in Vectashield medium. The F3H sample was analyzed with a TCS SP5 confocal fluorescence microscope. DAPI, EGFP and mCherry/RFP were excited by 405 nm diode laser, 488 nm argon laser and 561 nm diode-pumped solid-state laser, respectively. Images were recorded and further analyzed using imageJ.

2.2.14 Determination of global genomic 5hmC levels

Global 5hmC levels in genomic DNA from transiently transfected HEK293T cells were determined by the *in vitro* glucosylation assay as described previously (Szwagierczak, Bultmann et al. 2010, Frauer, Rottach et al. 2011) with minor modifications. Briefly, 50 ml reactions containing 150 mM NaCl, 20 mM Tris, pH 8.0, 25 mM $CaCl_2$, 1 mM DTT, 3.5 mM UDP-[^{3}H]glucose (20 Ci/mmol; Hartmann Analytic GmbH), 500 ng of sheared genomic DNA and 40 nM recombinant T4 β-glucosyltransferase were incubated for 20 min at room temperature and terminated by heating at 65°C for 10 min. DNA fragments were purified by silica column chromatography (Nucleospin, Macherey-Nagel) and radioactivity was determined by liquid scintillation. Radioactive counts were converted to percentages of 5hmC over total C using curves from PCR generated standards containing variable 5hmC/C ratios as previously described (Szwagierczak, Bultmann et al. 2010). The values for all GFP-TET constructs were corrected for differences in expression levels using GFP fluorescence measurements. This correction was not applied to control samples transfected with GFP as the latter is expressed at least at ten times higher levels than GFP-TET constructs, which would lead to artificially enhanced differences between basal 5hmC levels and those resulting by overexpression of TET constructs.

2.2.15 *In vitro* DNA binding assay

In vitro DNA binding assay was performed as described previously (Frauer and Leonhardt 2009, Frauer, Hoffmann et al. 2011, Frauer, Rottach et al. 2011). Briefly, three double stranded DNA oligonucleotides labeled with different ATTO fluorophores and with identical sequence, but either unmodified, symmetrically methylated or symmetrically hydroxymethylated cytosine at a single, central CpG site were used as substrates. GFP fusion constructs were expressed in HEK293T cells by transient transfection and immunopurified from cell lysates using the GFP-trap (ChromoTek). GFP-trap beads were washed three times before incubating with DNA substrates at a final concentration of 160 nM each. After removal of unbound substrates, amounts of protein (GFP fluorescence) and bound DNA were measured with Infinite M1000 plate reader (Tecan).

2.2.16 Statistical analysis

Results were expressed as mean values ± the standard deviation (SDEV) or as mean values ± the standard error of the mean (SEM) from the number of biological replicates indicated in the corresponding figure legend.

3 Results

3.1 Intrinsic and extrinsic connections of TET3 dioxygenase with CXXC modules

3.1.1 Identification and expression pattern of mouse *tet3* transcripts encoding a CXXC domain

In higher eukaryotes methylation of genomic cytosine to 5mC prominently contributes to epigenetic indexing of transcriptional activity. 5mC has long been regarded as a stable mark mediating permanent repression, but recently it has been shown that 5mC can be progressively oxidized to 5hmC, 5fC and 5caC by a three member family of TET proteins (Tahiliani, Koh et al. 2009, Ito, D'Alessio et al. 2010, He, Li et al. 2011, Ito, Shen et al. 2011), which has kindled the idea that they might represent intermediates in DNA demethylation pathways. Very few interactions involving TET proteins have so far been reported and even fewer known domains are identified in these proteins despite their relatively large size. As a consequence, little is known about how TET proteins are targeted to specific genomic loci in distinct cell types and developmental stages.

The N-terminal region of TET1 contains a CXXC domain (Tahiliani, Koh et al. 2009). In contrast, none of the human and mouse annotated genomic or transcript sequences for *tet2* and *tet3* include a sequence encoding such a domain. However, in both human and mouse genomes the *tet2* and *tet3* genes are adjacent to loci encoding *cxxc* domains, *cxxc4* and *cxxc10-1*, respectively (Fig. 11A) (Katoh and Katoh 2004, Frauer, Rottach et al. 2011). The *cxxc4* and *tet2* loci are 700 and 800 kb apart in the human and mouse genomes, respectively. These loci are transcribed in opposite orientations and encode distinct proteins, suggesting that they evolved through splitting of a *tet1*-like ancestral gene and intergenic inversion. The *cxxc10-1* ORF was identified *in silico* about 13 kb upstream of the annotated transcriptional start site of *tet3* and has the same orientation as the *tet3* ORF. Previously, we showed that the CXXC domains of TET1, CXXC10-1, CXXC4 and CXXC5 constitute a homology group distinct from CXXC domains present in several other factors with functions related to DNA or chromatin modification (Frauer, Rottach et al. 2011). The proximity and co-orientation of the *cxxc10-1* and *tet3* ORFs in human and

mouse genomes suggest that alternative *tet3* transcripts may include the *cxxc10-1* ORF. This is also suggested by GenBank entries of *tet3* orthologues encompassing an N-terminal *cxxc* domain from other vertebrate species, including a *Xenopus tet3* transcript and a TET3 protein homolog predicted from the genomic sequence of the naked mole rat (*Heterocephalus glaber*). Alignment of the CXXC domains from these TET3 homologues with the CXXC domains of mouse CXXC10-1, TET1, CXXC4 and CXXC5 shows that they all belong to the same homology subgroup that we identified previously (Fig. 11B). In addition, the *Hydra* genome encodes a single *tet* homolog and its predicted protein product contains an N-terminal CXXC domain with key features of this subgroup (Fig. 11B). These observations support the idea of a common ancestral *tet* gene encoding a CXXC domain and that in addition to *tet1*, this arrangement is preserved also in vertebrate *tet3*.

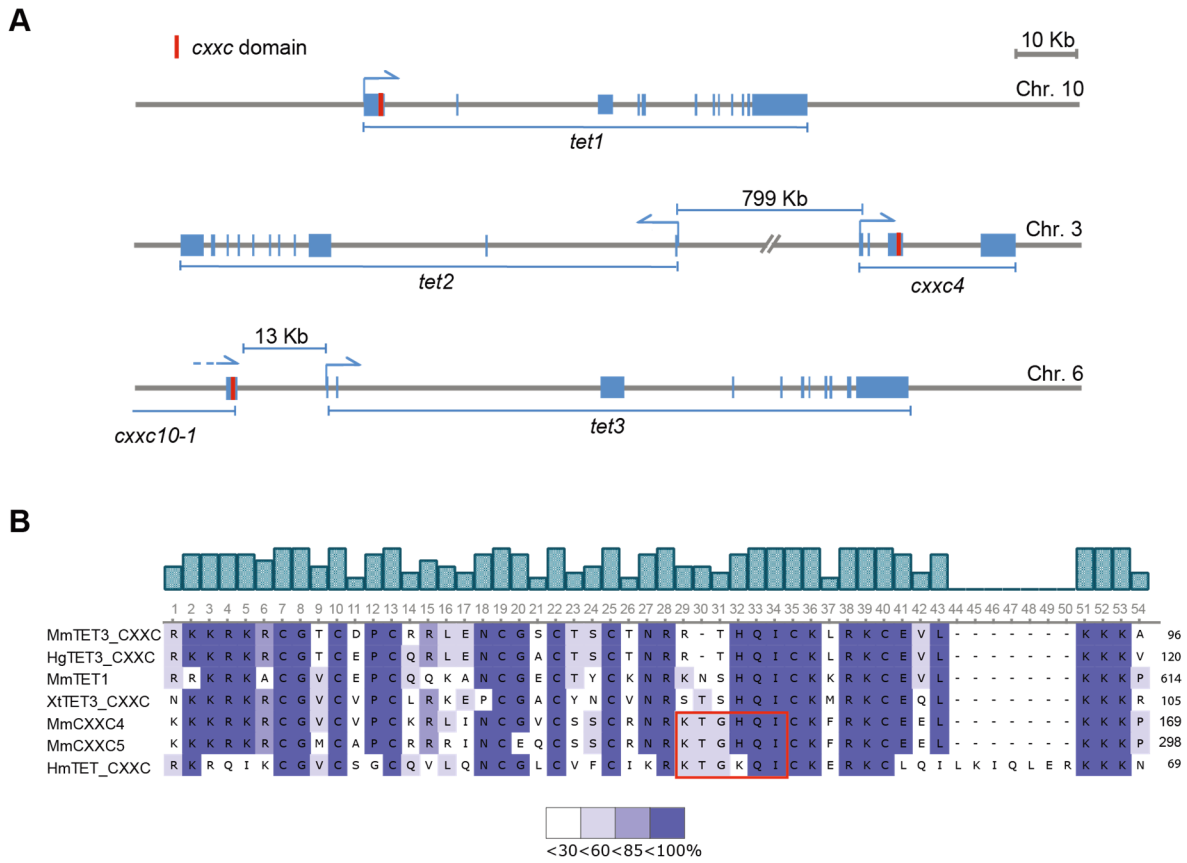

Figure 11. Genomic arrangement of mouse *tet* genes and adjacent *cxxc* loci and homology of *cxxc* domains from mouse *cxxc4*, *cxxc5* and *tet* homologues in various animal species. (A) Schematic representation of mouse *tet1*, *tet2/cxxc4* and *tet3/cxxc10* loci. Exons are depicted as blue rectangles. Annotated transcriptional start sites and transcription orientation are indicated with half arrows. (B) Alignment of CXXC domains from mouse CXXC4, CXXC5 and TET homologues in various animal species (Mm, *Mus Musculus*; Hg, *Heterocephalus glaber*; Xt, *Xenopus tropicalis*; Hm, *Hydra mangipallata*). The alignment was generated with Unipro UGENE (Okonechnikov, Golosova et al. 2012). Numbers on the right side indicate the position of the last amino acid in the corresponding protein. The KTXXXI motif, previously identified as determinant for the interaction of CXXC4 with DVL (London, Lee et al. 2004), is boxed. The scale at the bottom indicates the upper limit of percent identity represented by each color. GenBank accession numbers: MmCXXC10, JX946278; XtTET3, NP_001090656.1; HgTET3, EHB01729.1; MmTET1, NP_081660.1; MmCXXC4, NP_001004367; MmCXXC5, NP_598448; HmTET, XP_002161163.1. Data were published in Liu et al. 2013.

Thus, we set out to verify whether *tet3* transcripts including the *cxxc10-1* ORF are expressed in the mouse. To this aim we performed conventional PCR on total cDNA template from a neural stem cell (NSC) line derived by *in vitro* differentiation of E14 ESCs. We used primer pairs spanning from the *cxxc10-1* ORF to the *tet3* ORF in exon 3 according to the annotated *tet3* sequence. Cloning and sequencing of products identified two

alternative transcripts where the exon containing the *cxxc10-1* ORF is spliced to the first position of either exon 2 or exon 3 of the annotated *tet3* gene (Fig. 12A and 12B). These splicing events set the *cxxc10-1* ORF in frame with the annotated *tet3* coding sequence through its exon 2 and/or exon 3 sequences representing part of the 5'UTR in the annotated *tet3* transcript. Rapid amplification of cDNA 5' ends (RACE) identified a 5'UTR sequence upstream of the *cxxc10-1* ORF including an additional exon upstream of the one encoding the *cxxc10-1* ORF (Fig. 12A). To verify the expression and size of alternative *tet3* transcripts we first performed northern blotting of RNA from the same NSC line and parental ESCs (Fig. 12C). In NSCs a cDNA probe comprising exons 3-6 of the annotated *tet3* transcript detected two bands with estimated sizes of 10.9 and 11.6 kb, roughly corresponding to the sizes of the annotated *tet3* transcript and those encoding the *cxxc10-1* ORF, respectively, assuming the same splicing events downstream of the annotated exon 3 (Fig. 12A). A probe spanning the *cxxc10-1* ORF detected only the 11.6 kb band. Each of these probes detected the same respective bands in RNA from ESCs, but their intensity was much weaker than for NSCs (not visible in Fig. 12C) despite the same amount of RNA was loaded. We found no evidence for independent expression of the *cxxc10-1* sequence in these samples, as no other distinct band was detected in the blots (Fig. 12D). As final evidence for the expression of the *tet3* transcript including the *cxxc10-1* ORF and the annotated exon 2 (hereafter referred to as $tet3^{CXXC}L$) we amplified its entire coding sequence as a single fragment (5412 bp encoding a polypeptide of 1803 aa) using cDNA from NSCs as template and confirmed its primary structure by sequencing (NCBI accession number JX946278). These results show that the use of an alternative promoter and alternative splicing lead to the expression of *tet3* transcripts containing the *cxxc10-1* ORF (altogether referred to as $tet3^{CXXC}$) and that these transcripts share the same splicing organization with the previously annotated *tet3* transcript (hereafter referred to as *tet3*) downstream of its exons 2 ($tet3^{CXXC}L$) or 3 ($tet3^{CXXC}S$; Fig. 12A).

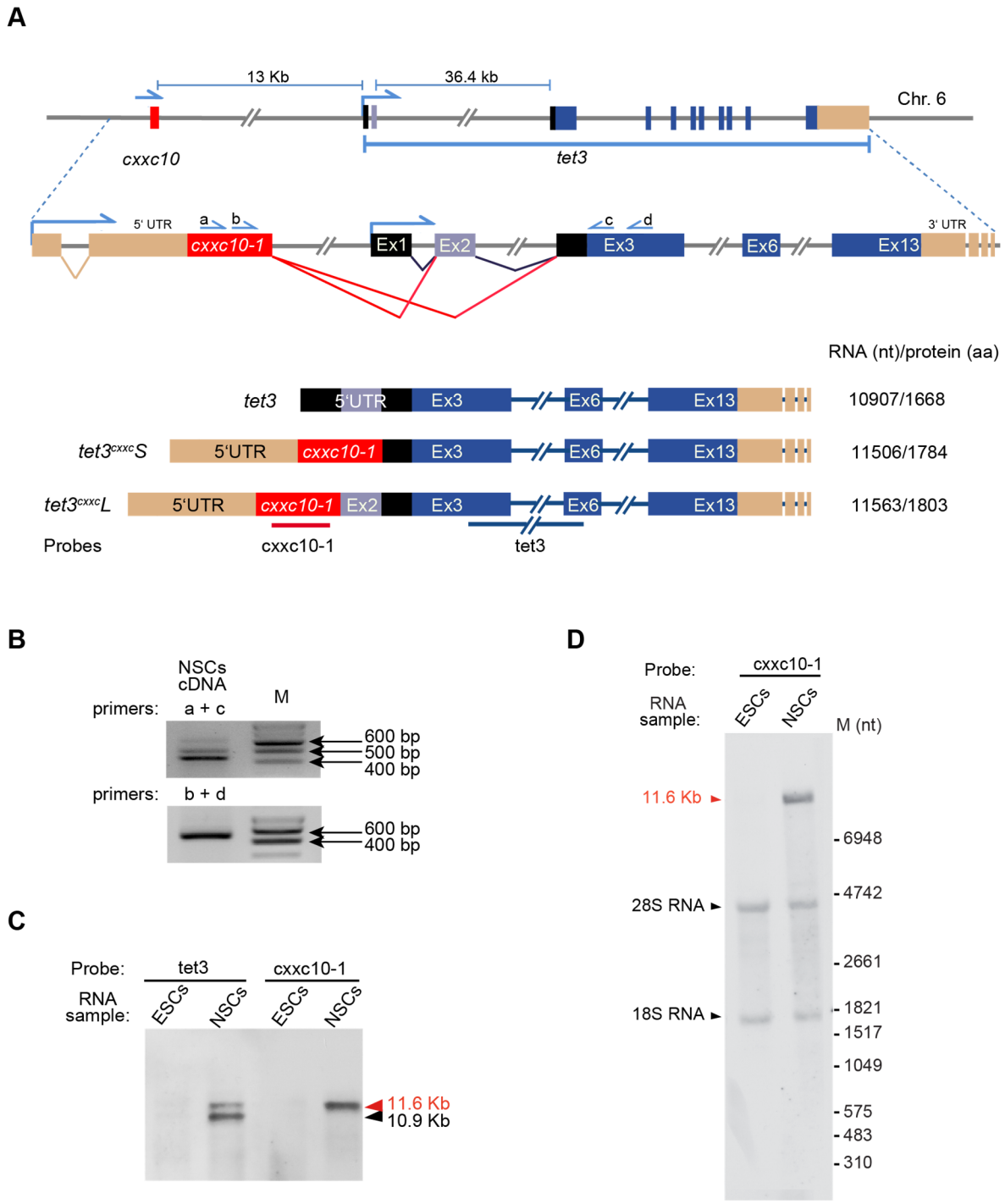

Figure 12. Identification of mouse *tet3* transcript variants encoding a *cxxc* domain. (A) Drawing illustrating the generation of alternative transcripts from the *tet3/cxxc10-1* locus. The positions of primers used in B are indicated. The lower part shows a schematic representation of alternative *tet3* transcripts. The positions of the probes used for northern blotting in C are indicated. (B) Amplification of fragments from NSCs cDNA identifying *tet3* transcripts that include the *cxxc10-1* ORF. (C) Northern blot detection of alternative *tet3* transcripts in ESCs and NSCs. (D) Northern blot analysis indicates no independent expression of the *cxxc10-1* sequence in these samples, as no distinct band was detected except alternative *tet3* transcripts. Data were published in Liu et al. 2013.

To characterize the expression patterns of *tet3* and *tet3*^{CXXC} transcripts we performed real time PCR (qPCR) on cDNAs from stem cell lines and various adult mouse tissues (Fig. 13A). We set primer pairs for selective amplification of the *tet3*^{CXXC} transcript including exon 2 of the *tet3* transcript, the *cxxc10-1* ORF and exons 1-3 of *tet3*. The levels of *tet3* and *tet3*^{CXXC} transcripts varied widely across the samples and were very low in ESCs, confirming our northern blot data. Notably, the ratio of *tet3* to *tet3*^{CXXC} transcripts was higher in brain regions relative to other tissues.

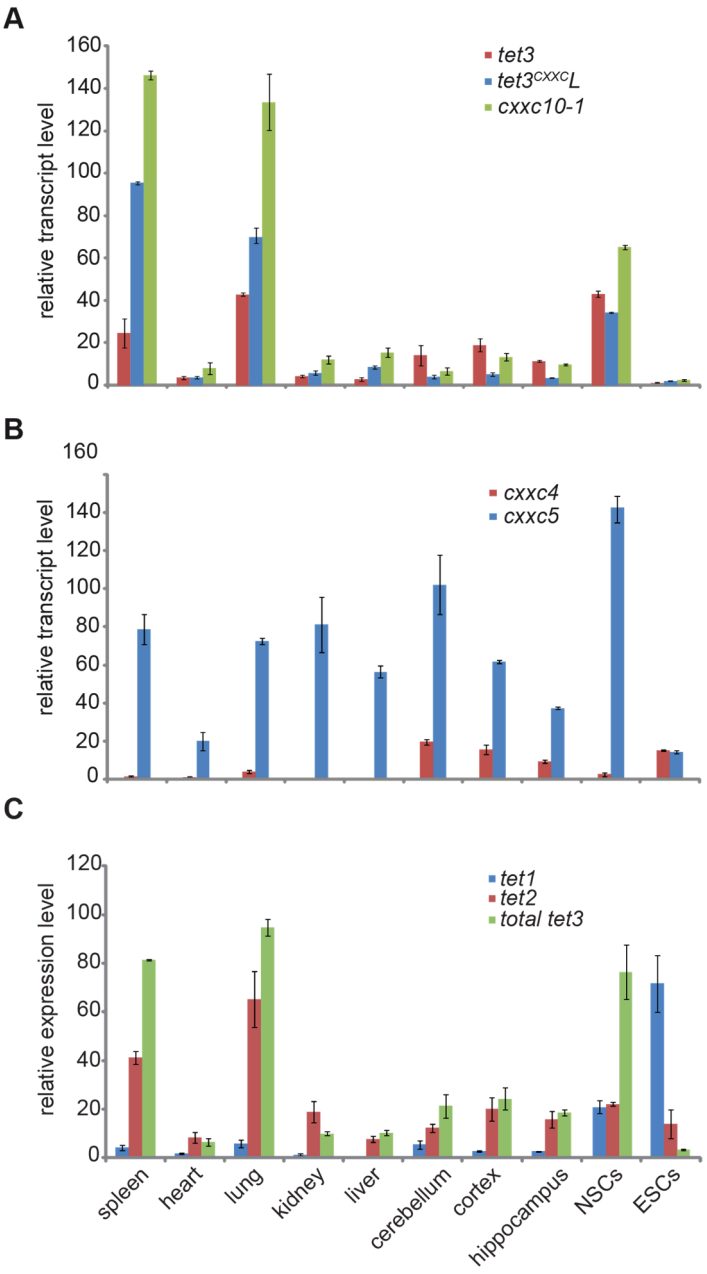

Figure 13. Transcript levels of *tet1-3*, *cxxc4* and *cxxc5* transcripts in mouse adult tissues, NSCs and ESCs. Transcript levels were determined by qPCR analysis of total cDNA. (A) Amplified fragments identify the *tet3* mRNA refseq NM_183138 (*tet3*), the alternative *tet3* transcript containing the *cxxc10-1* ORF and exon 2 of JX946278 (*tet3*^{CXXC}*L*) and all transcripts including the *cxxc10-1* ORF. The transcript levels of *tet3* and *tet3*^{CXXC}*L* varied widely across the samples and both were very low in ESCs, confirming our northern blot data. (B) Relative expression of *cxxc4* and *cxxc5* transcripts on mRNA level. *Cxxc4* transcript was detected mainly in the brain, while *cxxc5* was ubiquitously expressed across the samples. (C) Relative transcripts level of *tet1-3*. *Tet* transcripts show different expression levels across the samples. Low levels of *tet1* and high levels of *tet3* were detected in most differentiated tissues, while the opposite pattern was characterized in undifferentiated ESCs. Cumulative levels of all *tet3* transcripts were determined using a primer set spanning common sequences downstream exon 3 of the annotated *tet3* gene. Data relative to kidney, liver, cerebellum and cortex samples are from three biological replicates (two 6 week old 129Sv mice and a 30 week old C57BL/6 mouse). Data relative to spleen, heart, lung and hippocampus are from two biological replicates (a 6 week old 129/Sv mouse and a 30 week old C57BL/6 mouse). Data relative to NSCs and ESCs are from three independent cultures each. Shown are mean values and standard errors of the mean (SEM). Expression levels are all relative to *cxxc4* in heart (set to 1), so that values in these histograms are comparable. Data were published in Liu et al. 2013.

3.1.2 CXXC4 interacts with TET3 *in vivo* and is expressed in the adult brain

The evolutionary association of TET proteins with a distinct group of CXXC domains *in cis* raises the question as to whether they associate with this type of CXXC module also *in trans*. Therefore we probed the interaction of each of the three TET proteins with CXXC4 and CXXC5 using a mammalian fluorescent three hybrid assay (F3H). In this assay baits fused to GFP are anchored to a *lac* operator (*lac*O) array integrated in the genome of BHK cells and challenged with preys fused to a red fluorescent protein (Zolghadr, Mortusewicz et al. 2008, Meilinger, Fellinger et al. 2009, Dambacher, Deng et al. 2012, Herce, Deng et al. 2013). The colocalization of prey and bait at the *lac* operator array reflects their interaction (Fig. 14). CXXC4 showed interactions with all three TETs, but with a much stronger association with TET3 in comparison to TET1 and TET2 (Fig. 15). However, we could not detect coimmunoprecipitation of TET3 and CXXC4 fluorescent fusion constructs overexpressed in HEK293T cells. CXXC4 and 5 have been shown to antagonize canonical Wnt signaling by binding to cytoplasmic Disheveled (DVL) (Hino, Kishida et al. 2001, Michiue, Fukui et al. 2004, Andersson, Sodersten et al. 2009). However, expression of fluorescent fusions revealed a prevalently nuclear localization of CXXC4 in BHK cells, C2C12 myoblasts and ESCs (Liu, Wang et al. 2013). In this regard we noted that the KKKRK sequence (Fig. 11B) at the N-terminus of the CXXC domain in both CXXC4 and 5 is a

perfect match to the minimal prototypic nuclear localization sequence of the SV40 large T antigen (Jans, Xiao et al. 2000, Itoh, Brott et al. 2005), and that CXXC5 was also found to be predominantly nuclear in various cell types (Andersson, Sodersten et al. 2009, Pendino, Nguyen et al. 2009).

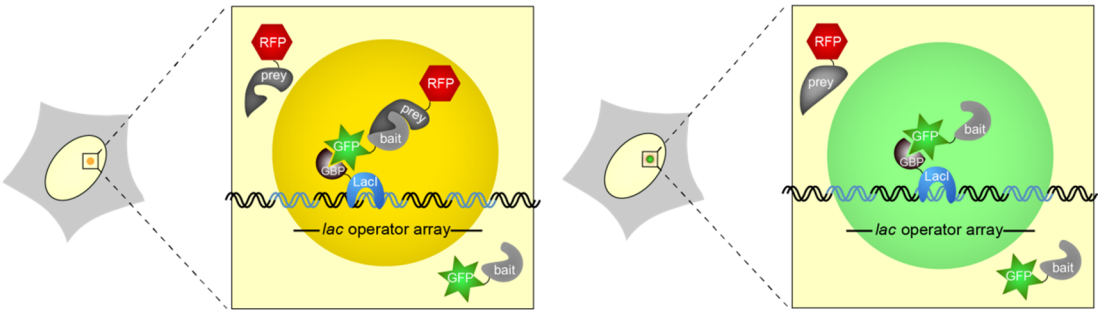

Figure 14. Schematic representation of the mammalian F3H assay. BHK cells containing *lac*O array on their genome are transfected with GBP-LacI fusion together with GFP-tagged bait and RFP-tagged prey. The Lac repressor (LacI) binds to its operator sequences and brings the GFP binding protein (GBP) to the *lac*O array. GFP-tagged bait proteins are recruited to the *lac*O loci by the GBP which is a single chain variable fragment antibody against GFP (Rothbauer, Zolghadr et al. 2006), thus giving out a green spot under microscope. If the RFP-tagged prey protein interacts with the bait, it would colocalize with the bait at *lac*O array, thus visualized as a yellow spot. While if there is no interaction between the bait and prey, the RFP-prey will disperse over the cell and only a green spot is observed (modified from Liu et al. 2013).

Next we determined the levels of *cxxc4* and *cxxc5* transcripts in adult mouse tissues and stem cell lines (Fig. 13B). Interestingly, among adult tissues *cxxc4* was expressed mainly in the brain, where *tet3* transcripts that do not encode the *cxxc* domain were more abundant relative to $tet3^{CXXC}$ transcripts. In contrast, *cxxc5* mRNA was detected ubiquitously and apart from ESCs its levels were substantially higher than those of *cxxc4*. No obvious correlation could be found between the levels of *cxxc5* transcripts and those of any of the *tet* transcripts analyzed (Fig. 13C).

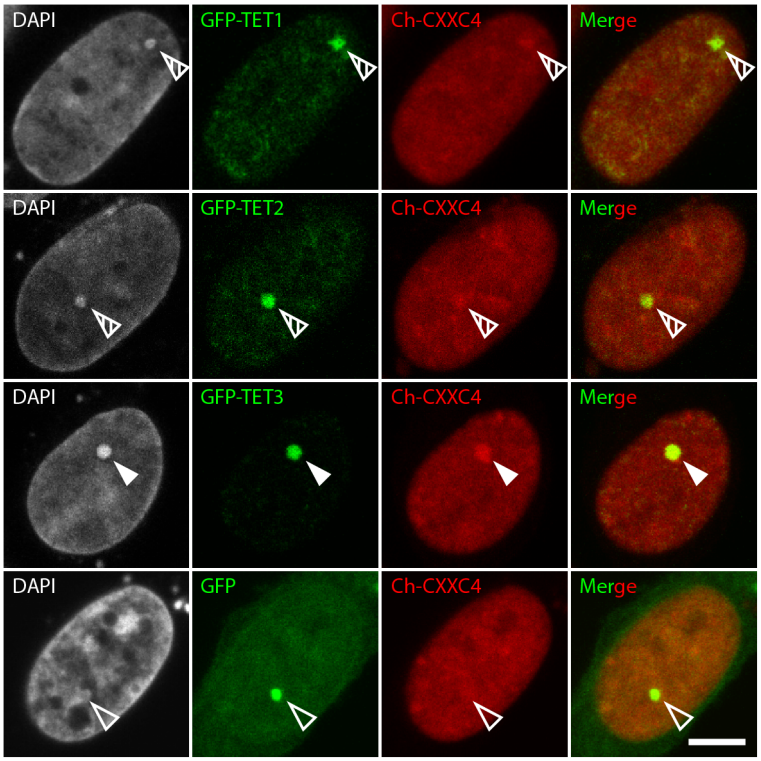

Figure 15. TETs and CXXC4 interact *in vivo*. Cells containing a *lac*O array were transfected with plasmids encoding a GBP-LacI fusion protein, GFP-TETs and Ch-CXXC4 proteins. These constructs were described in Table 2.1.5. The Lac repressor (LacI) binds to the *lac* operator array and recruits GFP-TETs through the GFP binding protein (GBP), which is visualized as a green spot. If Ch-CXXC4 protein intereacts with GFP-TETs, Ch-CXXC4 will consequently enrich at the *lac*O array and be visualized as yellow spots. Ch-CXXC4 shows a strong interaction (filled arrow) with GFP-TET3, a much weaker interaction (dashed arrow) with GFP-TET1 and GFP-TET2, and no unspecific binding with GFP (empty arrow). Scale bar: 5µm.

3.1.3 The CXXC domains of TET1, TET3^{CXXC}, CXXC4 and CXXC5 bind CpG containing DNA substrates

The high similarity shared by the CXXC domains of TET1, TET3^{CXXC} and the TET3 interactor CXXC4 prompted us to compare their DNA binding properties. CXXC4-GFP, CXXC5-GFP, GFP-TET1, CXXC^{TET1}-GFP as well as full length TET1, TET3 and TET3^{CXXC}L constructs with an N-terminal GFP tag were characterized with western blot assay (Fig. 16) and subjected to DNA binding assays (Fig. 17).

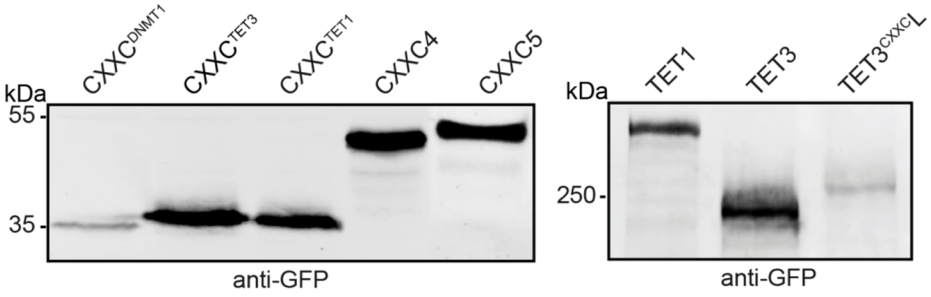

Figure 16. Western blot analysis of GFP fusion proteins for DNA binding assays. GFP-CXXC^{DNMT1}, CXXC^{TET3}-GFP, CXXC^{TET1}-GFP, CXXC4-GFP, CXXC5-GFP, GFP-TET1, GFP-TET3 and GFP-TET3^{CXXC}L were blotted with anti-GFP antibody. These constructs were described in Table 2.1.5. Data were published in Liu et al. 2013.

CXXC^{TET3}-GFP corresponds to the isolated CXXC domain of the CXXC10-1 ORF with GFP fused to its C-terminus and is therefore analogous to CXXC^{TET1}-GFP. Although we could not detect interactions between TET proteins and CXXC5, we investigated the DNA binding potential of the latter as its CXXC domain is also highly homologous to that of TET1. CXXC domains belonging to a distinct homology class, including the CXXC domain of DNMT1 (CXXC^{DNMT1}), were shown to preferentially bind CpG-containing sequences (Lee, Voo et al. 2001, Birke, Schreiner et al. 2002, Jorgensen, Ben-Porath et al. 2004, Pradhan, Esteve et al. 2008, Blackledge, Zhou et al. 2010, Frauer, Rottach et al. 2011). Therefore, we first determined the binding preference of our constructs with respect to DNA substrates differing only for the presence or absence of a single central CpG site and compared it to that of the CXXC domain of DNMT1. CXXC4, CXXC5 and all TET constructs showed higher DNA binding activity as well as similar and substantial preference for the substrate containing a CpG site as compared to GFP-CXXC^{DNMT1} (Fig. 17A). We then determined the binding preference with respect to substrates containing a single central CpG site with defined cytosine modifications as shown for CXXC^{TET1} constructs. CXXC4-GFP, CXXC5-GFP and CXXC^{TET3}-GFP displayed similar binding properties, with decreasing preference for substrates with the unmodified, symmetrically methylated and symmetrically hydroxymethylated CpG site. In contrast, CXXC^{TET1}-GFP did not discriminate between substrates with unmodified and symmetrically methylated CpG. In the case of full length TET1, TET3 and TET3^{CXXC}L constructs, incubation with a 4-fold molar excess of DNA substrates is expected to minimize potential competition among multiple DNA

binding sites. GFP-TET1 displayed the same substrate preference as the isolated CXXC domain of TET1 (CXXC^{TET1}-GFP), albeit with an 8-fold increase in binding activity, indicating that sequences outside the CXXC domain contribute to the affinity for DNA without altering the substrate preference. In contrast, both GFP-TET3 and GFP-TET3^{CXXC}L showed a relative increase in binding activity toward the substrate with the methylated CpG site as compared to CXXC^{TET3}-GFP (Fig. 17B). Thus, in TET3^{CXXC}L features outside the CXXC domain override the binding preference of the latter.

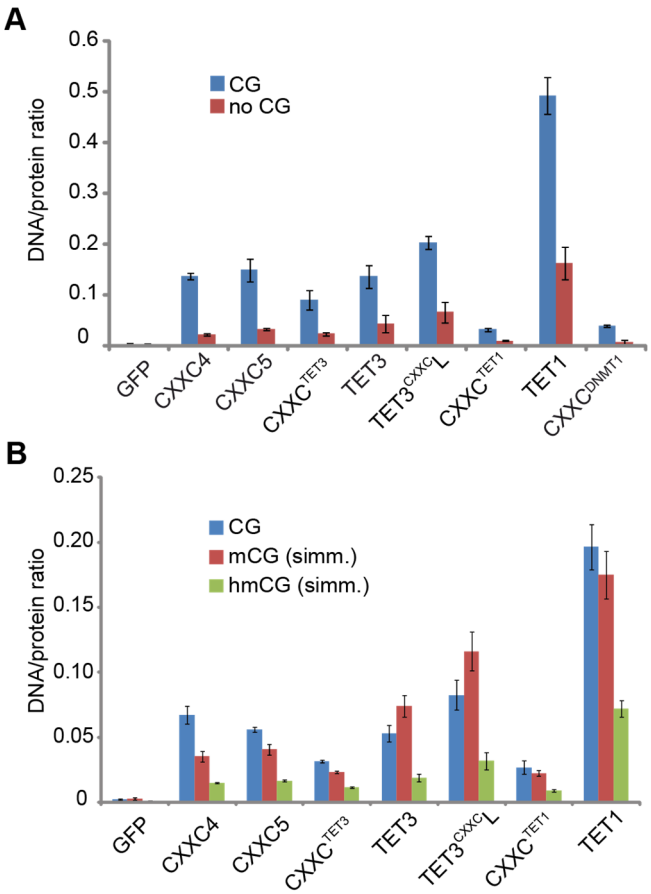

Figure 17. *In vitro* DNA binding properties of CXXC4 and 5, isolated CXXC domains and full length constructs of TET1 and TET3^{CXXC}. (A) All proteins were expressed as GFP fusion constructs in HEK293T cells and affinity purified using a GFP-trap. Fluorescently labeled DNA substrates containing one or no CG site were incubated in direct competition. Shown are mean values of bound substrate/protein ratios and SEM from n independent replicate experiments: GFP and CXXC^{TET3}-GFP, n = 5; GFP-TET1, CXXC4-GFP, CXXC5-GFP and GFP-CXXC^{DNMT1}, n = 4; GFP-TET3, GFP-TET3^{CXXC}L and CXXC^{TET1}-GFP, n = 3. (B) All constructs are GFP fusions. Fluorescently labeled DNA substrates with the same sequence and a single CpG site either unmethylated, symmetrically methylated or symmetrically hydroxymethylated were incubated in direct competition. Shown are mean values of bound substrate/protein ratios and SEM from n independent replicate experiments: TET1, n = 10; TET3, CXXC^{TET3}, n = 6; TET3^{CXXC}L, n = 7; CXXC^{TET1}, CXXC4 and GFP, n = 3; CXXC5, n = 2. Data were published in Liu et al. 2013.

3.1.4 TET3^{CXXC} oxidizes genomic 5mC *in vivo* and shows slightly lower mobility than the TET3 isoform lacking the CXXC domain

We then compared the activity of TET1 and TET3 isoforms with or without CXXC domain by determining global levels of genomic 5hmC in HEK293T cells transiently transfected with GFP-tagged constructs (Fig. 18A). A similar increase of 5hmC levels was observed in cells transfected with GFP-TET1, GFP-TET3 and GFP-TET3^{CXXC}L, the latter slightly showing higher conversion of 5mC to 5hmC.

As further characterization of TET3 isoforms we compared nuclear localization and mobility of GFP-TET3 and GFP-TET3^{CXXC}L in C2C12 myoblasts (Figure 18B and C). Both constructs were diffusely distributed throughout the nucleus with exclusion of nucleoli and large clusters of pericentric heterochromatin (chromocenters; Fig. 18B). After photobleaching half of the nucleus the fluorescence of GFP-TET3^{CXXC}L recovered more slowly and reached a plateau at a lower level than that of GFP-TET3 (Fig. 18C). These differences were small, but reproducible.

Thus, the presence of the CXXC domain in TET3 does not affect and perhaps promotes conversion of 5mC to 5hmC, while it reduces its mobility and slightly increases the immobile fraction, suggesting that the CXXC domain contributes to additional nuclear interactions.

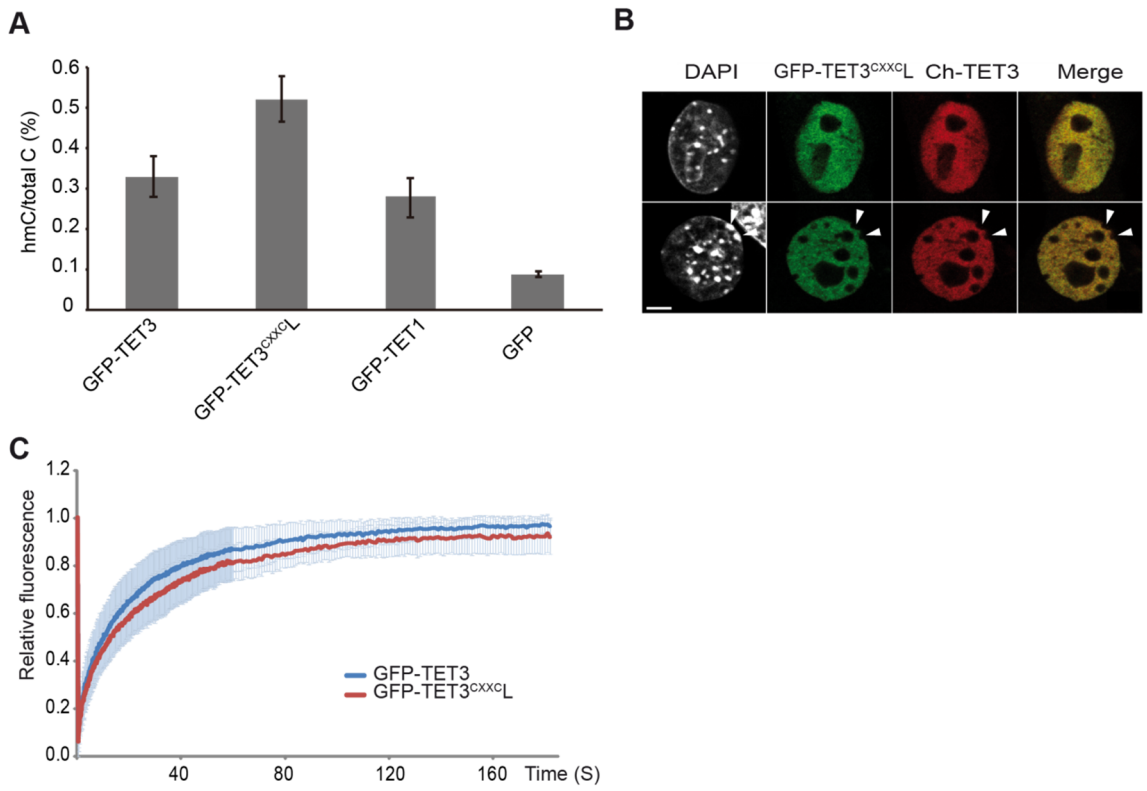

Figure 18. Enzymatic activity, localization and mobility of TET3^{CXXC} in somatic cells. (A) TET3^{CXXC} oxidizes genomic 5mC *in vivo*. GFP or GFP-TET fusions were transiently overexpressed in HEK293T cells and genomic 5hmC levels were determined using an *in vitro* glucosylation assay with T4 β-glucosyltransferase and UDP-[^{3}H] glucose. Shown are mean percentages and SEM of 5hmC over total C from 2 (GFP-TET1) or 3 (all others) independent transfections. (B) Optical sections of fixed C2C12 cotransfected with GFP-TET3^{CXXC}L and Ch-TET3 constructs as indicated. Arrowheads indicate the position of large chromocenters from which GFP-TET3^{CXXC}L and Ch-TET3 signals are excluded. Scale bar: 5µm. (C) FRAP curves of GFP-TET3 and GFP-TET3^{CXXC}L in transiently transfected C2C12 myoblasts. Images were taken every 150 ms in the first 60 s, and then at intervals of 1 s for the next 120 s. Shown are mean values and SEM from 12 (GFP-TET3) and 10 cells (GFP-TET3^{CXXC}L). Live cell imaging and FRAP analysis was performed as described (Schermelleh et al., 2007, Nucl Acids Res 35: 4301) with the following minor modifications. The images were Gauss-filtered (2 pixel radius) and data sets showing lateral movement were corrected by image registration using the StackReg plug-in of ImageJ, starting with a time frame where approximately half recovery was reached. Experiments for A and C were performed by Mengxi Wang and were published in Liu et al. 2013.

3.1.5 Detection of TET3^{CXXC} at the protein level

To detect TET3^{CXXC} at the protein level, we raised polyclonal antibodies against TET3^{CXXC}. Since TET3^{CXXC} shares the same exons downstream of the CXXC10-1 sequences with the annotated TET3, a peptide within CXXC10-1 region was selected as antigen (Fig. 19A). After expression and purification, His-tagged peptide was sent to Biogenes company for generating polyclonal antibodies against TET3^{CXXC}. From immunizations we obtained

specific rabbit polyclonal antibodies that recognize GFP-TET3^{CXXC} but not GFP-TET3 in immunofluorescence staining and western blot (Fig. 19B and 19C). Notebly, the TET3^{CXXC} polyclonal antibodies are able to detect endogenous TET3^{CXXC} in somatic cells (Fig. 19C). These results indicate that the TET3^{CXXC} polyclonal antibodies are well suited to further investigate the function of TET3^{CXXC}.

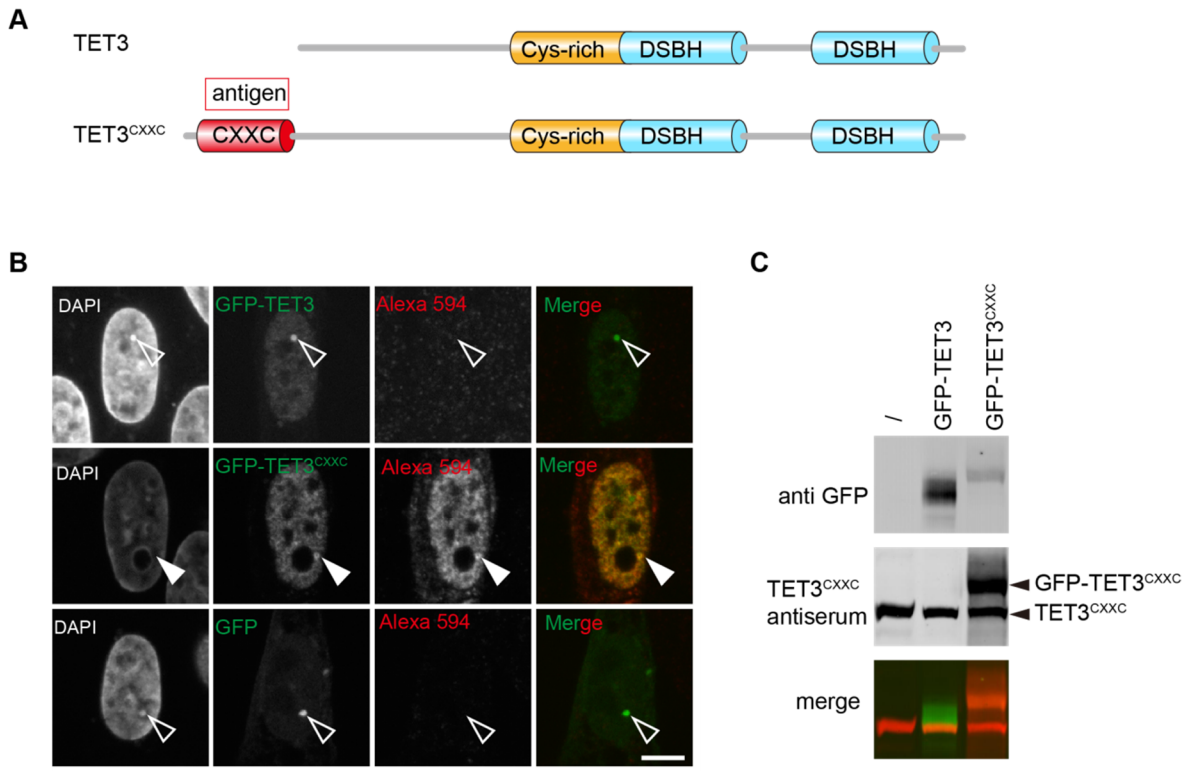

Figure 19. Characterization of TET3^{CXXC} polyclonal antiserum. (A) Antigen selected for immunization. TET3 and TET3^{CXXC} were showed as schematic outline, and CXXC motif of TET3^{CXXC} was chosen as antigen. (B) TET3^{CXXC} polyclonal antiserum recognized GFP-TET3^{CXXC} but not GFP-TET3 in immunofluorescence staining. Optical sections of fixed BHK cells containing a *lac*O array cotransfected with plasmids encoding a GBP-LacI fusion protein and GFP-TET fusions as indicated. The Lac repressor binds to the *lac* operator array and recruits GFP-TETs through the GBP, which is visualized as a green spot. TET3^{CXXC} polyclonal antiserum will enrich at the *lac*O array if it recognizes GFP-TETs, which will be visualized as yellow spot after the detection with Alexa Fluor 594 labeled anti-rabbit secondary antibody. TET3^{CXXC} polyclonal antiserum recognized GFP-TET3^{CXXC} specifically (filled arrow), and did not recognize GFP-TET3 or GFP alone (empty arrow). GFP was used to control unspecific binding and DAPI was used to show the nucleus. Scale bar: 5µm. (C) TET3^{CXXC} polyclonal antiserum recognized GFP-TET3^{CXXC} and endogenous TET3^{CXXC} in western blot assay. Cell lysates from HEK293T cells expressing GFP-TET3 or GFP-TET3^{CXXC} were detected with TET3^{CXXC} polyclonal antiserum. The same blot probed with an anti-GFP antibody was used as a loading control.

3.2 DNMT1 ubiquitin interacting motif UIM is required for maintenance of DNA methylation by binding to UHRF1-dependent H3 ubiquitination

The multi-functional protein UHRF1 is a key factor for regulation of maintenance DNA methylation by recruiting DNMT1 to its substrates. The association of UHRF1 with chromatin is mediated by its chromatin binding domains, the hemimethylated CpG binding SRA domain, the trimethylated H3K9 binding TTD domain and the unmethylated H3R2 binding PHD domain (Arita, Ariyoshi et al. 2008, Avvakumov, Walker et al. 2008, Hashimoto, Horton et al. 2008, Qian, Li et al. 2008, Nady, Lemak et al. 2011, Xie, Jakoncic et al. 2012, Liu, Gao et al. 2013, Rothbart, Dickson et al. 2013). Cooperative binding of the TTD domain of UHRF1 to di- and trimethylated histone H3K9 and of the SRA domain to hemimethylated DNA was described as a prerequisite for targeting DNMT1 to its substrate and for subsequent DNA methylation (Liu, Gao et al. 2013). Given the regulatory impact of these two domains, we were interested in how the PHD and RING domain of UHRF1 may functionally contribute to maintenance DNA methylation by DNMT1. Data from Dr. Weihua Qin and Patricia Wolf showed that the H3 ubiquitination in dependence on the PHD-mediated histone-binding and the RING-mediated ubiquitin E3 ligase activity of UHRF1 serves as an indirect targeting mechanism of DNMT1 to chromatin.

3.2.1 UHRF1 ubiquitinates histone H3 on K18 and K23 residues in mammalian cells

Using *Xenopus* extracts immunodepleted for DNMT1, H3 was shown to be ubiquitinated at K23 residue (Nishiyama, Yamaguchi et al. 2013). To map ubiquitination sites on histone H3 tails in mammalian cells, we performed immunoprecipitations followed by mass spectrometry. In contrast to the latest publication, the K18 residue of histone H3 was identified as the ubiquitination site (data not shown; manuscript sumitted). In consistence with the study on *Xenopus* extracts, comparison of ubiquitination levels showed that both, single K18 and K23 mutations to alanine (A) decreased the ubiquitination, suggesting that

both residues are potential UHRF1 ubiquitination targets (Fig. 20). Notably, the decrease in H3 ubiquitination was most prominent for the double mutant GFP-H3 K18A-K23A. Furthermore, we could confirm that the R2 residue of H3 plays a role in H3 ubiquitination by UHRF1 as the mutant GFP-H3 R2A resulted in reduced ubiquitination levels (Fig. 20).

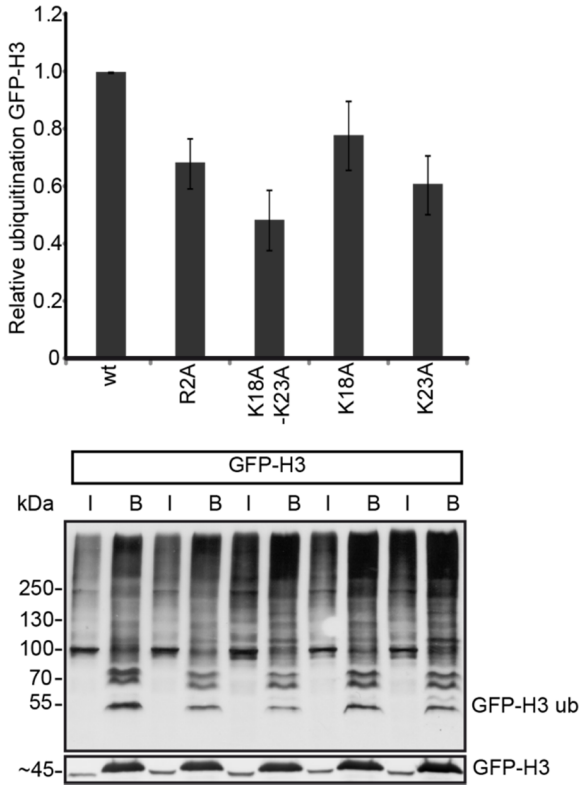

Figure 20. UHRF1 ubiquitinates histone H3 at K18 and K23 residues in dependence on R2. Mapping and quantification of UHRF1 ubiquitination target sites in H3 N-terminal tails and dependence of the ubiquitination on R2. GFP-H3 constructs carrying R2A, K18A, K23A and K18A-K23A mutations were transiently co-expressed in HEK293T cells with HA-ubiquitin and after immunoprecipitation with the GFP-Trap, ubiquitinated GFP-H3 was detected by western blot with an anti-HA antibody. Equal loading of GFP-H3 (~45 kDa) is shown by the anti-GFP blot below. Quantifications were done with ImageJ. Shown are mean values of four to five independent experiments, error bars indicate SEM.

3.2.2 The DNMT1 UIM mediated binding to ubiquitinated H3 is required for methylation activity *in vivo*

To unravel how H3 ubiquitination may contribute to maintenance DNA methylation, we screened DNMT1 for potential binding motifs. Intriguingly, we found that the N-terminal regulatory domain of DNMT1 contains an ubiquitin interacting motif (UIM). This motif is

located in a region spanning from amino acid 380 to 399 of mouse DNMT1 and shows striking similarity to UIMs of known ubiquitin interacting proteins (Fig. 21A). We were interested in investigating whether UIM of DNMT1 plays a role in DNA methylation maintenance. Therefore, we generated GFP-DNMT1 mutants lacking the UIM (GFP-DNMT1 Δ356-404). We first tested the interaction between DNMT1 UIM mutant and UHRF1 and the result indicated that GFP-DNMT1 Δ356-404 is still able to interact with Ch-UHRF1 (Fig. 21B). Notably, in comparison to GFP-DNMT1 wt, the interaction between DNMT1 and UHRF1 is dramatically increased when lacking the UIM domain, which could be because the deletion of UIM somehow released the autoinhibitory role of TS domain, therefore leading to the enhanced association between DNMT1 and UHRF1. However, further comparison on the ubiquitin binding properties between GFP-DNMT1 wt and GFP-DNMT1 Δ356-404 as well as functional complementation assay in *dnmt1*$^{-/-}$ ESC lines clearly revealed the GFP-DNMT1 Δ356-404 showed a defect in the association with ubiquitinated histone H3 and was not able to reestablish DNA methylation patterns (data not shown; manuscript sumitted), which indicated a key role of the UIM in DNMT1 targeting via ubiquitinated histone H3 binding and for maintenance DNA methylation in mammalian cells.

A

DNMT1. 380-399	V	D	E	P	Q	M	L	T	S	E	K	L	S	I	Y	D	S	T	S	T
EPS15. 852-871	S	E	E	D	M	I	E	W	A	K	R	E	S	E	R	E	E	E	Q	R
EPS15. 878-897	Q	E	Q	E	D	L	E	L	A	I	A	L	S	K	S	E	I	S	E	A
EP15R. 863-882	N	E	E	Q	Q	L	A	W	A	K	R	E	S	E	K	A	E	Q	E	R
EP15R. 889-907	Q	E	Q	E	D	L	E	L	A	I	A	L	S	K	A	D	M	P	A	-
STAM2. 165-184	K	E	D	E	D	I	A	K	A	I	E	L	S	L	Q	E	Q	K	Q	Q
PSMD4. 211-230	S	A	D	P	E	L	A	L	A	L	R	V	S	M	E	E	Q	R	Q	R
PSMD4. 282-301	T	E	E	E	Q	I	A	Y	A	M	Q	M	S	L	Q	G	T	E	F	S
ATX3. 224-243	E	D	E	D	D	L	Q	R	A	L	A	I	S	R	Q	E	I	D	M	E
ATX3. 244-263	D	E	E	A	D	L	R	R	A	I	Q	L	S	M	Q	G	S	S	R	S
DNJB2. 250-269	S	E	D	E	D	L	Q	L	A	M	A	Y	S	L	S	E	M	E	A	A
UBP37. 806-825	R	E	E	Q	E	L	Q	Q	A	L	A	Q	S	L	Q	E	Q	E	A	W
Consensus	X	e	e	e	X	Φ	X	X	A	X	X	X	S	X	X	e	X	X	X	X

B

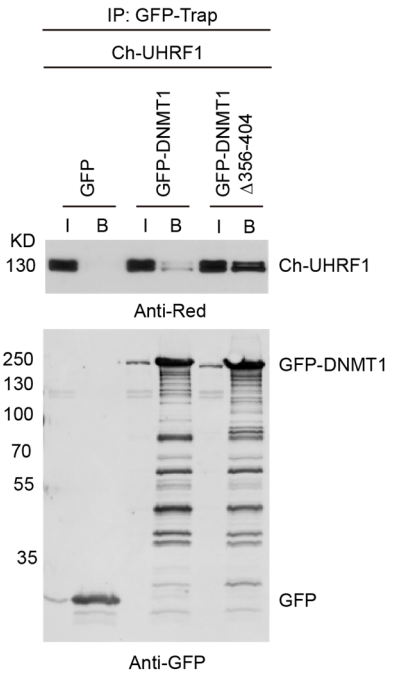

Figure 21. The TS domain of DNMT1 harbors an ubiquitin interacting motif which is not required for the interaction between UHRF1 and DNMT1. (A) Schematic outline of the ubiquitin interacting motif UIM in the TS domain of DNMT1. A peptide sequence of DNMT1 encompassing amino acid 380 to 399 was aligned with peptide sequences of proteins previously known to carry UIMs. The UIM consensus sequence is shown below, 'Φ' denotes hydrophobic residue and 'e' denotes negatively charged residues. UniProtKB accession numbers: DNMT1, P13864; EPS15 (epidermal growth factor receptor substrate 15), P42567; EP15R (epidermal growth factor receptor substrate 15-like 1), Q60902; STAM2 (signal transducing adapter molecule 2), O88811; PSMD4 (26S proteasome non-ATPase regulatory subunit 4), O35226; ATX3 (ataxin-3), Q9CVD2; DNJB2 (DnaJ homolog subfamily B member 2), Q9QYI5; UBP37 (ubiquitin carboxyl-terminal hydrolase 37), Q8C0R0. (B) Co-immunoprecipitation assay of GFP-DNMT1 and Ch-UHRF1 *in vitro*. GFP-DNMT1 wt or Δ 356-404 were co-expressed with Ch-UHRF1 and after co-immunoprecipitation using the GFP-Trap, the bound fractions were detected by western blot with specific antibodies against GFP and RFP. GFP was used as negative control.

3.3 UHRF1 and USP7 control stability of heterochromatin protein CBX1 by polyubiquitination

3.3.1 UHRF1 associates and colocalizes with all three CBXs

UHRF1 is an essential factor of DNA methylation, which recruits and modifies DNMT1, but little is known about other cellular targets of its RING type ubiquitin E3 ligase activity. To identify novel ubiquitination substrates, we fused two Uba domains with GFP for detection and precipitation with a nanobody based GFP-Trap. To test the precipitation efficiency, we transiently expressed HA-tagged ubiquitin together with GFP-2Uba in HEK293T cells and analyzed the coprecipitated ubiquitinated proteins with an HA antibody. The comparison with the GFP control showed that GFP-2Uba was highly efficient in the specific precipitation of ubiquitinated proteins (Dr. Weihua Qin, personal communication).

We next combined this GFP-2Uba pull down approach with quantitative mass spectrometry (MS) and identified many UHRF1-dependent ubiquitination targets which relate to cell cycle, development, RNA processing and heterochromatin formation pathways. In this study we focused on the heterochromatin proteins CBX1 and CBX3 as they represent possible new links to UHRF1 and epigenetic regulation.

To independently confirm the MS results, we tested the physical interactions between UHRF1 and CBX proteins. Since CBX5 shares high sequence and structure similarity with CBX1 and CBX3, it was also included in the test. We transiently co-expressed UHRF1-GFP with RFP-CBXs or RFP in HEK293T cells and immunoprecipitated protein complexes with the RFP-trap, and UHRF1 coimmunoprecipitated with all three CBXs (Fig. 22A). In addition, with immunofuorescence we also detected the colocalization at heterochromatin region in the cells between all CBXs and UHRF1 (Fig. 22B).

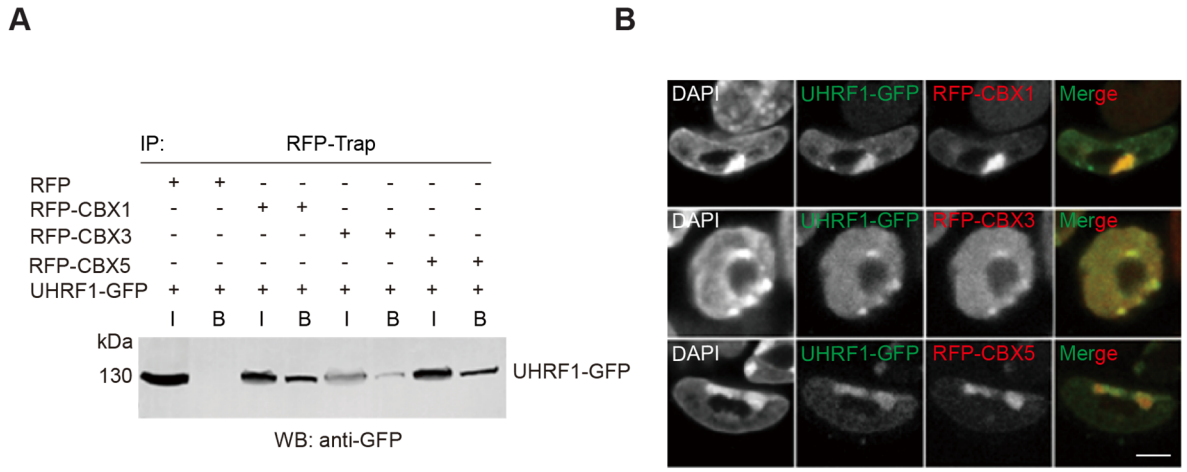

Figure 22. UHRF1 associates and colocalizes with all three CBXs. (A) UHRF1 interacts with all three CBX proteins. RFP-CBX proteins were immunoprecipitated from cells coexpressing UHRF1-GFP with the RFP-Trap and subjected to western blot analysis with an anti-GFP antibody. (B) Confocal images show colocalization of all three CBX proteins and UHRF1 at heterochromatin region. Wild type mESCs were transiently cotransfected with constructs coding for RFP-CBX proteins and UHRF1-GFP and were prepared for confocal laser scanning microscopy. Nuclear staining was performed with DAPI. Scale bar: 5 µm.

3.3.2 UHRF1 is the E3 ligase for ubiquitination of CBX1

Though UHRF1 shows association and colocalization with all three CBX proteins, only GFP-CBX1 and GFP-CBX3 showed strong ubiquitination *in vivo* with a F3H assay, and no ubiquitination was detected for GFP-CBX5 (data not shown). Therefore, we focused on CBX1 and CBX3 in the following investigations.

The interaction and colocalization data raise the possibility that UHRF1 is the ubiquitin E3 ligase for CBX1 and CBX3 modification. To investigate the ubiquitination of CBX1 and CBX3 *in vivo*, we coexpressed GFP-CBX1 or GFP-CBX3 with combinations of HA-ubiquitin and Ch-UHRF1 in HEK293T cells. Expression of UHRF1 showed a clear increase in ubiquitinated GFP-CBX1 (Fig. 23A). Using the same assay we could not detect ubiquitination of CBX3 by UHRF1 (Fig. 23B). These experiments demonstrated that UHRF1 acts as a ubiquitin E3 ligase for the modification of CBX1.

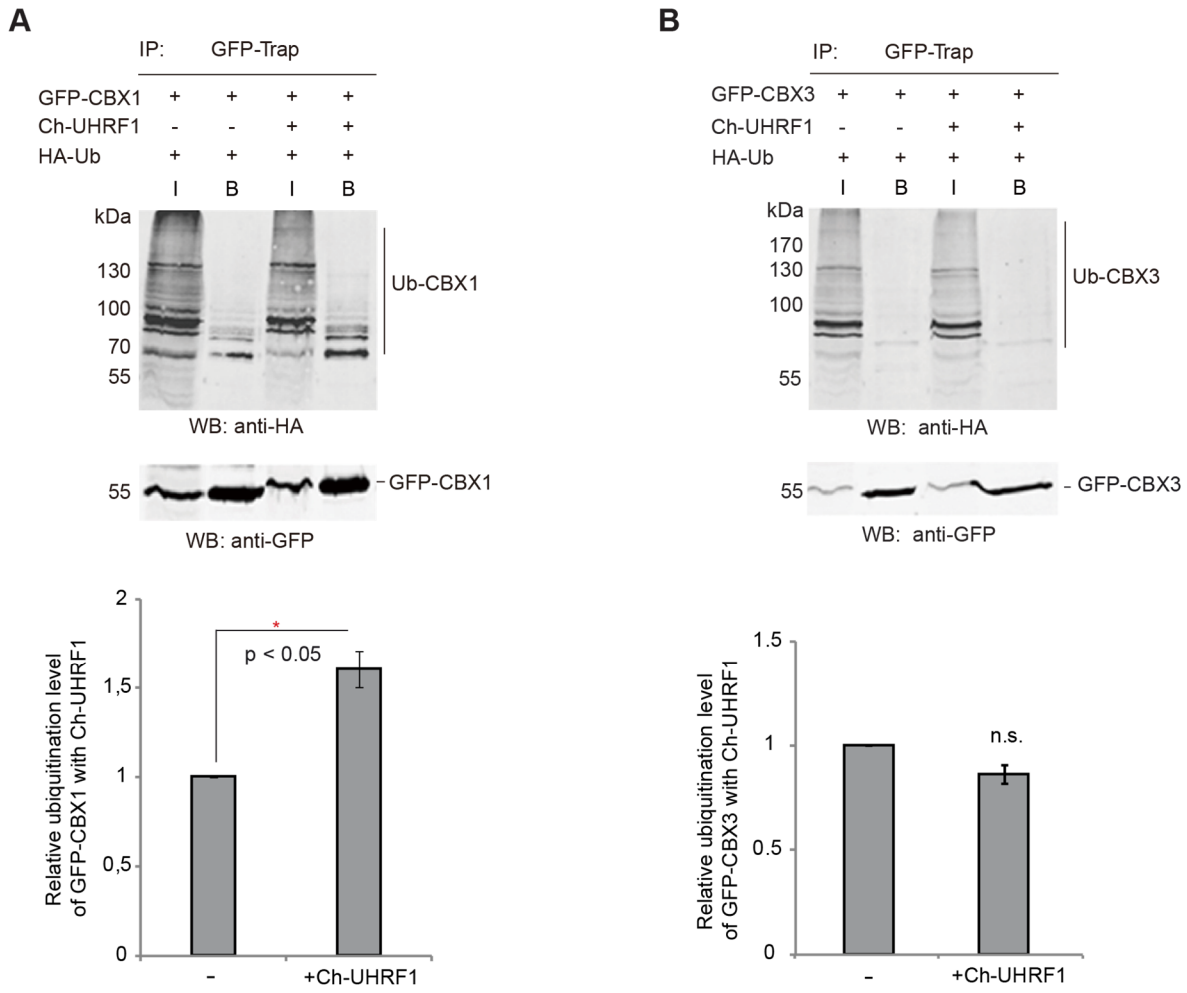

Figure 23. UHRF1 specifically ubiquitinates CBX1. (A) Ubiquitination of CBX1 by UHRF1. GFP-Trap immunoprecipitations from HEK293T cells expressing indicated combinations of HA-ubiquitin, GFP-CBX1 and Ch-UHRF1 were detected with an anti-HA antibody. (B) UHRF1 cannot ubiquitinate CBX3 *in vivo*. GFP-CBX3 was coexpressed in HEK293T cells with combinations of HA-ubiquitin and Ch-UHRF1. GFP-Trap pulldowns were probed with an anti-HA antibody. The blot probed with an anti-GFP antibody was used as a loading control. Quantifications were done with ImageJ. Shown are mean values of three independent experiments, error bars indicate SEM. Statistical difference between values was estimated by t test; statistical differences are marked by asterisks (*<0.05) or not significant (n.s.).

To identify potential ubiquitination sites we aligned the three CBX protein sequences and found three lysine residues at the C-terminus of CBX1 not present in CBX3 or CBX5 (Fig. 24A). To investigate whether the C-terminal lysine residues are the ubiquitination sites, we tested the ubiquitination levels of GFP-CBX1 carrying single lysine to alanine mutations (K180R, K181R or K184R) or the deletion of the last six amino acids including all

three lysine residues (GFP-CBX1^{delC}) (Fig. 24B). While single lysine mutations had little to no effect, GFP-CBX1^{delC} caused a clear reduction in CBX1 ubiquitination (Fig. 25A and 25B).

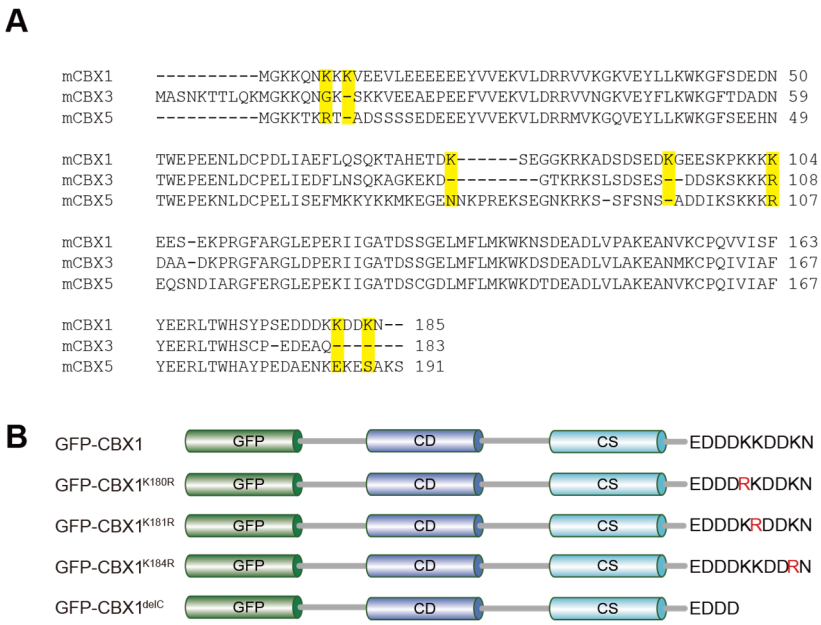

Figure 24. Alignment of mouse heterochromatin protein CBXs and schematic structure of CBX1 and its mutants. (A) Alignment of mouse heterochromatin proteins CBX1, CBX3 and CBX5. The lysine residues specific for CBX1 are highlighed in yellow. Accession number: CBX1 NP_031648.1; CBX3 AAI10377; CBX5 AAH04707. (B) Schematic structure of CBX1 and its mutants. Single lysine to alanine mutations (K180R, K181R or K184R) or the deletion of the last six amino acids including all three lysine residues (GFP-CBX1^{delC}) were generated to test whether these mutations affect the ubiquitination of GFP-CBX1.

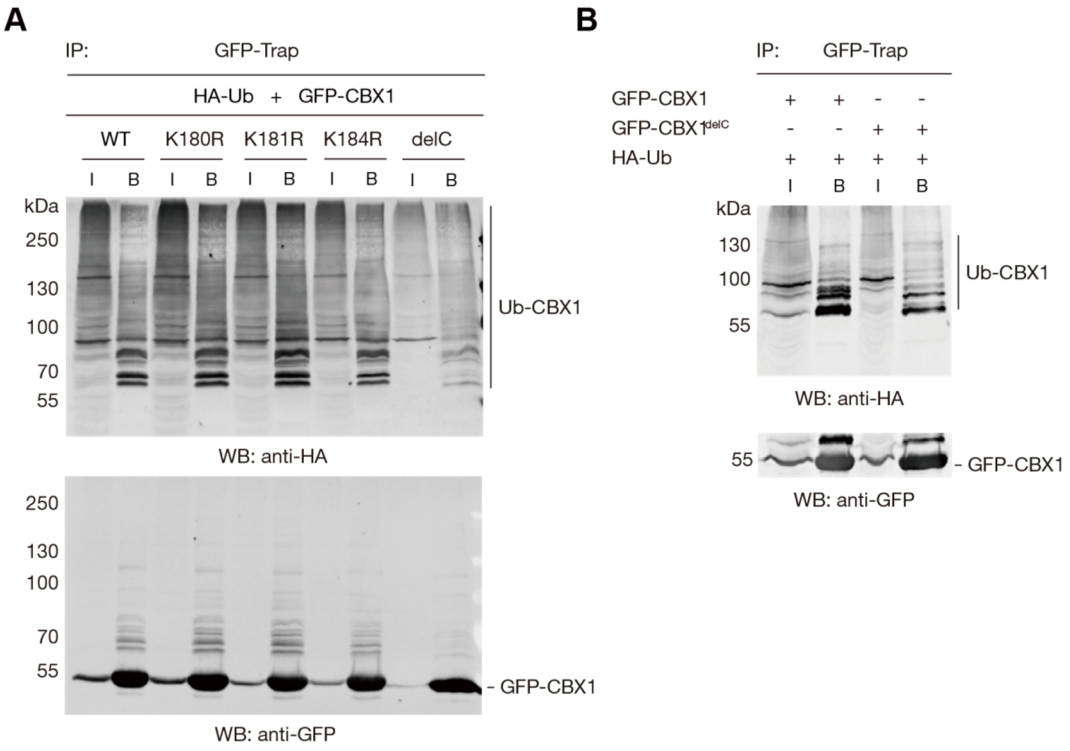

Figure 25. Characterization of CBX1 substitution mutants to narrow down ubiquitination sites in CBX1 by *in vivo* ubiquitination assay. (A) GFP-CBX1 and its mutants were coexpressed with HA-ubiquitin in HEK293T cells and immunoprecipitated using the GFP-Trap. Immunoprecipitations were analyzed with an anti-HA antibody. The GFP-CBX1^{delC} had a clear effect on the ubiquitination of CBX1, whereas single lysine substitutions (K180R, K181R or K184R) had little to no effect with ubiquitination levels similar to the wild type protein. (B) Immunoprecipitations comparing GFP-CBX1 wt and C-terminal deletion mutant (GFP-CBX1^{delC}) further indicated that CBX1 ubiquitination occurs at its C-terminus.

3.3.3 Association between UHRF1 and CBX1 is independent of PxVxV motif

To map the domains that mediate the interaction with CBX1, we generated several expression vectors encoding the five distinct UHRF1 domains for Co-IP and found a weak interaction of the SRA domain with CBX1 (Fig. 26A and 26B). Most CBX1-interacting proteins contain a consensus peptide pentamer PxVx[M/L/V] that is also present in the SRA domain (Fig. 27A, Thiru et al., 2004; Huang et al., 2006). To test whether the interaction between CBX1 and UHRF1 is PxVxV-dependent, we mutated the conserved valine residues to aspartic and glutamic acid residues (Fig. 27B). Co-IP experiments showed the interaction was not affected by disruption of PxVx[M/L/V] motif, suggesting that the association between UHRF1 and CBX1 is independent of this motif (Fig. 27C).

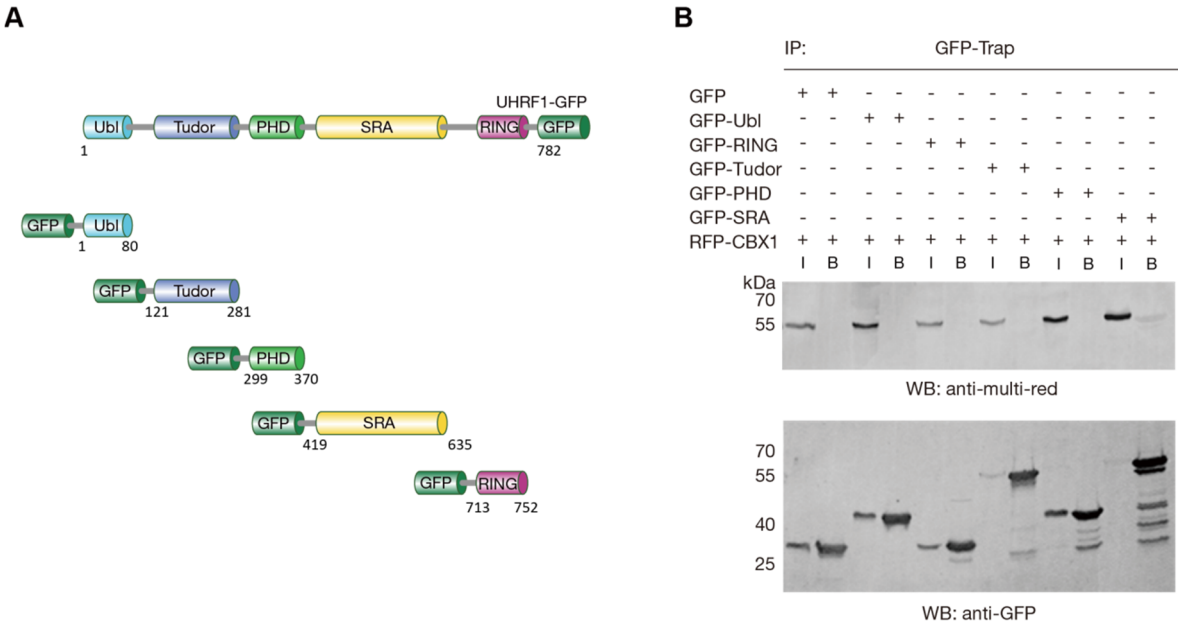

Figure 26. Mapping the domains of UHRF1 responsible for interaction with CBX1. (A) Schematic structure of UHRF1 and its domains fused to GFP used for co-immunoprecipitation. (B) Co-immunoprecipitation to map the interaction between UHRF1 and CBX1. UHRF1 and its domains were coexpressed with RFP-CBX1 in HEK293T cells. Immunoprecipitations performed with GFP-Trap were analyzed using indicated antibodies Fig. 26B was from Dr. Weihua Qin.

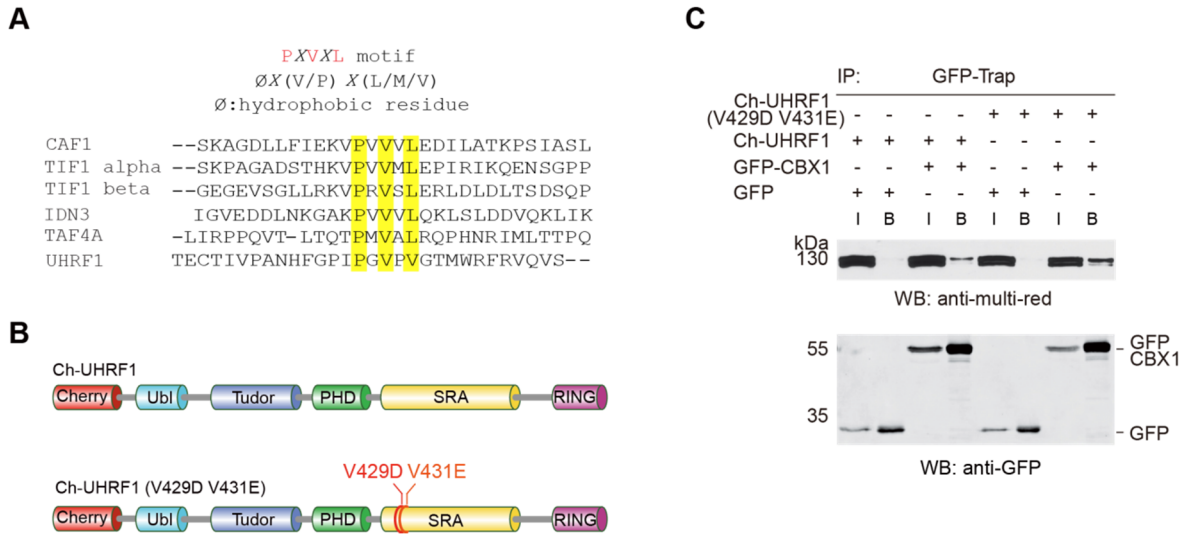

Figure 27. Association between UHRF1 and CBX1 is independent of PxVx[M/L/V] motif. (A) Alignment of HP1/CBX binding motif in different proteins (Thiru, Nietlispach et al. 2004, Huang, Myers et al. 2006). Yellow shaded amino acids are highly conserved. Accession numbers: CAF1 NP_038761; TIF1 alpha NP_056989; TIF1 beta NP_035718; IDN3 NP_056199; TAF4A NP_001074561; UHRF1 NP_001104548. (B) Schematic structure of UHRF1 and its substitution mutant in the PxVx[M/L/V] motif. (C) The interaction between CBX1 and UHRF1 is independent of the PxVx[M/L/V] motif. GFP-CBX1 was coexpressed in HEK293T cell with Ch-UHRF1 or Ch-UHRF1 mutant and immunoprecipitated with the GFP-Trap. Immunoprecipitations were subjected to western blot analysis and probed with indicated antibodies.

3.3.4 USP7 stabilizes CBX1 by deubiquitination

We and other groups previously showed that USP7 physically interacts with both DNMT1 and UHRF1 and controls their ubiquitination status (Du, Song et al. 2010, Qin, Leonhardt et al. 2011, Ma, Chen et al. 2012). Interestingly, Co-IP experiments revealed interactions of USP7 with all three of CBX proteins (Fig. 28A). To determine whether USP7 affects the ubiquitination level of CBX1, HEK293T cells were cotransfected with GFP-CBX1 in combination with HA-ubiquitin and Ch-USP7 expression constructs. Western blot analyses showed that GFP-CBX1 polyubiquitination was reduced to undetectable levels by coexpression of Ch-USP7, but not the catalytically inactive point mutant Ch-USP7^{C224S} (Fig. 28B).

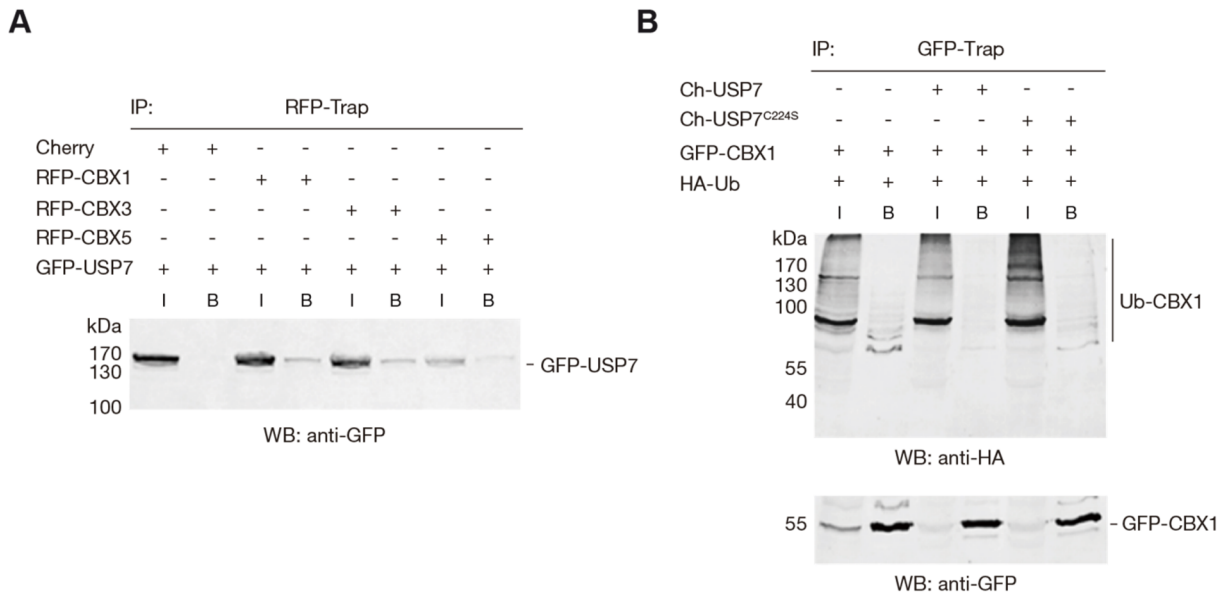

Figure 28. USP7 stabilizes CBX1 by deubiquitination. (A) USP7 interacts with all three CBX proteins. RFP-CBX proteins were immunoprecipitated with RFP-Trap from HEK293T cells coexpressing GFP-USP7 followed by western blot analysis with the indicated antibody. (B) Deubiquitination of CBX1 by USP7. GFP-Trap pulldowns from HEK293T cells expressing indicated combinations of HA-ubiquitin, GFP-CBX1, Ch-USP7 and Ch-USP7^{C224S} were analyzed by western blot analysis with an anti-HA antibody. These data were from Dr. Weihua Qin.

4 Discussion

4 Discussion

4.1 Regulation of TET enzymes via CXXC domains *in cis* and *in trans*

Since the first description of their enzymatic activity of oxidizing 5mC to 5hmC in 2009, TET family proteins have become a focus of substantial interest. In the past five years, TET enzymes have been extensively studied from several aspects including identification of interacting partners, crystal structural analyses, loss of function studies in ESCs and mouse, genome-wide profile determination of binding sites as well as the oxidized bases 5hmC, 5fC and 5caC. Growing evidence suggests that TET proteins play an important role in gene expression and cell differentiation, but how TET proteins regulate their target genes remains to be elucidated.

4.1.1 Regulation of TET proteins via CXXC domains *in cis*

TET proteins are relatively large size protein, but very few domains have so far been reported. The only well characterized module is a C-terminal catalytic domain, composed of a cysteine-rich and a DSBH domain, which oxidizes 5mC in a 2-OG and Fe^{2+}-dependent manner and is present in all three TET family members (Tahiliani, Koh et al. 2009). The DSBH domain has a central DSBH core which forms the catalytic cavity for binding the DNA substrate and catalyzing the base modification (Hu, Li et al. 2013). The cysteine-rich domain locating next to DSBH domain, though it does not form an independent structural unit, has an essential role for the enzymatic activity of TET proteins by stabilizing the DNA above the DSBH core (Iyer, Tahiliani et al. 2009, Tahiliani, Koh et al. 2009, Hu, Li et al. 2013).

While the C-terminal part of TET proteins is relatively well characterized, little is known about the N-terminal regulatory domain. The only domain identified is the N-terminal CXXC domain in TET1, in contrast, none of human and mouse TET2 and TET3 containing such domains. As CXXC domains generally serve as selective DNA binding domain, it raises the possibility that CXXC-mediated DNA binding might target TET1 to its DNA substrates. With *in vitro* DNA binding assays, we and others showed the slight binding preference of CXXC to unmethylated DNA (Xu, Wu et al. 2011, Liu, Wang et al. 2013). These results are also consistent with previous report that $CXXC^{TET1}$ has strong binding property to CpG-rich

DNA sequences and are consistent with genome-wide occupancy analyses of TET1 proteins preferentially localize to unmodified CpG-rich promoters and genic regions in mouse ESCs (Fig. 29A) (Williams, Christensen et al. 2011, Wu and Zhang 2011). Although the CXXC domain of TET1 shows binding affinity to unmethylated DNA, it is not required for oxidation activity *in vivo* (Tahiliani, Koh et al. 2009, Zhang, Zhang et al. 2010), suggesting that there might be other elements in the regulatory domain involved in targeting TET1 to its DNA substrates.

In a search of functional domains, none of the known DNA binding motifs was identified in the regulatory domain of annotated TET2 and TET3. Interestingly, we here showed a novel isoform of TET3 with a CXXC motif in the N-terminal domain ($TET3^{CXXC}$). Consistent with our study, $TET3^{CXXC}$ protein was also identified in *Xenopus* and human cell lines (Xu, Xu et al. 2012). Similar to other CXXC domains, TET3 CXXC motif is also able to bind unmethylated DNA in *in vitro* DNA binding assays (Fig. 29C). We showed that $TET3^{CXXC}$ reduces mobility in nuclei and shows higher conversion of 5mC to 5hmC in comparison to TET3 without CXXC zinc finger domain, suggesting that CXXC-mediated DNA binding might facilitate TET3 enzymatic activity which is consistent with a recent report that TET3 carrying a deletion of the CXXC domain resulted in a globally reduced 5hmC level in the genome of *Xenopus* embryos (Xu, Xu et al. 2012). With qPCR analyses, the transcript level of $TET3^{CXXC}$ significantly increased upon differentiation from ESCs to neural stem cells (NSCs), suggesting that $TET3^{CXXC}$ might be functionally important for neuronal differentiation. Indeed, several studies have reported a role of TET3 in mouse brain development. In embryonic cortex, an increase of TET2/$TET3^{CXXC}$ expression level is observed when neural progenitor cells differentiate into neurons, and knockdown of TET2/TET3 results in neuronal differentiation defects (Hahn, Qiu et al. 2013). TET3 knockout mice are viable through development and die on postnatal day one, but it is still unknown whether the perinatal lethality is linked to neurological development defects or not (Gu, Guo et al. 2011). A recent paper reported that TET3-mediated accumulation of 5hmC in adult cortex is associated with rapid behavioral adaptation (Li, Wei et al. 2014). Deletion of $TET3^{CXXC}$ in *Xenopus* causes defects in early eye and neural development, and importantly, both CXXC domain and catalytic domain are essential for its role in gene

regulation and embryonic development (Xu, Xu et al. 2012). According to these data, both mouse and *Xenopus* TET3^{CXXC} are likely to play a role in the development of the neuronal system. However, current functional studies on TET3 did not distinguish these two isoforms especially for knockdown or knockout assays, therefore cannot tell whether these effects are due to one or more specific isoforms. Further work focusing on specific gene deletion of different TET3 isoforms should provide insight into the functional similarity and differences of TET3^{CXXC} and TET3 during development.

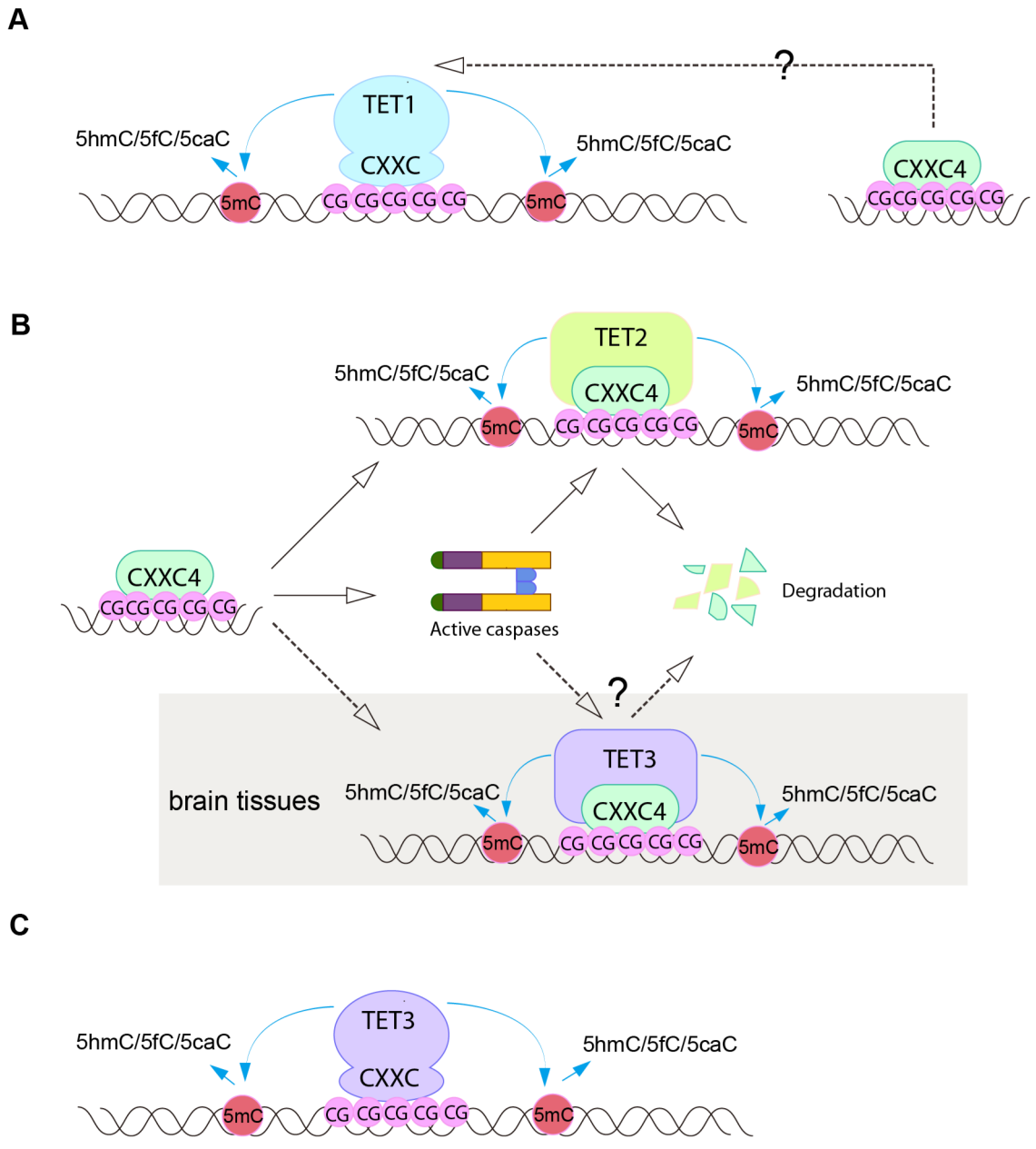

Figure 29. Schematic representation of CXXC domains in genomic targeting of TET proteins. (A) TET1 contains an N-terminal CXXC domain targeting TET1 protein to CpG-rich region where TET1 catalyzes 5mC oxidation. It is unclear whether CXXC4 is involved in regulating TET1 function. (B) CXXC4 might be involved in the genome targeting of both TET2 and TET3 without CXXC domain. CXXC4 preferentially binds chromatin regions containing unmethylated CpG dinucleotides, and then recruits TET2 to these specific DNA regions to oxidize 5mC. Concomitantly, DNA-bound CXXC4 activates caspases which cleave both TET2 and CXXC4 resulting in subsequent TET2 degradation (modified from Ko et al. 2013). Moreover, CXXC4 also interacts with TET3. Based on the high expression level of CXXC4 and TET3 in brain tissues we proposed that CXXC4 might constitute another targeting mechanism of TET3 in brain cells. But it is unknown whether CXXC4 also activates caspases to degrade TET3. (C) $TET3^{CXXC}$ can be targeted to CpG-rich regions by its CXXC domain where it exerts its enzymatic activity.

4.1.2 Regulation of TET proteins via CXXC domains *in trans*

In contrast to TET1 and $TET3^{CXXC}$, TET2 does not contain any known DNA binding motif, suggesting the existence of an alternative way to target TET2 to its DNA substrates. Supporting this idea, another research suggested that the CXXC domain of TET2 was separated from the catalytic domain during vertebrate evolution, and the separated region now encodes CXXC4 (Ko, An et al. 2013). CXXC4 physically interacts with and targets TET2 for destruction via caspase-dependent mechanisms (Fig. 29B), which depends on the DNA binding affinity of the CXXC4 CXXC domain (Ko, An et al. 2013). Interestingly, with genomic analyses we showed that *tet2* and *tet3* genes are adjacent to *cxxc4* and *cxxc10* respectively in human and mouse (Katoh and Katoh 2004, Frauer, Rottach et al. 2011), which suggests CXXC4 and CXXC10 might play a role in the recruitment of TET proteins to chromatin. Indeed, besides the association of TETs with a distinct group of CXXC domains *in cis*, we observed interactions between TET proteins and CXXC4, supporting the idea that CXXC4 might function as potential targeting partners. We observed that TET3 shows the strongest association with CXXC4 *in vivo* in comparison with TET1 and TET2 using a mammalian F3H assay, however, we failed to detect coimmunoprecipitation of any mouse TET proteins with CXXC4 fluorescent fusion constructs overexpressed in HEK293T cells, which may be due to the lack or limiting endogenous levels of bridging factors or posttranslational modifications in these cells. CXXC4 was shown to regulate TET2 enzymatic activity by controlling its abundance (Ko,

An et al. 2013), but it is unclear whether CXXC4 also plays a role in regulating TET3 stability *in vivo* (Fig. 31B). Furthermore, we noted that the ratio of TET3 to TET3^{CXXC} transcripts was much higher in brain regions relative to other tissues where CXXC4 transcripts were more abundant, which suggests that CXXC4 might contribute to context specific functions of TET3 in distinct cell types and developmental stages (Fig. 29B). Further investigations are needed to assess how CXXC4 regulates TET3 function *in vivo*. Taken together, our work and other recent studies proved that there are general connections between CXXC type zinc finger modules and TET proteins, either containing their own zinc finger domain (TET1 and TET3) or regulated by a closely related independent zinc finger protein (TET2 and TET3). To address the role of CXXC in recruitment of TET proteins, loss-of-function experiments have to be performed in the future.

Association with distinct CXXC domains may also modulate TET protein function by additional mechanisms. CXXC4 was shown to antagonize Wnt signaling by competing with Axin for binding to DVL, thus leading to destabilization of β-catenin (Hino, Kishida et al. 2001, Michiue, Fukui et al. 2004). β-catenin stabilization by DVL occurs in the cytoplasm, and nuclear DVL has been shown to interact with the C-JUN/TCF/β-catenin transcriptional complex and to be required for activation of Wnt signaling pathway (Fig. 30A). (Torres and Nelson 2000, Itoh, Brott et al. 2005, Gan, Wang et al. 2008). As CXXC4 is predominantly nuclear it would be interesting to investigate whether CXXC4 and TET proteins are involved in nuclear TCF/β-catenin transcriptional complexes and affect transcription of their target genes. A KTXXXI motif within the CXXC domain of CXXC4 was previously shown to be minimally required for the interaction with DVL (London, Lee et al. 2004), though poorly conserved in the CXXC domain of vertebrate TET3^{CXXC} isoforms, it is present in TET3 protein rather than TET1 and TET2. TET3 and CXXC4 might compete with the C-JUN to interact with DVL, thus inhibit β-catenin-dependent gene transcription (Fig. 30B). Differential expression of TET3 isoforms and interaction with CXXC4 may therefore modulate the recruitment of TET3 to TCF/β-catenin complexes.

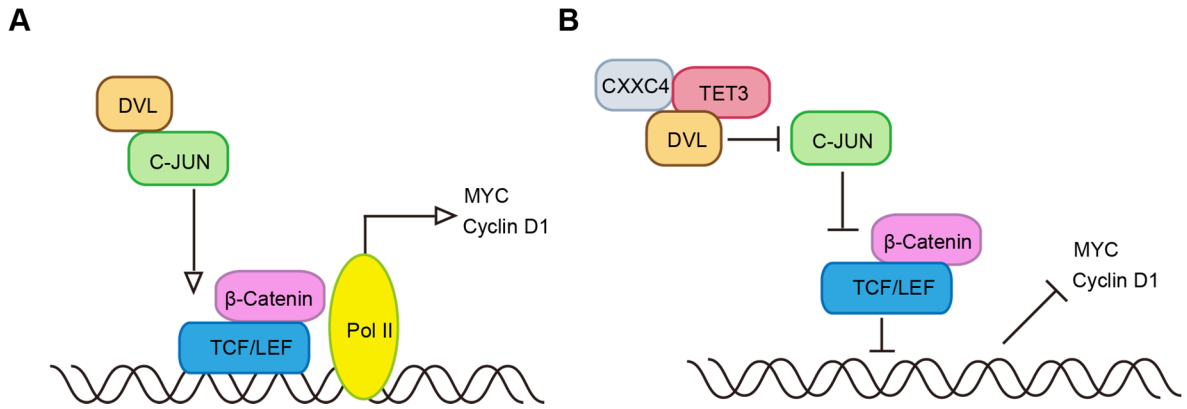

Figure 30. A model for involvement of CXXC4 and TET3 in the transcriptional complex in the canonical Wnt signaling pathway. (A) In the canonical Wnt signaling pathway, nuclear DVL interacts with C-JUN to form a complex with β-catenin/TCF/LEF on the promoter of Wnt target genes and regulates gene transcription such as MYC and Cyclin D1. (B) TET3 might interact with DVL directly or be recruited by CXXC4 to DVL to inhibit its association with C-JUN, thus inhibit the β-catenin/ TCF/LEF transcriptional activity.

In addition to zinc finger proteins, TET proteins have been found to interact with many differentiation-related factors such as NANOG, PPARγ and PRDM14 (Costa, Ding et al. 2013, Fujiki, Shinoda et al. 2013, Okashita, Kumaki et al. 2014), and chromatin-associated proteins involved in transcription activation including OGT and SET/COMPASS complex or transcription repression such as SIN3A and NURD complexes (Williams, Christensen et al. 2011, Chen, Chen et al. 2013, Deplus, Delatte et al. 2013, Shi, Kim et al. 2013, Vella, Scelfo et al. 2013, Zhang, Liu et al. 2014). Further research is necessary to untangle the intricate networks of TET proteins with the chromatin environment and fully understand the biological functions of TET proteins as well as their oxidized bases. These studies might also facilitate the development of therapeutic drugs for TETs related diseases.

4.2 Functions of UHRF1 in targeting DNMT1 to hemimethylated DNA substrates

Spatial and temporal control of DNA methylation in the genome is essential for normal development and is associated with a number of key cellular processes including genomic imprinting, X-chromosome inactivation, suppression of repetitive elements, and carcinogenesis. As maintenance DNA methyltransferase, DNMT1 is ubiquitously expressed in proliferation cells and shows an aberrant expression in certain tumors (Leonhardt, Page et al. 1992, Robert, Morin et al. 2003). Thus, to understand the mechanism of DNMT1 catalyzed maintenance DNA methylation has been of interest for researchers in the past two decades. The hemimethylated DNA is the prefered substrate of DNMT1 and is produced through semi-conservative DNA replication. Though DNMT1 colocalizes with DNA replication foci during S phase in living cells, it shows low preferential binding to hemimethylated DNA substrates *in vitro*, suggesting that additional cofactors must be required for recruitment of DNMT1 to hemimethylated CpG sites. The multi-domain protein UHRF1 was identified as an essential factor for the regulation of maintenance DNA methylation by targeting DNMT1 to its substrates. UHRF1 contains several chromatin binding domains, including a TTD and a PHD domain that bind H3K9me2/3, and a SRA domain that binds hemimethylated DNA (Arita, Ariyoshi et al. 2008, Avvakumov, Walker et al. 2008, Hashimoto, Horton et al. 2008, Qian, Li et al. 2008, Nady, Lemak et al. 2011, Xie, Jakoncic et al. 2012, Liu, Gao et al. 2013, Rothbart, Dickson et al. 2013). These chromatin binding domains of UHRF1 together with the Ubl and RING domain are essential for properly recruiting DNMT1 to chromatin for maintenance DNA methylation.

4.2.1 Direct targeting mechanism mediated by the SRA domain

In comparison to *de novo* DNA methyltransferases DNMT3A and 3B, DNMT1 contains a relatively large regulatory N-terminal domain that is essential for its methylation activity. Each N-terminal functional domain in DNMT1 was tested in order to determine the essential domain for recruiting DNMT1 to its DNA substrates. Although PBD domain

mediated PCNA association recruits DNMT1 to DNA replication foci, this interaction is not required for maintenance DNA methylation (Spada, Haemmer et al. 2007). The CXXC zinc finger shows binding preference for unmethylated DNA substrates, but is also not required for DNMT1 activity (Frauer, Rottach et al. 2011). However, based on crystal structure of DNMT1, DNMT1 CXXC domain seems to have an autoinhibition role in DNA methylation maintenance, in which the CXXC domain specifically binds to unmethylated CpG dinucleotides and occludes it from the active sites of DNMT1 to prevent *de novo* methylation activity (Song, Rechkoblit et al. 2011, Song, Teplova et al. 2012). In contrast to the PBD and CXXC domains, we found here that the TS domain is the only functional domain essential for directly targeting DNMT1 to hemimethyalted DNA substrates and is required for maintenance DNA methylation. The DNMT1 TS domain also shows binding ability to trimethylated H3K9, which is one possibility to explain its function in recruitment of DNMT1 to DNA substrates (data not shown; manuscript sumitted). Consistently, TS domain deletion mutants show a faster kinetics and a weaker association with chromocenters in S phase when transiently expressed in mouse fibroblasts, suggesting that the TS domain is involved in DNMT1 targeting (Schneider, Fuchs et al. 2013).. We propose that TS domain mediated chromatin association might be one prerequisite for DNMT1 recruitment and maintenance DNA methylation activity.

Similar to DNMT1, cofactor UHRF1 also shows significant binding preference for hemimethylated DNA substrates that is mediated by its SRA domain (Sharif, Muto et al. 2007, Arita, Ariyoshi et al. 2008, Avvakumov, Walker et al. 2008, Delagoutte, Lallous et al. 2008, Hashimoto, Horton et al. 2008, Qian, Li et al. 2008). It is possible that hemimethylated DNA binding by UHRF1 allows DNMT1 to specifically methylate the unmodified cytosine base on the opposite strand (Hashimoto, Horton et al. 2009, Hashimoto, Vertino et al. 2010). This model is quite tempting since hemimethylated CpG is the substrate of DNMT1, but it is unlikely that DNMT1 catalytic domain and UHRF1 SRA domain are able to bind simultaneously to the same site due to the steric clashes between these two proteins (Arita, Ariyoshi et al. 2008, Song, Teplova et al. 2012). Somehow the SRA domain of UHRF1 must be released from the reaction site to allow DNMT1 access to its substrate (Hashimoto, Vertino et al. 2010). This suggests that UHRF1

might first bind to hemimethylated sites which are then bound by DNMT1, after that UHRF1 is displaced and DNA methylation occurs (Hashimoto, Vertino et al. 2010). Consistent with previous publications, we also showed that the SRA domain mediates the association with DNMT1 and pointed out that the interaction between DNMT1 and UHRF1 is essential for maintenance DNA methylation (data not shown; manuscript sumitted). Thus, TS domain and SRA domain were characterized as essential elements for directly targeting DNMT1 to its substrates.

4.2.2 Indirect targeting mechanism coordinated by RING and PHD domains

The physical interaction between DNMT1 and UHRF1 is essential but not sufficient for maintenance of DNA methylation, suggesting that there might be an intermediate connecting UHRF1 and DNMT1 at replication sites (Song, Rechkoblit et al. 2011, Arita, Isogai et al. 2012, Nishiyama, Yamaguchi et al. 2013). A recent paper showed that UHRF1 mutants deficient in either SRA domain mediated hemimethylated CpG binding or TTD-PHD mediated H3K9me2/3 binding, were still able to target DNMT1 to replication sites for DNA methylation (Liu, Gao et al. 2013). Different from this finding, we and others found that point mutations in the PHD domain abolished the function of UHRF1 in maintenance DNA methylation (Rothbart, Dickson et al. 2013). These distinct results might be due to the different mutations in the PHD domain of UHRF1 tested. Besides the PHD domain, our data also show that the RING domain mutant still interacts with DNMT1 and does not affect the localization of UHRF1, but DNMT1 does not colocalize with these two UHRF1 mutants and shows a dispersed distribution in the nucleus (data not shown; manuscript sumitted). Also, these two mutants are unable to fully rescue the DNA methylation level in *uhrf1*$^{-/-}$ ESCs. Therefore, though RING and PHD domains do not affect interaction with DNMT1 directly, they are indispensable for targeting DNMT1 for maintenance DNA methylation.

Considering that the RING domain of UHRF1 has been reported to preferentially ubiquitinate histone H3 and DNMT1 (Citterio, Papait et al. 2004, Karagianni, Amazit et al. 2008, Du, Song et al. 2010, Qin, Leonhardt et al. 2011), we made use of mass

spectrometry to identify UHRF1-dependent ubiquitination targets and found that histone H3 is the substrate of UHRF1 in core histones. In addition, we further mapped the ubiquitination sites of histone H3 by UHRF1 with mass spectrometry and mutational analyses. Our data demonstrated that UHRF1 ubiquitinates histone H3 both on the K18 and K23 sites which is in consistence with a recent report showing that K23 residue of histone H3 is an ubiquitination target site of UHRF1 in *Xenopus* eggs (Nishiyama, Yamaguchi et al. 2013). Besides K23, we showed that K18 is the major ubiquitination target in mammalian cells. Considering that histone H3 is highly conserved throughout evolution, these different mapping results might result from the different materials used in the experiments. Ubiquitinated histone H3 was proposed to provide a docking site for DNMT1 binding and thereby maintain DNA methylation (Nishiyama, Yamaguchi et al. 2013). Remarkably, we found, to our knowledge for the first time, that DNMT1 contains a ubiquitin interacting motif (UIM) in the N-terminal part of the TS domain. This UIM is responsible for recruitment of DNMT1 to chromatin as UIM mutants showed a defect in the association with ubiquitinated histone H3 and H2A. Also UIM mutants of DNMT1 were unable to restore DNA methylation pattern in *dnmt1*$^{-/-}$ ESCs. Although the ubiquitination of H2A is one of the most abundant modifications in core histones, its functions in recruitment of DNMT1 to chromatin are still unknown. An interesting question to be answered is whether other essential factors are involved in targeting DNMT1 to chromatin in addition to UHRF1. We proposed that UHRF1 mediates ubiquitination of H3 at K18 and K23 residues providing the binding sites for the UIM of DNMT1, which is essential for propagation of DNA methylation after replication.

The PHD domain is one of the elements in UHRF1 for recruiting DNMT1 to its DNA substrates because of its binding affinity for unmethylated H3R2. Interestingly, we also observed that the ubiquitination level of histone H3 could not be rebuilt in *uhrf1*$^{-/-}$ ESCs stably expressing UHRF1 PHD mutant (data not shown; manuscript sumitted). This finding suggests that the PHD domain mediated chromatin association of UHRF1 might be a prerequisite for ubiquitination of histone H3. Thus, all functional domains in UHRF1 contribute to recruit DNMT1 to its substrates via either direct binding to chromatin or setting up docking sites for DNA methyltransferase. Although we do not have strong

supporting evidence, the functional domains of UHRF1 might coordinately work together to precisely regulate maintenance DNA methylation. To sum up, we proposed a direct and an indirect mechanism of DNMT1 targeting by UHRF1. In addition to SRA domain mediated recognition of hemimethylated DNA by UHRF1, TTD and PHD mediated histone binding serves as the first step in both mechanisms to ensure the proper recruitment of UHRF1 to appropriate genomic loci. Then the second step of the direct targeting mechanism is the TS domain of DNMT1 interacting with the SRA domain of UHRF1. In the second step of the indirect targeting mechanism, UHRF1 ubiquitinates histone H3 at K18 and K23 residues and thus provides a platform for the binding of DNMT1. In the third step, the UIM in the TS domain mediates binding of DNMT1 to ubiquitinated H3.

4.2.3 Chromatin association of DNMT1 is regulated by dynamic posttranslational modifications

As component of the DNMT1-UHRF1 complex, the ubiquitin specific protease USP7 was reported to function in stabilizing DNMT1 and UHRF1 via deubiquitination (Du, Song et al. 2010, Felle, Joppien et al. 2011, Qin, Leonhardt et al. 2011, Ma, Chen et al. 2012). This raises the possibility that USP7 might also control the ubiquitination status of histone H3 as an antagonist of UHRF1, thus dynamically regulating the association of DNMT1 with chromatin. To answer this question, the ubiquitination of histone H3 and DNA methylation levels should be checked in USP7 depleted cells. In addition to the potential function in regulation of H3 ubiquitination status, USP7 is able to enhance DNMT1 activity *in vitro* (Du, Song et al. 2010, Felle, Joppien et al. 2011, Qin, Leonhardt et al. 2011), which might be due to the interaction between the Ubl domain of USP7 and the TS domain of DNMT1 (Felle, Joppien et al. 2011, Qin, Leonhardt et al. 2011). Crystal structural analyses revealed that the TS domain likely acts as an autoinhibitory region for DNMT1 methyltransferase activity (Fig. 31A, 31B and 31C). Thus UIM-mediated binding to ubiquitinated H3 or Ubl domains might release the TS domain from the DNA binding pocket and thereby contribute to the activation of its methyltransferase activity. All these data indicate that USP7 likely plays a role in controlling chromatin structure. USP7 and

UHRF1 might work together to modulate the ubiquitination status of histone H3 as well as DNMT1 and dynamically regulate the stability and association of DNMT1 with chromatin (Fig. 31A).

In addition to ubiquitination, K18 and K23 residues on histone H3 were previously reported to be mostly modified by acetylation (Kurdistani, Tavazoie et al. 2004, Wang, Zang et al. 2008). Lysine acetylation is generally associated with gene activation by weakening electrostatic interactions between histone tails and DNA, thus providing a more open chromatin structure. H3K18ac has been reported mainly locating in the region surrounding transcription start sites and positively correlates with gene expression (Wang, Zang et al. 2008). It is very likely that acetylation at these sites prevents further ubiquitination of H3K18 through a direct competition, thereby prevents the binding and silencing effect of DNMT1 (Fig. 31A). Besides the direct competition between these two lysine modifications, more complicated and indirect processes might also connect these two signaling pathways. For example, treatment with HDAC inhibitors not only causes accumulation of acetylated histones, but also induces global and gene-specific DNA demethylation (Ou, Torrisani et al. 2007), which occurs even when DNA replication is blocked. More importantly, DNMT1 was also reported to be decreased both at mRNA and protein level after HDAC inhibitors treatment, suggesting that acetylation has different profound effects on maintenance DNA methylation (Arzenani, Zade et al. 2011). Taken together, UHRF1-mediated histone ubiquitination might work together with histone acetylation to regulate chromatin structure and thereby affect gene expression. It would be interesting to further investigate the site-specific and global acetylation level in cell lines stably expressing RING and PHD mutations of UHRF1.

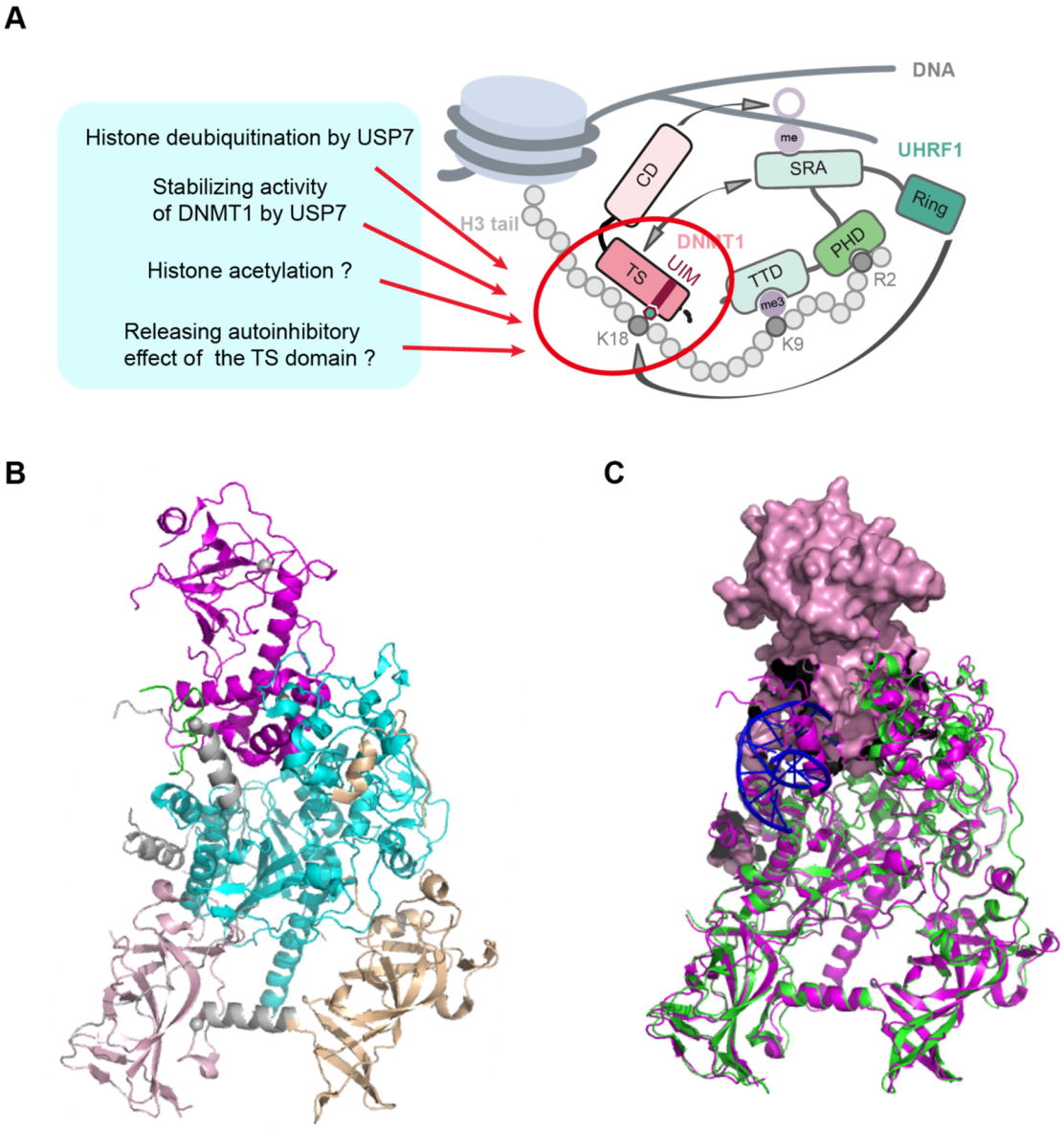

Figure 31. Potential regulatory factors in targeting DNMT to DNA substrates. (A) Histone acetylation on K18 and K23 residues and USP7-mediated deubiquitination might work together with UHRF1 to modulate the ubiquitination status of histone H3 and thereby dynamically regulate association of DNMT1 with chromatin. In addition, the binding to ubiquitinated H3 might lead to conformational change of TS domain and release the inhibitory effect on methyltransferase activity of DNMT1. (B) Cartoon model of mouse DNMT1 (amino acids 291-1620). The TS domain (magenta) plugs into the DNA binding pocket of the catalytic domain (cyan) (PDB: 3AV4). (C) Superposition of DNMT1 structure showing the sterical clash between the TS domain and DNA substrate. The structure from Takeshita et al. starts with the TS domain (pink; TS domain in surface representation) (Takeshita, Suetake et al. 2011), whereas the shorter structure from Song et al. starts with the CXXC domain (Song, Rechkoblit et al. 2011) and is solved in complex with unmethylated DNA (DNA in blue and remaining structure in green; PDB: 3PT6).

4.3 Potential roles of UHRF2 in epigenetics

UHRF2, the second member of the UHRF family proteins, is highly similar to UHRF1 both in sequence and structure. UHRF2 also interacts with DNMT1, DNMT3A/3B and G9A, and binds to H3K9me2/me3 through its TTD domain (Pichler, Wolf et al. 2011, Zhang, Gao et al. 2011). However, UHRF2 is unable to target DNMT1 to replication foci possibly due to a wider binding pocket of the SRA domain which shows optimal binding for fully hydroxymethylated DNA (Zhou, Xiong et al. 2014), thus providing a mechanistic explanation for UHRF2 inability to rescue DNA methylation defects in *uhrf1*$^{-/-}$ ESCs. Surprisingly, the recruitment of DNMT1 to DNA replication foci by UHRF1 still occurs with SRA domain mutant that lacks hemi-methylated CpG binding (Liu, Gao et al. 2013), which suggests that this is unlikely to be the only reason for UHRF2 incompetence for DNMT1 targeting. We here showed that RING domain mediated histone H3 ubiquitination is indispensable for recruitment of DNMT1 to appropriate genomic loci. Although the RING domain in UHRF family is highly similar, it is possible that E3 ligase activity on H3 might be also limited to UHRF1, which would be another explanation for the failure of UHRF2 to target DNMT1 to replication foci. The comparison of ubiquitin targets between UHRF1 and UHRF2 should shed new light on the distinct functions of UHRF2 in epigenetics.

Recently, UHRF2 was reported as a specific reader of 5hmC (Spruijt, Gnerlich et al. 2013, Zhou, Xiong et al. 2014), providing new insights into its biological functions in DNA demethylation. The direct binding of UHRF2 to 5hmC might protect it from further oxidization by TET enzymes (Fig. 32A). Compared with the 5hmC level, 5fC and 5caC are present in mammalian cells at much lower levels, only 20 5fC and 3 5caC are found in every 10^6 cytosines (He, Li et al. 2011, Ito, Shen et al. 2011, Pfaffeneder, Hackner et al. 2011). How TET enzymes select 5hmC for further processing is still unclear, but it is possible that most 5hmC are blocked by specific 5hmC readers and TET proteins can only work on the bare and accessible 5hmC bases.

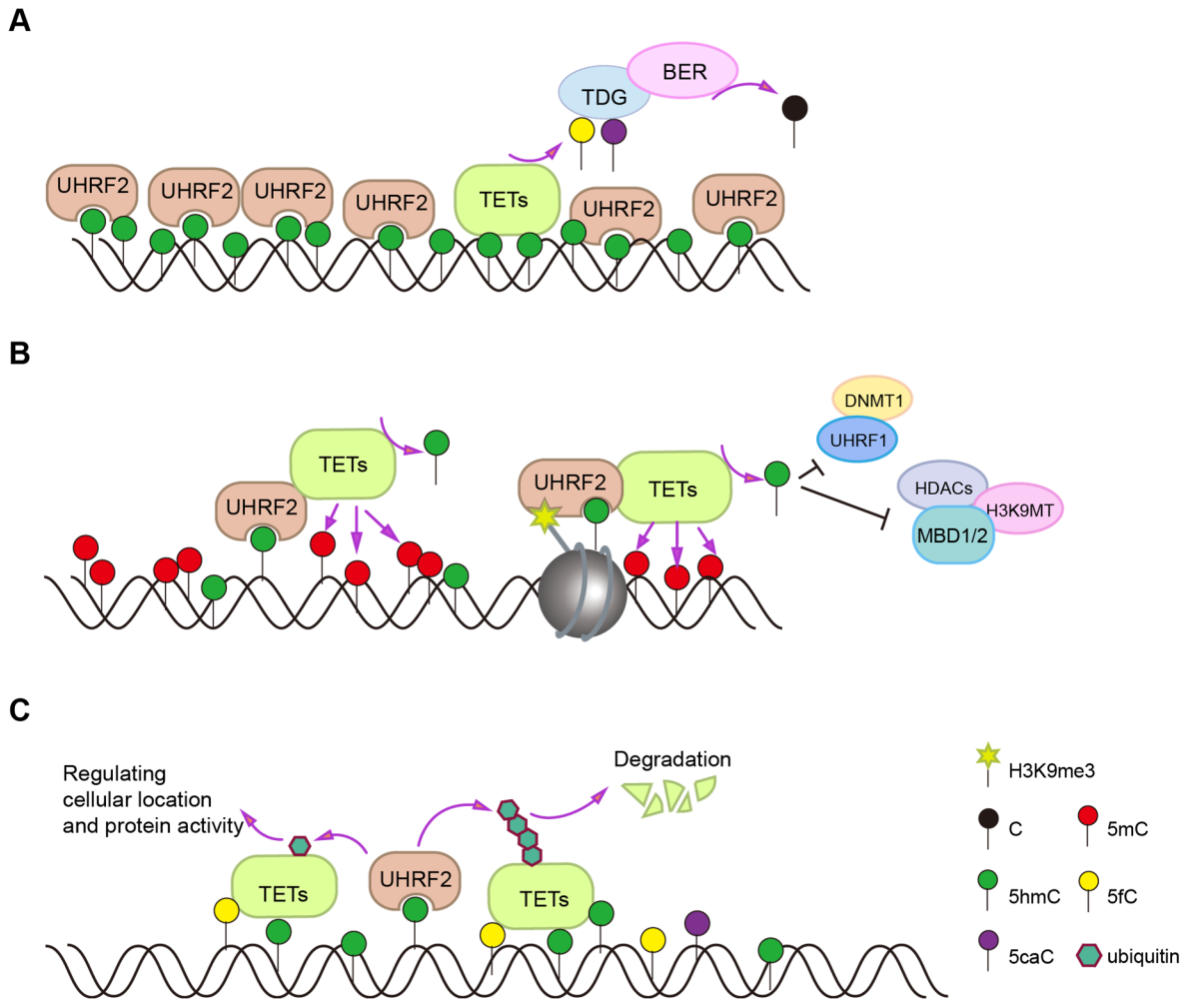

Figure 32. Diverse roles of UHRF2 in regulating TETs functions. (A) UHRF2 binding to 5hmC might prevent further oxidation. TETs can only oxidize unprotected 5hmC to 5fC and 5caC, which would be recognized and removed by TDG, and finally restore cytosine by BER. (B) SRA domain mediated UHRF2 binding to 5hmC might target TETs to oxidize adjacent 5mC, or TTD and PHD mediated H3K9me3 binding might work together with the SRA domain to recruit TETs to oxidize specific regions of chromatin. 5hmC strongly inhibits the binding of MBD1/2, DNMT1, UHRF1, HDACs and H3K9MT to chromatin, therefore is thought to activate gene expression. (C) UHRF2 harbors a RING domain which might ubiquitinate TETs leading to degradation (polyubiquitination) or further regulating their cellular location or enzymatic activity (monoubiquitination).

Since UHRF2 tightly associates with chromatin via its multiple domains, it might target TET proteins to chromatin to regulate gene expression (Fig. 32B). In support of this idea, it was shown that overexpression of UHRF2 together with TET1 catalytic domain increases the levels of 5hmC, 5fC and 5caC (Spruijt, Gnerlich et al. 2013), suggesting that UHRF2 might promote repetitive oxidation of 5mC. UHRF2 might recruit TET proteins to these

regions to oxidize 5mC, which will block the binding of UHRF1, DNMT1, HDACs, H3K9 methyltransferases (H3K9MT) and MBD proteins to chromatin, thus induce transcriptional activation.

In addition to function as a targeting partner, UHRF2 might also regulate the function of TETs as a protein modifier. TET proteins all have a large N-terminal regulatory domain which might be subjected to many posttranslational modifications. All three TET proteins can be O-GlcNAcylated by OGT which regulates TET protein function either by controlling protein abundance or subcellular localization (Chen, Chen et al. 2013, Shi, Kim et al. 2013, Vella, Scelfo et al. 2013, Zhang, Liu et al. 2014). UHRF2 contains a RING domain which has auto-ubiquitination activity (Mori, Li et al. 2004), and ubiquitinates different cyclins and PEST proteolytic signal-containing nuclear protein (PCNP), thus plays an essential role in cell cycle regulation (Mori, Li et al. 2004, Mori, Ikeda et al. 2011). It is tempting to propose that after setting up 5hmC pattern, TET proteins might also be ubiquitinated by UHRF2 which tags TETs for proteasomal degradation or altering their cellular location, activity or protein interactions (Fig. 32C). Interestingly, ubiquitin E3 ligase UHRF2 was reported as a SUMO E3 ligase (Oh and Chung 2013), so further studies should investigate whether TETs could be SUMOylated by UHRF2.

To verify these hypotheses, localization of UHRF2, TETs, TDG and different oxidized cytosine derivatives could be studied during the cell cycle to have first hints whether these proteins interact with each other or not, then global- and site-specific 5hmC changes in UHRF2 depleted cells could be investigated to determine whether UHRF2 negatively or positively affects enzymatic TET activity. Besides, ChIP assay is also needed to determine UHRF2 associated specific genomic regions which might indicate its biological functions.

4.4 Regulatory role of UHRF1 in chromatin dynamics

In addition to its role in controlling the maintenance of DNA methylation, UHRF1 is also involved in the control of large-scale reorganization and aggregation of chromocenters (Papait, Pistore et al. 2007, Papait, Pistore et al. 2008). Depletion of UHRF1 led to a reduced number and increased size of chromocenters, whereas ectopic expression of UHRF1 led to decondensation of chromocenters (Papait, Pistore et al. 2008), which suggests that UHRF1 contributes to dynamic changes of chromatin. To gain insights into the E3 ligase activity of UHRF1 in chromatin organization, we established a novel approach to detect ubiquitination *in vitro* and in living cells. Applying this method, we successfully identified UHRF1-dependent targets related to different regulatory pathways.

4.4.1 A novel approach to detect ubiquitination *in vitro* and in living cells

Ubiquitination is a key posttranslational modification of protein, which is essential for regulation of protein homeostasis and plays a central role in numerous processes like DNA damage repair, signal transduction pathways and innate immune responses (Haglund and Dikic 2005, Ravid and Hochstrasser 2008, Bianchi and Meier 2009). To better understand the biological significance of ubiquitin chains, it is essential to develop novel approach to detect the polymeric ubiquitin signal *in vitro* as well as in endogenous settings. Immunoprecipitation of target proteins followed by anti-ubiquitin immunoblotting are normally used (Peng, Schwartz et al. 2003, Choo and Zhang 2009, Udeshi, Mertins et al. 2013), and specific antibodies against target protein are needed in most of reported methods. We developed a novel approach to detect ubiquitination which circumvents the requirement of specific antibodies against proteins of interest. We fused two ubiquitin association domains (Uba) with GFP (GFP-2Uba) for detection and precipitation of ubiquitinated proteins with a nanobody based GFP-Trap, and the following detection could be done with anti-ubiquitin immunoblotting. This approach is highly efficient for precipitation of ubiquitinated proteins from crude cell extracts. Thus, with the bait GFP-Uba, E3 ligase dependent substrates and site-specific ubiquitination could be identified in combination with mass spectrometry.

Though techniques for detection of ubiquitinated protein have been relatively well developed, visualizing endogenous ubiquitination events in cells remains problematic. Recent technical advances in Uba based biosensor, linkage-specific polyubiquitin and ubiquitination-induced fluorescence complementation (UiFC) assay are able to monitor different types of ubiquitin chain in living cells (Sims, Scavone et al. 2012, van Wijk, Fiskin et al. 2012, Chen, Zhong et al. 2013), however, these methods focused on the ubiquitination at the whole cell level rather than specific ubiquitination of target protein. In our study, based on F3H assay (Herce, Deng et al. 2013), our approach is able to visualize specific ubiquitinated protein in cells. GFP fusion proteins are anchored at a *lac* operator (*lac*O) array inserted in the genome and visible as a spot of enriched green fluorescence in the nucleus. Cherry tagged 2xUba fusion protein will accumulate at the *lac*O spot if the immobilized GFP fusion proteins are ubiquitinated (Dr. Weihua Qin, personal communication). With combination, we broaden the application of F3H from only studying protein-protein interactions to functional screening in living cells. This method can also be applied to study ubiquitination dynamics using a combination with fluorescence recovery after photobleaching (FRAP) experiments in living cells.

4.4.2 Chromatin structure and heterochromatin protein HP1

With this novel approach, heterochromatin protein CBX1 and CBX3 were identified as UHRF1 ubiquitination targets, suggesting a role of UHRF1 in the regulation of chromatin structure as HP1 plays important roles in heterochromatin formation.

In mammals, there are three homologues of HP1, termed HP1α, HP1β and HP1γ, also known as CBX5, CBX1 and CBX3, respectively. Despite the high similarities in amino acid sequence and structural organization, HP1 homologues also display differences with respect to heterochromatin formation as well as gene regulation (Nielsen, Sanchez et al. 2002, Black, Allen et al. 2010, Black, Manning et al. 2013, Ryu, Lee et al. 2014). Several reports indicated that different PTMs might be associated with distinct functional properties of the HP1 family proteins. For example, phosphorylation of CBX3 was shown to correlate with mitosis and inhibits transcriptional repression (Minc, Allory et al. 1999,

Koike, Maita et al. 2000), while phosphorylation of threonine 51 of CBX1 led to a loss of H3K9me binding (Ayoub, Jeyasekharan et al. 2008). SUMOylation of CBX5 promotes initial targeting of CBX5 to pericentric heterochromatin (Maison, Bailly et al. 2011, Maison, Romeo et al. 2012). Our results show that in contrast to CBX3 and CBX5, only CBX1 is strongly ubiquitinated by UHRF1 *in vitro* and *in vivo*, raising the possibility that ubiquitination by UHRF1 may endow CBX1 distinct functions. Different from the functions of ubiquitination in signaling pathways, the ubiquitination of CBX1 by UHRF1 controls its abundance, suggesting that UHRF1 might function in heterochromatin formation. However, the question is still open whether the ubiquitination of CBX1 also functions as docking sites for protein interactions.

USP7 physically interacts with both DNMT1 and UHRF1 and controls their ubiquitination status (Du, Song et al. 2010, Felle, Joppien et al. 2011, Qin, Leonhardt et al. 2011, Ma, Chen et al. 2012). In this study, we showed that USP7 also associates with all three CBX proteins and controls CBX1 stability by deubiquitination. As UHRF1 is most abundant in S phase (Ma, Chen et al. 2012), accordingly CBX1 level also decreased in S phase (Dr. Weihua Qin, personal communication), suggesting that CBX1 might be subjected to cell cycle dependent ubiquitination regulation by UHRF1. Studies on composition changes of UHRF1-USP7 complexes during the cell cycle will help to better understand how the chromatin structure is dynamically regulated via controlling CBX1 abundance.

5 Annex

5.1 References

(2005). "Rapid amplification of 5' complementary DNA ends (5' RACE)." Nat Meth **2**(8): 629-630.
Aasland, R., T. J. Gibson and A. F. Stewart (1995). "The PHD finger: implications for chromatin-mediated transcriptional regulation." Trends Biochem Sci **20**(2): 56-59.
Allen, M. D., C. G. Grummitt, C. Hilcenko, S. Y. Min, L. M. Tonkin, C. M. Johnson, S. M. Freund, M. Bycroft and A. J. Warren (2006). "Solution structure of the nonmethyl-CpG-binding CXXC domain of the leukaemia-associated MLL histone methyltransferase." Embo j **25**(19): 4503-4512.
Allfrey, V. G., R. Faulkner and A. E. Mirsky (1964). "ACETYLATION AND METHYLATION OF HISTONES AND THEIR POSSIBLE ROLE IN THE REGULATION OF RNA SYNTHESIS." Proc Natl Acad Sci U S A **51**: 786-794.
Andersson, T., E. Sodersten, J. K. Duckworth, A. Cascante, N. Fritz, P. Sacchetti, I. Cervenka, V. Bryja and O. Hermanson (2009). "CXXC5 is a novel BMP4-regulated modulator of Wnt signaling in neural stem cells." J Biol Chem **284**(6): 3672-3681.
Arita, K., M. Ariyoshi, H. Tochio, Y. Nakamura and M. Shirakawa (2008). "Recognition of hemi-methylated DNA by the SRA protein UHRF1 by a base-flipping mechanism." Nature **455**(7214): 818-821.
Arita, K., S. Isogai, T. Oda, M. Unoki, K. Sugita, N. Sekiyama, K. Kuwata, R. Hamamoto, H. Tochio, M. Sato, M. Ariyoshi and M. Shirakawa (2012). "Recognition of modification status on a histone H3 tail by linked histone reader modules of the epigenetic regulator UHRF1." Proc Natl Acad Sci U S A **109**(32): 12950-12955.
Arzenani, M. K., A. E. Zade, Y. Ming, S. J. Vijverberg, Z. Zhang, Z. Khan, S. Sadique, L. Kallenbach, L. Hu, V. Vukojevic and T. J. Ekstrom (2011). "Genomic DNA hypomethylation by histone deacetylase inhibition implicates DNMT1 nuclear dynamics." Mol Cell Biol **31**(19): 4119-4128.
Avvakumov, G. V., J. R. Walker, S. Xue, Y. Li, S. Duan, C. Bronner, C. H. Arrowsmith and S. Dhe-Paganon (2008). "Structural basis for recognition of hemi-methylated DNA by the SRA domain of human UHRF1." Nature **455**(7214): 822-825.
Ayoub, N., A. D. Jeyasekharan, J. A. Bernal and A. R. Venkitaraman (2008). "HP1-beta mobilization promotes chromatin changes that initiate the DNA damage response." Nature **453**(7195): 682-686.
Bakker, J., X. Lin and W. G. Nelson (2002). "Methyl-CpG binding domain protein 2 represses transcription from hypermethylated pi-class glutathione S-transferase gene promoters in hepatocellular carcinoma cells." J Biol Chem **277**(25): 22573-22580.
Bateman, A., E. Birney, L. Cerruti, R. Durbin, L. Etwiller, S. R. Eddy, S. Griffiths-Jones, K. L. Howe, M. Marshall and E. L. Sonnhammer (2002). "The Pfam protein families database." Nucleic Acids Res **30**(1): 276-280.
Becker, P. B., S. Ruppert and G. Schutz (1987). "Genomic footprinting reveals cell type-specific DNA binding of ubiquitous factors." Cell **51**(3): 435-443.
Bednarik, D. P., C. Duckett, S. U. Kim, V. L. Perez, K. Griffis, P. C. Guenthner and T. M. Folks (1991). "DNA CpG methylation inhibits binding of NF-kappa B proteins to the HIV-1 long terminal repeat cognate DNA motifs." New Biol **3**(10): 969-976.
Beerli, R. R., U. Schopfer, B. Dreier and C. F. Barbas, 3rd (2000). "Chemically regulated zinc finger transcription factors." J Biol Chem **275**(42): 32617-32627.
Beumer, K. J., J. K. Trautman, M. Christian, T. J. Dahlem, C. M. Lake, R. S. Hawley, D. J. Grunwald, D. F. Voytas and D. Carroll (2013). "Comparing zinc finger nucleases and transcription activator-like effector nucleases for gene targeting in Drosophila." G3 (Bethesda) **3**(10): 1717-1725.

Bianchi, K. and P. Meier (2009). "A tangled web of ubiquitin chains: breaking news in TNF-R1 signaling." Mol Cell **36**(5): 736-742.
Bibikova, M. (2001). "Stimulation of homologous recombination through targeted cleavage by chimeric nucleases." Mol. Cell Biol. **21**: 289-297.
Bibikova, M., M. Golic, K. G. Golic and D. Carroll (2002). "Targeted chromosomal cleavage and mutagenesis in Drosophila using zinc-finger nucleases." Genetics **161**: 1169-1175.
Bird, A. (2002). "DNA methylation patterns and epigenetic memory." Genes Dev **16**(1): 6-21.
Birke, M., S. Schreiner, M. P. Garcia-Cuellar, K. Mahr, F. Titgemeyer and R. K. Slany (2002). "The MT domain of the proto-oncoprotein MLL binds to CpG-containing DNA and discriminates against methylation." Nucleic Acids Res **30**(4): 958-965.
Biswas, M., K. Voltz, J. C. Smith and J. Langowski (2011). "Role of histone tails in structural stability of the nucleosome." PLoS Comput Biol **7**(12): e1002279.
Black, J. C., A. Allen, C. Van Rechem, E. Forbes, M. Longworth, K. Tschop, C. Rinehart, J. Quiton, R. Walsh, A. Smallwood, N. J. Dyson and J. R. Whetstine (2010). "Conserved antagonism between JMJD2A/KDM4A and HP1gamma during cell cycle progression." Mol Cell **40**(5): 736-748.
Black, J. C., A. L. Manning, C. Van Rechem, J. Kim, B. Ladd, J. Cho, C. M. Pineda, N. Murphy, D. L. Daniels, C. Montagna, P. W. Lewis, K. Glass, C. D. Allis, N. J. Dyson, G. Getz and J. R. Whetstine (2013). "KDM4A lysine demethylase induces site-specific copy gain and rereplication of regions amplified in tumors." Cell **154**(3): 541-555.
Blackledge, N. P., J. C. Zhou, M. Y. Tolstorukov, A. M. Farcas, P. J. Park and R. J. Klose (2010). "CpG islands recruit a histone H3 lysine 36 demethylase." Mol Cell **38**(2): 179-190.
Blau, H. M., G. K. Pavlath, E. C. Hardeman, C. P. Chiu, L. Silberstein, S. G. Webster, S. C. Miller and C. Webster (1985). "Plasticity of the differentiated state." Science **230**(4727): 758-766.
Boch, J., H. Scholze, S. Schornack, A. Landgraf, S. Hahn, S. Kay, T. Lahaye, A. Nickstadt and U. Bonas (2009). "Breaking the code of DNA binding specificity of TAL-type III effectors." Science **326**(5959): 1509-1512.
Bohgaki, M., A. Hakem, M. J. Halaby, T. Bohgaki, Q. Li, P. A. Bissey, J. Shloush, T. Kislinger, O. Sanchez, Y. Sheng and R. Hakem (2013). "The E3 ligase PIRH2 polyubiquitylates CHK2 and regulates its turnover." Cell Death Differ **20**(6): 812-822.
Borden, K. L. (2000). "RING domains: master builders of molecular scaffolds?" J Mol Biol **295**(5): 1103-1112.
Borden, K. L., M. N. Boddy, J. Lally, N. J. O'Reilly, S. Martin, K. Howe, E. Solomon and P. S. Freemont (1995). "The solution structure of the RING finger domain from the acute promyelocytic leukaemia proto-oncoprotein PML." Embo j **14**(7): 1532-1541.
Borden, K. L. and P. S. Freemont (1996). "The RING finger domain: a recent example of a sequence-structure family." Curr Opin Struct Biol **6**(3): 395-401.
Bostick, M., J. K. Kim, P. O. Esteve, A. Clark, S. Pradhan and S. E. Jacobsen (2007). "UHRF1 plays a role in maintaining DNA methylation in mammalian cells." Science **317**(5845): 1760-1764.
Boyes, J. and A. Bird (1991). "DNA methylation inhibits transcription indirectly via a methyl-CpG binding protein." Cell **64**(6): 1123-1134.
Bransteitter, R., P. Pham, M. D. Scharff and M. F. Goodman (2003). "Activation-induced cytidine deaminase deaminates deoxycytidine on single-stranded DNA but requires the action of RNase." Proc Natl Acad Sci U S A **100**(7): 4102-4107.
Brayer, K. J. and D. J. Segal (2008). "Keep your fingers off my DNA: protein-protein interactions mediated by C2H2 zinc finger domains." Cell Biochem Biophys **50**(3): 111-131.

Cairns, B. R. (2009). "The logic of chromatin architecture and remodelling at promoters." Nature **461**(7261): 193-198.
Capuano, F., M. Mulleder, R. Kok, H. J. Blom and M. Ralser (2014). "Cytosine DNA Methylation Is Found in Drosophila melanogaster but Absent in Saccharomyces cerevisiae, Schizosaccharomyces pombe, and Other Yeast Species." Anal Chem **86**(8): 3697-3702.
Cedar, H. and Y. Bergman (2009). "Linking DNA methylation and histone modification: patterns and paradigms." Nat Rev Genet **10**(5): 295-304.
Chakravarty, S., L. Zeng and M. M. Zhou (2009). "Structure and site-specific recognition of histone H3 by the PHD finger of human autoimmune regulator." Structure **17**(5): 670-679.
Chedin, F., M. R. Lieber and C. L. Hsieh (2002). "The DNA methyltransferase-like protein DNMT3L stimulates de novo methylation by Dnmt3a." Proc Natl Acad Sci U S A **99**(26): 16916-16921.
Chen, C. C., K. Y. Wang and C. K. Shen (2012). "The mammalian de novo DNA methyltransferases DNMT3A and DNMT3B are also DNA 5-hydroxymethylcytosine dehydroxymethylases." J Biol Chem **287**(40): 33116-33121.
Chen, Q., Y. Chen, C. Bian, R. Fujiki and X. Yu (2013). "TET2 promotes histone O-GlcNAcylation during gene transcription." Nature **493**(7433): 561-564.
Chen, R. Z., U. Pettersson, C. Beard, L. Jackson-Grusby and R. Jaenisch (1998). "DNA hypomethylation leads to elevated mutation rates." Nature **395**(6697): 89-93.
Chen, T. and E. Li (2004). "Structure and function of eukaryotic DNA methyltransferases." Curr Top Dev Biol **60**: 55-89.
Chen, Z., Y. Zhong, Y. Wang, S. Xu, Z. Liu, I. V. Baskakov, M. J. Monteiro, M. Karbowski, Y. Shen and S. Fang (2013). "Ubiquitination-induced fluorescence complementation (UiFC) for detection of K48 ubiquitin chains in vitro and in live cells." PLoS One **8**(9): e73482.
Cheng, J., Y. Yang, J. Fang, J. Xiao, T. Zhu, F. Chen, P. Wang, Z. Li, H. Yang and Y. Xu (2013). "Structural insight into coordinated recognition of trimethylated histone H3 lysine 9 (H3K9me3) by the plant homeodomain (PHD) and tandem tudor domain (TTD) of UHRF1 (ubiquitin-like, containing PHD and RING finger domains, 1) protein." J Biol Chem **288**(2): 1329-1339.
Chignola, F., M. Gaetani, A. Rebane, T. Org, L. Mollica, C. Zucchelli, A. Spitaleri, V. Mannella, P. Peterson and G. Musco (2009). "The solution structure of the first PHD finger of autoimmune regulator in complex with non-modified histone H3 tail reveals the antagonistic role of H3R2 methylation." Nucleic Acids Res **37**(9): 2951-2961.
Choo, Y., I. Sanchez-Garcia and A. Klug (1994). "In vivo repression by a site-specific DNA-binding protein designed against an oncogenic sequence." Nature **372**(6507): 642-645.
Choo, Y. S. and Z. Zhang (2009). "Detection of protein ubiquitination." J Vis Exp(30).
Chuang, L. S., H. I. Ian, T. W. Koh, H. H. Ng, G. Xu and B. F. Li (1997). "Human DNA-(cytosine-5) methyltransferase-PCNA complex as a target for p21WAF1." Science **277**(5334): 1996-2000.
Cierpicki, T., L. E. Risner, J. Grembecka, S. M. Lukasik, R. Popovic, M. Omonkowska, D. D. Shultis, N. J. Zeleznik-Le and J. H. Bushweller (2010). "Structure of the MLL CXXC domain-DNA complex and its functional role in MLL-AF9 leukemia." Nat Struct Mol Biol **17**(1): 62-68.
Citterio, E., R. Papait, F. Nicassio, M. Vecchi, P. Gomiero, R. Mantovani, P. P. Di Fiore and I. M. Bonapace (2004). "Np95 is a histone-binding protein endowed with ubiquitin ligase activity." Mol Cell Biol **24**(6): 2526-2535.

Cliffe, L. J., R. Kieft, T. Southern, S. R. Birkeland, M. Marshall, K. Sweeney and R. Sabatini (2009). "JBP1 and JBP2 are two distinct thymidine hydroxylases involved in J biosynthesis in genomic DNA of African trypanosomes." Nucleic Acids Res **37**(5): 1452-1462.
Cokus, S. J., S. Feng, X. Zhang, Z. Chen, B. Merriman, C. D. Haudenschild, S. Pradhan, S. F. Nelson, M. Pellegrini and S. E. Jacobsen (2008). "Shotgun bisulphite sequencing of the Arabidopsis genome reveals DNA methylation patterning." Nature **452**(7184): 215-219.
Colot, V. and J. L. Rossignol (1999). "Eukaryotic DNA methylation as an evolutionary device." Bioessays **21**(5): 402-411.
Conti, L., S. M. Pollard, T. Gorba, E. Reitano, M. Toselli, G. Biella, Y. Sun, S. Sanzone, Q.-L. Ying, E. Cattaneo and A. Smith (2005). "Niche-Independent Symmetrical Self-Renewal of a Mammalian Tissue Stem Cell." PLoS Biol **3**(9): e283.
Corcoran, C. A., J. Montalbano, H. Sun, Q. He, Y. Huang and M. S. Sheikh (2009). "Identification and characterization of two novel isoforms of Pirh2 ubiquitin ligase that negatively regulate p53 independent of RING finger domains." J Biol Chem **284**(33): 21955-21970.
Cosgrove, M. S., J. D. Boeke and C. Wolberger (2004). "Regulated nucleosome mobility and the histone code." Nat Struct Mol Biol **11**(11): 1037-1043.
Costa, Y., J. Ding, T. W. Theunissen, F. Faiola, T. A. Hore, P. V. Shliaha, M. Fidalgo, A. Saunders, M. Lawrence, S. Dietmann, S. Das, D. N. Levasseur, Z. Li, M. Xu, W. Reik, J. C. Silva and J. Wang (2013). "NANOG-dependent function of TET1 and TET2 in establishment of pluripotency." Nature **495**(7441): 370-374.
Cross, S. H., R. R. Meehan, X. Nan and A. Bird (1997). "A component of the transcriptional repressor MeCP1 shares a motif with DNA methyltransferase and HRX proteins." Nat Genet **16**(3): 256-259.
Dambacher, S., W. Deng, M. Hahn, D. Sadic, J. Frohlich, A. Nuber, C. Hoischen, S. Diekmann, H. Leonhardt and G. Schotta (2012). "CENP-C facilitates the recruitment of M18BP1 to centromeric chromatin." Nucleus **3**(1): 101-110.
Daniel, J. M., C. M. Spring, H. C. Crawford, A. B. Reynolds and A. Baig (2002). "The p120(ctn)-binding partner Kaiso is a bi-modal DNA-binding protein that recognizes both a sequence-specific consensus and methylated CpG dinucleotides." Nucleic Acids Res **30**(13): 2911-2919.
Davey, C. A., D. F. Sargent, K. Luger, A. W. Maeder and T. J. Richmond (2002). "Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 a resolution." J Mol Biol **319**(5): 1097-1113.
Delagoutte, B., N. Lallous, C. Birck, P. Oudet and J. P. Samama (2008). "Expression, purification, crystallization and preliminary crystallographic study of the SRA domain of the human UHRF1 protein." Acta Crystallogr Sect F Struct Biol Cryst Commun **64**(Pt 10): 922-925.
Deplus, R., B. Delatte, M. K. Schwinn, M. Defrance, J. Mendez, N. Murphy, M. A. Dawson, M. Volkmar, P. Putmans, E. Calonne, A. H. Shih, R. L. Levine, O. Bernard, T. Mercher, E. Solary, M. Urh, D. L. Daniels and F. Fuks (2013). "TET2 and TET3 regulate GlcNAcylation and H3K4 methylation through OGT and SET1/COMPASS." Embo j **32**(5): 645-655.
Di Fiore, B., A. Palena, A. Felsani, F. Palitti, M. Caruso and P. Lavia (1999). "Cytosine methylation transforms an E2F site in the retinoblastoma gene promoter into a binding site for the general repressor methylcytosine-binding protein 2 (MeCP2)." Nucleic Acids Res **27**(14): 2852-2859.
Du, Z., J. Song, Y. Wang, Y. Zhao, K. Guda, S. Yang, H. Y. Kao, Y. Xu, J. Willis, S. D. Markowitz, D. Sedwick, R. M. Ewing and Z. Wang (2010). "DNMT1 stability is regulated by proteins coordinating deubiquitination and acetylation-driven ubiquitination." Sci Signal **3**(146): ra80.

DuBridge, R. B., P. Tang, H. C. Hsia, P. M. Leong, J. H. Miller and M. P. Calos (1987). "Analysis of mutation in human cells by using an Epstein-Barr virus shuttle system." Mol Cell Biol **7**(1): 379-387.
Dulac, C. (2010). "Brain function and chromatin plasticity." Nature **465**(7299): 728-735.
Dupont, C., D. R. Armant and C. A. Brenner (2009). "Epigenetics: definition, mechanisms and clinical perspective." Semin Reprod Med **27**(5): 351-357.
Ehrlich, M., M. A. Gama-Sosa, L. H. Huang, R. M. Midgett, K. C. Kuo, R. A. McCune and C. Gehrke (1982). "Amount and distribution of 5-methylcytosine in human DNA from different types of tissues of cells." Nucleic Acids Res **10**(8): 2709-2721.
Elrod-Erickson, M., M. A. Rould, L. Nekludova and C. O. Pabo (1996). "Zif268 protein-DNA complex refined at 1.6 A: a model system for understanding zinc finger-DNA interactions." Structure **4**(10): 1171-1180.
Eskat, A., W. Deng, A. Hofmeister, S. Rudolphi, S. Emmerth, D. Hellwig, T. Ulbricht, V. Doring, J. M. Bancroft, A. D. McAinsh, M. C. Cardoso, P. Meraldi, C. Hoischen, H. Leonhardt and S. Diekmann (2012). "Step-wise assembly, maturation and dynamic behavior of the human CENP-P/O/R/Q/U kinetochore sub-complex." PLoS One **7**(9): e44717.
Esteve, P. O., Y. Chang, M. Samaranayake, A. K. Upadhyay, J. R. Horton, G. R. Feehery, X. Cheng and S. Pradhan (2011). "A methylation and phosphorylation switch between an adjacent lysine and serine determines human DNMT1 stability." Nat Struct Mol Biol **18**(1): 42-48.
Feinberg, A. P., R. Ohlsson and S. Henikoff (2006). "The epigenetic progenitor origin of human cancer." Nat Rev Genet **7**(1): 21-33.
Felle, M., S. Joppien, A. Nemeth, S. Diermeier, V. Thalhammer, T. Dobner, E. Kremmer, R. Kappler and G. Langst (2011). "The USP7/Dnmt1 complex stimulates the DNA methylation activity of Dnmt1 and regulates the stability of UHRF1." Nucleic Acids Res **39**(19): 8355-8365.
Fellinger, K., U. Rothbauer, M. Felle, G. Langst and H. Leonhardt (2009). "Dimerization of DNA methyltransferase 1 is mediated by its regulatory domain." J Cell Biochem **106**(4): 521-528.
Feng, Q., R. Cao, L. Xia, H. Erdjument-Bromage, P. Tempst and Y. Zhang (2002). "Identification and functional characterization of the p66/p68 components of the MeCP1 complex." Mol Cell Biol **22**(2): 536-546.
Feng, Q. and Y. Zhang (2001). "The MeCP1 complex represses transcription through preferential binding, remodeling, and deacetylating methylated nucleosomes." Genes Dev **15**(7): 827-832.
Ferreira, H., J. Somers, R. Webster, A. Flaus and T. Owen-Hughes (2007). "Histone tails and the H3 alphaN helix regulate nucleosome mobility and stability." Mol Cell Biol **27**(11): 4037-4048.
Filion, G. J., S. Zhenilo, S. Salozhin, D. Yamada, E. Prokhortchouk and P. A. Defossez (2006). "A family of human zinc finger proteins that bind methylated DNA and repress transcription." Mol Cell Biol **26**(1): 169-181.
Finerty, P. J., Jr. and B. L. Bass (1997). "A Xenopus zinc finger protein that specifically binds dsRNA and RNA-DNA hybrids." J Mol Biol **271**(2): 195-208.
Frauer, C., T. Hoffmann, S. Bultmann, V. Casa, M. C. Cardoso, I. Antes and H. Leonhardt (2011). "Recognition of 5-hydroxymethylcytosine by the Uhrf1 SRA domain." PLoS One **6**(6): e21306.
Frauer, C. and H. Leonhardt (2009). "A versatile non-radioactive assay for DNA methyltransferase activity and DNA binding." Nucleic Acids Research **37**(3): e22.

Frauer, C., A. Rottach, D. Meilinger, S. Bultmann, K. Fellinger, S. Hasenoder, M. Wang, W. Qin, J. Soding, F. Spada and H. Leonhardt (2011). "Different binding properties and function of CXXC zinc finger domains in Dnmt1 and Tet1." PLoS One **6**(2): e16627.
Fujiki, K., A. Shinoda, F. Kano, R. Sato, K. Shirahige and M. Murata (2013). "PPARgamma-induced PARylation promotes local DNA demethylation by production of 5-hydroxymethylcytosine." Nat Commun **4**: 2262.
Fujita, N., S. Takebayashi, K. Okumura, S. Kudo, T. Chiba, H. Saya and M. Nakao (1999). "Methylation-mediated transcriptional silencing in euchromatin by methyl-CpG binding protein MBD1 isoforms." Mol Cell Biol **19**(9): 6415-6426.
Fujita, N., S. Watanabe, T. Ichimura, S. Tsuruzoe, Y. Shinkai, M. Tachibana, T. Chiba and M. Nakao (2003). "Methyl-CpG binding domain 1 (MBD1) interacts with the Suv39h1-HP1 heterochromatic complex for DNA methylation-based transcriptional repression." J Biol Chem **278**(26): 24132-24138.
Fuks, F., W. A. Burgers, A. Brehm, L. Hughes-Davies and T. Kouzarides (2000). "DNA methyltransferase Dnmt1 associates with histone deacetylase activity." Nat Genet **24**(1): 88-91.
Gan, X. Q., J. Y. Wang, Y. Xi, Z. L. Wu, Y. P. Li and L. Li (2008). "Nuclear Dvl, c-Jun, beta-catenin, and TCF form a complex leading to stabilization of beta-catenin-TCF interaction." J Cell Biol **180**(6): 1087-1100.
Gashler, A. L., S. Swaminathan and V. P. Sukhatme (1993). "A novel repression module, an extensive activation domain, and a bipartite nuclear localization signal defined in the immediate-early transcription factor Egr-1." Mol Cell Biol **13**(8): 4556-4571.
Goldberg, A. D., C. D. Allis and E. Bernstein (2007). "Epigenetics: a landscape takes shape." Cell **128**(4): 635-638.
Goll, M. G. and T. H. Bestor (2005). "Eukaryotic cytosine methyltransferases." Annu Rev Biochem **74**: 481-514.
Goll, M. G., F. Kirpekar, K. A. Maggert, J. A. Yoder, C. L. Hsieh, X. Zhang, K. G. Golic, S. E. Jacobsen and T. H. Bestor (2006). "Methylation of tRNAAsp by the DNA methyltransferase homolog Dnmt2." Science **311**(5759): 395-398.
Grishin, N. V. (2001). "Treble clef finger--a functionally diverse zinc-binding structural motif." Nucleic Acids Res **29**(8): 1703-1714.
Gu, T. P., F. Guo, H. Yang, H. P. Wu, G. F. Xu, W. Liu, Z. G. Xie, L. Shi, X. He, S. G. Jin, K. Iqbal, Y. G. Shi, Z. Deng, P. E. Szabo, G. P. Pfeifer, J. Li and G. L. Xu (2011). "The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes." Nature **477**(7366): 606-610.
Guo, J. U., Y. Su, C. Zhong, G. L. Ming and H. Song (2011). "Hydroxylation of 5-methylcytosine by TET1 promotes active DNA demethylation in the adult brain." Cell **145**(3): 423-434.
Haaf, T. (2006). "Methylation dynamics in the early mammalian embryo: implications of genome reprogramming defects for development." Curr Top Microbiol Immunol **310**: 13-22.
Haglund, K. and I. Dikic (2005). "Ubiquitylation and cell signaling." Embo j **24**(19): 3353-3359.
Hahn, M. A., R. Qiu, X. Wu, A. X. Li, H. Zhang, J. Wang, J. Jui, S. G. Jin, Y. Jiang, G. P. Pfeifer and Q. Lu (2013). "Dynamics of 5-hydroxymethylcytosine and chromatin marks in Mammalian neurogenesis." Cell Rep **3**(2): 291-300.
Hajkova, P., S. Erhardt, N. Lane, T. Haaf, O. El-Maarri, W. Reik, J. Walter and M. A. Surani (2002). "Epigenetic reprogramming in mouse primordial germ cells." Mech Dev **117**(1-2): 15-23.
Halaby, M. J., R. Hakem and A. Hakem (2013). "Pirh2: an E3 ligase with central roles in the regulation of cell cycle, DNA damage response, and differentiation." Cell Cycle **12**(17): 2733-2737.

Hashimoto, H., J. R. Horton, X. Zhang, M. Bostick, S. E. Jacobsen and X. Cheng (2008). "The SRA domain of UHRF1 flips 5-methylcytosine out of the DNA helix." Nature **455**(7214): 826-829.
Hashimoto, H., J. R. Horton, X. Zhang and X. Cheng (2009). "UHRF1, a modular multi-domain protein, regulates replication-coupled crosstalk between DNA methylation and histone modifications." Epigenetics **4**(1): 8-14.
Hashimoto, H., Y. Liu, A. K. Upadhyay, Y. Chang, S. B. Howerton, P. M. Vertino, X. Zhang and X. Cheng (2012). "Recognition and potential mechanisms for replication and erasure of cytosine hydroxymethylation." Nucleic Acids Res **40**(11): 4841-4849.
Hashimoto, H., P. M. Vertino and X. Cheng (2010). "Molecular coupling of DNA methylation and histone methylation." Epigenomics **2**(5): 657-669.
He, X. J., T. Chen and J. K. Zhu (2011). "Regulation and function of DNA methylation in plants and animals." Cell Res **21**(3): 442-465.
He, Y. F., B. Z. Li, Z. Li, P. Liu, Y. Wang, Q. Tang, J. Ding, Y. Jia, Z. Chen, L. Li, Y. Sun, X. Li, Q. Dai, C. X. Song, K. Zhang, C. He and G. L. Xu (2011). "Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA." Science **333**(6047): 1303-1307.
Heard, E. and C. M. Disteche (2006). "Dosage compensation in mammals: fine-tuning the expression of the X chromosome." Genes Dev **20**(14): 1848-1867.
Hendrich, B., J. Guy, B. Ramsahoye, V. A. Wilson and A. Bird (2001). "Closely related proteins MBD2 and MBD3 play distinctive but interacting roles in mouse development." Genes Dev **15**(6): 710-723.
Hendrich, B., U. Hardeland, H. H. Ng, J. Jiricny and A. Bird (1999). "The thymine glycosylase MBD4 can bind to the product of deamination at methylated CpG sites." Nature **401**(6750): 301-304.
Herce, H. D., W. Deng, J. Helma, H. Leonhardt and M. C. Cardoso (2013). "Visualization and targeted disruption of protein interactions in living cells." Nat Commun **4**: 2660.
Hino, S., S. Kishida, T. Michiue, A. Fukui, I. Sakamoto, S. Takada, M. Asashima and A. Kikuchi (2001). "Inhibition of the Wnt signaling pathway by Idax, a novel Dvl-binding protein." Mol Cell Biol **21**(1): 330-342.
Hirai, H., T. Tani and N. Kikyo (2010). "Structure and functions of powerful transactivators: VP16, MyoD and FoxA." Int J Dev Biol **54**(11-12): 1589-1596.
Holt, N., J. Wang, K. Kim, G. Friedman, X. Wang, V. Taupin, G. M. Crooks, D. B. Kohn, P. D. Gregory, M. C. Holmes and P. M. Cannon (2010). "Human hematopoietic stem/progenitor cells modified by zinc-finger nucleases targeted to CCR5 control HIV-1 in vivo." Nat Biotechnol **28**(8): 839-847.
Hu, L., Z. Li, J. Cheng, Q. Rao, W. Gong, M. Liu, Y. G. Shi, J. Zhu, P. Wang and Y. Xu (2013). "Crystal structure of TET2-DNA complex: insight into TET-mediated 5mC oxidation." Cell **155**(7): 1545-1555.
Hu, L., Z. Li, P. Wang, Y. Lin and Y. Xu (2011). "Crystal structure of PHD domain of UHRF1 and insights into recognition of unmodified histone H3 arginine residue 2." Cell Res **21**(9): 1374-1378.
Huang, Y., M. P. Myers and R. M. Xu (2006). "Crystal structure of the HP1-EMSY complex reveals an unusual mode of HP1 binding." Structure **14**(4): 703-712.
Ideraabdullah, F. Y., S. Vigneau and M. S. Bartolomei (2008). "Genomic imprinting mechanisms in mammals." Mutat Res **647**(1-2): 77-85.
Ito, S., A. C. D'Alessio, O. V. Taranova, K. Hong, L. C. Sowers and Y. Zhang (2010). "Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification." Nature **466**(7310): 1129-1133.

Ito, S., L. Shen, Q. Dai, S. C. Wu, L. B. Collins, J. A. Swenberg, C. He and Y. Zhang (2011). "Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine." Science **333**(6047): 1300-1303.
Itoh, K., B. K. Brott, G. U. Bae, M. J. Ratcliffe and S. Y. Sokol (2005). "Nuclear localization is required for Dishevelled function in Wnt/beta-catenin signaling." J Biol **4**(1): 3.
Iuchi, S. (2001). "Three classes of C2H2 zinc finger proteins." Cell Mol Life Sci **58**(4): 625-635.
Iwata, A., Y. Nagashima, L. Matsumoto, T. Suzuki, T. Yamanaka, H. Date, K. Deoka, N. Nukina and S. Tsuji (2009). "Intranuclear degradation of polyglutamine aggregates by the ubiquitin-proteasome system." J Biol Chem **284**(15): 9796-9803.
Iyer, L. M., M. Tahiliani, A. Rao and L. Aravind (2009). "Prediction of novel families of enzymes involved in oxidative and other complex modifications of bases in nucleic acids." Cell Cycle **8**(11): 1698-1710.
Jackson, D. A. and A. Pombo (1998). "Replicon clusters are stable units of chromosome structure: evidence that nuclear organization contributes to the efficient activation and propagation of S phase in human cells." J Cell Biol **140**(6): 1285-1295.
Jaenisch, R. and A. Bird (2003). "Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals." Nat Genet **33 Suppl**: 245-254.
Jakovcevski, M. and S. Akbarian (2012). "Epigenetic mechanisms in neurological disease." Nat Med **18**(8): 1194-1204.
Jans, D. A., C. Y. Xiao and M. H. Lam (2000). "Nuclear targeting signal recognition: a key control point in nuclear transport?" Bioessays **22**(6): 532-544.
Ji, Q., A. L. Fischer, C. R. Brown, E. R. Eastlund, T. Dvash, B. Zhong, M. A. Gerber, I. Lyons, S. W. Knight and C. A. Kreader (2014). "Engineered zinc-finger transcription factors activate OCT4 (POU5F1), SOX2, KLF4, c-MYC (MYC) and miR302/367." Nucleic Acids Res **42**(10): 6158-6167.
Jones, D. A., I. G. Cowell and P. B. Singh (2000). "Mammalian chromodomain proteins: their role in genome organisation and expression." BioEssays **22**: 124-127.
Jones, P. L., G. J. Veenstra, P. A. Wade, D. Vermaak, S. U. Kass, N. Landsberger, J. Strouboulis and A. P. Wolffe (1998). "Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription." Nat Genet **19**(2): 187-191.
Jorgensen, H. F., I. Ben-Porath and A. P. Bird (2004). "Mbd1 is recruited to both methylated and nonmethylated CpGs via distinct DNA binding domains." Mol Cell Biol **24**(8): 3387-3395.
Jung, Y. S., Y. Qian and X. Chen (2011). "The p73 tumor suppressor is targeted by Pirh2 RING finger E3 ubiquitin ligase for the proteasome-dependent degradation." J Biol Chem **286**(41): 35388-35395.
Jung, Y. S., Y. Qian and X. Chen (2012). "Pirh2 RING-finger E3 ubiquitin ligase: its role in tumorigenesis and cancer therapy." FEBS Lett **586**(10): 1397-1402.
Kalb, R., S. Latwiel, H. I. Baymaz, P. W. T. C. Jansen, C. W. Müller, M. Vermeulen and J. Müller (2014). "Histone H2A monoubiquitination promotes histone H3 methylation in Polycomb repression." Nat Struct Mol Biol **21**(6): 569-571.
Kaneda, M., M. Okano, K. Hata, T. Sado, N. Tsujimoto, E. Li and H. Sasaki (2004). "Essential role for de novo DNA methyltransferase Dnmt3a in paternal and maternal imprinting." Nature **429**(6994): 900-903.
Karagianni, P., L. Amazit, J. Qin and J. Wong (2008). "ICBP90, a novel methyl K9 H3 binding protein linking protein ubiquitination with heterochromatin formation." Mol Cell Biol **28**(2): 705-717.
Katoh, M. and M. Katoh (2004). "Identification and characterization of human CXXC10 gene in silico." Int J Oncol **25**(4): 1193-1199.

Kim, J., A. Kollhoff, A. Bergmann and L. Stubbs (2003). "Methylation-sensitive binding of transcription factor YY1 to an insulator sequence within the paternally expressed imprinted gene, Peg3." Hum Mol Genet **12**(3): 233-245.
Kim, J. K., P. O. Esteve, S. E. Jacobsen and S. Pradhan (2009). "UHRF1 binds G9a and participates in p21 transcriptional regulation in mammalian cells." Nucleic Acids Res **37**(2): 493-505.
Kim, S. S., Y. M. Chen, E. O'Leary, R. Witzgall, M. Vidal and J. V. Bonventre (1996). "A novel member of the RING finger family, KRIP-1, associates with the KRAB-A transcriptional repressor domain of zinc finger proteins." Proc Natl Acad Sci U S A **93**(26): 15299-15304.
Kim, Y. G., J. Cha and S. Chandrasegaran (1996). "Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain." Proc. Natl Acad. Sci. USA **93**: 1156-1160.
Klug, A. (2010). "The discovery of zinc fingers and their applications in gene regulation and genome manipulation." Annu Rev Biochem **79**: 213-231.
Ko, M., J. An, H. S. Bandukwala, L. Chavez, T. Aijo, W. A. Pastor, M. F. Segal, H. Li, K. P. Koh, H. Lahdesmaki, P. G. Hogan, L. Aravind and A. Rao (2013). "Modulation of TET2 expression and 5-methylcytosine oxidation by the CXXC domain protein IDAX." Nature **497**(7447): 122-126.
Kobayashi, N., T. G. Boyer and A. J. Berk (1995). "A class of activation domains interacts directly with TFIIA and stimulates TFIIA-TFIID-promoter complex assembly." Mol Cell Biol **15**(11): 6465-6473.
Koike, N., H. Maita, T. Taira, H. Ariga and S. M. Iguchi-Ariga (2000). "Identification of heterochromatin protein 1 (HP1) as a phosphorylation target by Pim-1 kinase and the effect of phosphorylation on the transcriptional repression function of HP1(1)." FEBS Lett **467**(1): 17-21.
Kouzarides, T. (2007). "Chromatin modifications and their function." Cell **128**(4): 693-705.
Krishna, S. S., I. Majumdar and N. V. Grishin (2003). "Structural classification of zinc fingers: survey and summary." Nucleic Acids Res **31**(2): 532-550.
Kundu, T. K., V. B. Palhan, Z. Wang, W. An, P. A. Cole and R. G. Roeder (2000). "Activator-dependent transcription from chromatin in vitro involving targeted histone acetylation by p300." Mol Cell **6**(3): 551-561.
Kurdistani, S. K., S. Tavazoie and M. Grunstein (2004). "Mapping global histone acetylation patterns to gene expression." Cell **117**(6): 721-733.
Kwan, A. H., D. A. Gell, A. Verger, M. Crossley, J. M. Matthews and J. P. Mackay (2003). "Engineering a protein scaffold from a PHD finger." Structure **11**(7): 803-813.
Lachner, M., D. O'Carroll, S. Rea, K. Mechtler and T. Jenuwein (2001). "Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins." Nature **410**(6824): 116-120.
Lange, M., B. Kaynak, U. B. Forster, M. Tonjes, J. J. Fischer, C. Grimm, J. Schlesinger, S. Just, I. Dunkel, T. Krueger, S. Mebus, H. Lehrach, R. Lurz, J. Gobom, W. Rottbauer, S. Abdelilah-Seyfried and S. Sperling (2008). "Regulation of muscle development by DPF3, a novel histone acetylation and methylation reader of the BAF chromatin remodeling complex." Genes Dev **22**(17): 2370-2384.
Latham, T., N. Gilbert and B. Ramsahoye (2008). "DNA methylation in mouse embryonic stem cells and development." Cell Tissue Res **331**(1): 31-55.
Law, J. A. and S. E. Jacobsen (2010). "Establishing, maintaining and modifying DNA methylation patterns in plants and animals." Nat Rev Genet **11**(3): 204-220.
Le Guezennec, X., M. Vermeulen, A. B. Brinkman, W. A. Hoeijmakers, A. Cohen, E. Lasonder and H. G. Stunnenberg (2006). "MBD2/NuRD and MBD3/NuRD, two distinct complexes with different biochemical and functional properties." Mol Cell Biol **26**(3): 843-851.

Lechner, M. S., G. E. Begg, D. W. Speicher and F. J. Rauscher, 3rd (2000). "Molecular determinants for targeting heterochromatin protein 1-mediated gene silencing: direct chromoshadow domain-KAP-1 corepressor interaction is essential." Mol Cell Biol **20**(17): 6449-6465.
Lee, J. H., K. S. Voo and D. G. Skalnik (2001). "Identification and characterization of the DNA binding domain of CpG-binding protein." J Biol Chem **276**(48): 44669-44676.
Lei, H., S. P. Oh, M. Okano, R. Juttermann, K. A. Goss, R. Jaenisch and E. Li (1996). "De novo DNA cytosine methyltransferase activities in mouse embryonic stem cells." Development **122**(10): 3195-3205.
Leng, R. P., Y. Lin, W. Ma, H. Wu, B. Lemmers, S. Chung, J. M. Parant, G. Lozano, R. Hakem and S. Benchimol (2003). "Pirh2, a p53-induced ubiquitin-protein ligase, promotes p53 degradation." Cell **112**(6): 779-791.
Leonhardt, H., A. W. Page, H. U. Weier and T. H. Bestor (1992). "A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei." Cell **71**(5): 865-873.
Lewis, J. D., R. R. Meehan, W. J. Henzel, I. Maurer-Fogy, P. Jeppesen, F. Klein and A. Bird (1992). "Purification, sequence, and cellular localization of a novel chromosomal protein that binds to methylated DNA." Cell **69**(6): 905-914.
Li, E. (2002). "Chromatin modification and epigenetic reprogramming in mammalian development." Nat Rev Genet **3**(9): 662-673.
Li, E., T. H. Bestor and R. Jaenisch (1992). "Targeted mutation of the DNA methyltransferase gene results in embryonic lethality." Cell **69**(6): 915-926.
Li, H., S. Ilin, W. Wang, E. M. Duncan, J. Wysocka, C. D. Allis and D. J. Patel (2006). "Molecular basis for site-specific read-out of histone H3K4me3 by the BPTF PHD finger of NURF." Nature **442**(7098): 91-95.
Li, X., M. Ito, F. Zhou, N. Youngson, X. Zuo, P. Leder and A. C. Ferguson-Smith (2008). "A maternal-zygotic effect gene, Zfp57, maintains both maternal and paternal imprints." Dev Cell **15**(4): 547-557.
Li, X., W. Wei, Q. Y. Zhao, J. Widagdo, D. Baker-Andresen, C. R. Flavell, A. D'Alessio, Y. Zhang and T. W. Bredy (2014). "Neocortical Tet3-mediated accumulation of 5-hydroxymethylcytosine promotes rapid behavioral adaptation." Proc Natl Acad Sci U S A **111**(19): 7120-7125.
Lin, Y. S., I. Ha, E. Maldonado, D. Reinberg and M. R. Green (1991). "Binding of general transcription factor TFIIB to an acidic activating region." Nature **353**(6344): 569-571.
Lister, R., M. Pelizzola, R. H. Dowen, R. D. Hawkins, G. Hon, J. Tonti-Filippini, J. R. Nery, L. Lee, Z. Ye, Q. M. Ngo, L. Edsall, J. Antosiewicz-Bourget, R. Stewart, V. Ruotti, A. H. Millar, J. A. Thomson, B. Ren and J. R. Ecker (2009). "Human DNA methylomes at base resolution show widespread epigenomic differences." Nature **462**(7271): 315-322.
Liu, N., M. Wang, W. Deng, C. S. Schmidt, W. Qin, H. Leonhardt and F. Spada (2013). "Intrinsic and extrinsic connections of Tet3 dioxygenase with CXXC zinc finger modules." PLoS One **8**(5): e62755.
Liu, P. Q., E. J. Rebar, L. Zhang, Q. Liu, A. C. Jamieson, Y. Liang, H. Qi, P. X. Li, B. Chen, M. C. Mendel, X. Zhong, Y. L. Lee, S. P. Eisenberg, S. K. Spratt, C. C. Case and A. P. Wolffe (2001). "Regulation of an endogenous locus using a panel of designed zinc finger proteins targeted to accessible chromatin regions. Activation of vascular endothelial growth factor A." J Biol Chem **276**(14): 11323-11334.
Liu, X., Q. Gao, P. Li, Q. Zhao, J. Zhang, J. Li, H. Koseki and J. Wong (2013). "UHRF1 targets DNMT1 for DNA methylation through cooperative binding of hemi-methylated DNA and methylated H3K9." Nat Commun **4**: 1563.

Liu, Y., S. Figley, S. K. Spratt, G. Lee, D. Ando, R. Surosky and M. G. Fehlings (2010). "An engineered transcription factor which activates VEGF-A enhances recovery after spinal cord injury." Neurobiol Dis **37**(2): 384-393.
Liu, Y., H. Toh, H. Sasaki, X. Zhang and X. Cheng (2012). "An atomic model of Zfp57 recognition of CpG methylation within a specific DNA sequence." Genes Dev **26**(21): 2374-2379.
Liutkeviciute, Z., G. Lukinavicius, V. Masevicius, D. Daujotyte and S. Klimasauskas (2009). "Cytosine-5-methyltransferases add aldehydes to DNA." Nat Chem Biol **5**(6): 400-402.
Livak, K. J. and T. D. Schmittgen (2001). "Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method." Methods **25**(4): 402-408.
Lloyd, A., C. L. Plaisier, D. Carroll and G. N. Drews (2005). "Targeted mutagenesis using zinc-finger nucleases in Arabidopsis." Proc. Natl Acad. Sci. USA **102**: 2232-2237.
Lo Conte, L., B. Ailey, T. J. Hubbard, S. E. Brenner, A. G. Murzin and C. Chothia (2000). "SCOP: a structural classification of proteins database." Nucleic Acids Res **28**(1): 257-259.
Loh, Y. H., W. Zhang, X. Chen, J. George and H. H. Ng (2007). "Jmjd1a and Jmjd2c histone H3 Lys 9 demethylases regulate self-renewal in embryonic stem cells." Genes Dev **21**(20): 2545-2557.
London, T. B., H. J. Lee, Y. Shao and J. Zheng (2004). "Interaction between the internal motif KTXXXI of Idax and mDvl PDZ domain." Biochem Biophys Res Commun **322**(1): 326-332.
Long, H. K., N. P. Blackledge and R. J. Klose (2013). "ZF-CxxC domain-containing proteins, CpG islands and the chromatin connection." Biochem Soc Trans **41**(3): 727-740.
Lorsbach, R. B., J. Moore, S. Mathew, S. C. Raimondi, S. T. Mukatira and J. R. Downing (2003). "TET1, a member of a novel protein family, is fused to MLL in acute myeloid leukemia containing the t(10;11)(q22;q23)." Leukemia **17**(3): 637-641.
Ma, H., H. Chen, X. Guo, Z. Wang, M. E. Sowa, L. Zheng, S. Hu, P. Zeng, R. Guo, J. Diao, F. Lan, J. W. Harper, Y. G. Shi, Y. Xu and Y. Shi (2012). "M phase phosphorylation of the epigenetic regulator UHRF1 regulates its physical association with the deubiquitylase USP7 and stability." Proc Natl Acad Sci U S A **109**(13): 4828-4833.
Maison, C., D. Bailly, D. Roche, R. M. de Oca, A. V. Probst, I. Vassias, F. Dingli, B. Lombard, D. Loew, J.-P. Quivy and G. Almouzni (2011). "SUMOylation promotes de novo targeting of HP1[alpha] to pericentric heterochromatin." Nat Genet **43**(3): 220-227.
Maison, C., K. Romeo, D. Bailly, M. Dubarry, J. P. Quivy and G. Almouzni (2012). "The SUMO protease SENP7 is a critical component to ensure HP1 enrichment at pericentric heterochromatin." Nat Struct Mol Biol **19**(4): 458-460.
Maiti, A. and A. C. Drohat (2011). "Thymine DNA glycosylase can rapidly excise 5-formylcytosine and 5-carboxylcytosine: potential implications for active demethylation of CpG sites." J Biol Chem **286**(41): 35334-35338.
Mansfield, R. E., C. A. Musselman, A. H. Kwan, S. S. Oliver, A. L. Garske, F. Davrazou, J. M. Denu, T. G. Kutateladze and J. P. Mackay (2011). "Plant homeodomain (PHD) fingers of CHD4 are histone H3-binding modules with preference for unmodified H3K4 and methylated H3K9." J Biol Chem **286**(13): 11779-11791.
Masumoto, H., D. Hawke, R. Kobayashi and A. Verreault (2005). "A role for cell-cycle-regulated histone H3 lysine 56 acetylation in the DNA damage response." Nature **436**(7048): 294-298.
Matthews, A. G., A. J. Kuo, S. Ramon-Maiques, S. Han, K. S. Champagne, D. Ivanov, M. Gallardo, D. Carney, P. Cheung, D. N. Ciccone, K. L. Walter, P. J. Utz, Y. Shi, T. G. Kutateladze, W. Yang, O. Gozani and M. A. Oettinger (2007). "RAG2 PHD finger couples histone H3 lysine 4 trimethylation with V(D)J recombination." Nature **450**(7172): 1106-1110.

Mayer, W., A. Niveleau, J. Walter, R. Fundele and T. Haaf (2000). "Demethylation of the zygotic paternal genome." Nature **403**(6769): 501-502.
Meilinger, D., K. Fellinger, S. Bultmann, U. Rothbauer, I. M. Bonapace, W. E. Klinkert, F. Spada and H. Leonhardt (2009). "Np95 interacts with de novo DNA methyltransferases, Dnmt3a and Dnmt3b, and mediates epigenetic silencing of the viral CMV promoter in embryonic stem cells." EMBO Rep **10**(11): 1259-1264.
Mellen, M., P. Ayata, S. Dewell, S. Kriaucionis and N. Heintz (2012). "MeCP2 binds to 5hmC enriched within active genes and accessible chromatin in the nervous system." Cell **151**(7): 1417-1430.
Michiue, T., A. Fukui, A. Yukita, K. Sakurai, H. Danno, A. Kikuchi and M. Asashima (2004). "Xldax, an inhibitor of the canonical Wnt pathway, is required for anterior neural structure formation in Xenopus." Dev Dyn **230**(1): 79-90.
Millar, C. B., J. Guy, O. J. Sansom, J. Selfridge, E. MacDougall, B. Hendrich, P. D. Keightley, S. M. Bishop, A. R. Clarke and A. Bird (2002). "Enhanced CpG mutability and tumorigenesis in MBD4-deficient mice." Science **297**(5580): 403-405.
Miller, J., A. D. McLachlan and A. Klug (1985). "Repetitive zinc-binding domains in the protein transcription factor IIIA from Xenopus oocytes." Embo j **4**(6): 1609-1614.
Miller, J. C. (2007). "An improved zinc-finger nuclease architecture for highly specific genome editing." Nature Biotech. **25**: 778-785.
Minc, E., Y. Allory, H. J. Worman, J. C. Courvalin and B. Buendia (1999). "Localization and phosphorylation of HP1 proteins during the cell cycle in mammalian cells." Chromosoma **108**(4): 220-234.
Mori, T., D. D. Ikeda, T. Fukushima, S. Takenoshita and H. Kochi (2011). "NIRF constitutes a nodal point in the cell cycle network and is a candidate tumor suppressor." Cell Cycle **10**(19): 3284-3299.
Mori, T., D. D. Ikeda, Y. Yamaguchi and M. Unoki (2012). "NIRF/UHRF2 occupies a central position in the cell cycle network and allows coupling with the epigenetic landscape." FEBS Lett **586**(11): 1570-1583.
Mori, T., Y. Li, H. Hata and H. Kochi (2004). "NIRF is a ubiquitin ligase that is capable of ubiquitinating PCNP, a PEST-containing nuclear protein." FEBS Lett **557**(1-3): 209-214.
Morton, J., M. W. Davis, E. M. Jorgensen and D. Carroll (2006). "Induction and repair of zinc-finger nuclease-targeted double-strand breaks in Caenorhabditis elegans somatic cells." Proc Natl Acad Sci U S A **103**(44): 16370-16375.
Moscou, M. J. and A. J. Bogdanove (2009). "A simple cipher governs DNA recognition by TAL effectors." Science **326**(5959): 1501.
Mueller, U., C. Bauer, M. Siegl, A. Rottach and H. Leonhardt (2014). "TET-mediated oxidation of methylcytosine causes TDG or NEIL glycosylase dependent gene reactivation." Nucleic Acids Res **42**(13): 8592-8604.
Murzin, A. G., S. E. Brenner, T. Hubbard and C. Chothia (1995). "SCOP: a structural classification of proteins database for the investigation of sequences and structures." J Mol Biol **247**(4): 536-540.
Musselman, C. A., R. E. Mansfield, A. L. Garske, F. Davrazou, A. H. Kwan, S. S. Oliver, H. O'Leary, J. M. Denu, J. P. Mackay and T. G. Kutateladze (2009). "Binding of the CHD4 PHD2 finger to histone H3 is modulated by covalent modifications." Biochem J **423**(2): 179-187.
Nabel, C. S., H. Jia, Y. Ye, L. Shen, H. L. Goldschmidt, J. T. Stivers, Y. Zhang and R. M. Kohli (2012). "AID/APOBEC deaminases disfavor modified cytosines implicated in DNA demethylation." Nat Chem Biol **8**(9): 751-758.
Nady, N., A. Lemak, J. R. Walker, G. V. Avvakumov, M. S. Kareta, M. Achour, S. Xue, S. Duan, A. Allali-Hassani, X. Zuo, Y. X. Wang, C. Bronner, F. Chedin, C. H. Arrowsmith and

S. Dhe-Paganon (2011). "Recognition of multivalent histone states associated with heterochromatin by UHRF1 protein." J Biol Chem **286**(27): 24300-24311.
Nakamura, Y., T. Umehara, H. Hamana, Y. Hayashizaki, M. Inoue, T. Kigawa, M. Shirouzu, T. Terada, A. Tanaka, B. Padmanabhan and S. Yokoyama (2007). "Crystal structure analysis of the PHD domain of the transcription co-activator Pygopus." J Mol Biol **370**(1): 80-92.
Nakao, M., S. Matsui, S. Yamamoto, K. Okumura, M. Shirakawa and N. Fujita (2001). "Regulation of transcription and chromatin by methyl-CpG binding protein MBD1." Brain Dev **23 Suppl 1**: S174-176.
Nakayama, J., J. C. Rice, B. D. Strahl, C. D. Allis and S. I. Grewal (2001). "Role of histone H3 lysine 9 methylation in epigenetic control of heterochromatin assembly." Science **292**(5514): 110-113.
Nan, X., R. R. Meehan and A. Bird (1993). "Dissection of the methyl-CpG binding domain from the chromosomal protein MeCP2." Nucleic Acids Res **21**(21): 4886-4892.
Nan, X., H. H. Ng, C. A. Johnson, C. D. Laherty, B. M. Turner, R. N. Eisenman and A. Bird (1998). "Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex." Nature **393**(6683): 386-389.
Narayan, V. A., R. W. Kriwacki and J. P. Caradonna (1997). "Structures of zinc finger domains from transcription factor Sp1. Insights into sequence-specific protein-DNA recognition." J Biol Chem **272**(12): 7801-7809.
Newell-Price, J., A. J. Clark and P. King (2000). "DNA methylation and silencing of gene expression." Trends Endocrinol Metab **11**(4): 142-148.
Nielsen, A. L., C. Sanchez, H. Ichinose, M. Cervino, T. Lerouge, P. Chambon and R. Losson (2002). "Selective interaction between the chromatin-remodeling factor BRG1 and the heterochromatin-associated protein HP1alpha." Embo j **21**(21): 5797-5806.
Nishiyama, A., L. Yamaguchi, J. Sharif, Y. Johmura, T. Kawamura, K. Nakanishi, S. Shimamura, K. Arita, T. Kodama, F. Ishikawa, H. Koseki and M. Nakanishi (2013). "Uhrf1-dependent H3K23 ubiquitylation couples maintenance DNA methylation and replication." Nature **502**(7470): 249-253.
Norton, V. G., B. S. Imai, P. Yau and E. M. Bradbury (1989). "Histone acetylation reduces nucleosome core particle linking number change." Cell **57**(3): 449-457.
Novembre, J., A. P. Galvani and M. Slatkin (2005). "The geographic spread of the CCR5 Delta32 HIV-resistance allele." PLoS Biol **3**(11): e339.
Oh, Y. and K. C. Chung (2013). "UHRF2, a ubiquitin E3 ligase, acts as a small ubiquitin-like modifier E3 ligase for zinc finger protein 131." J Biol Chem **288**(13): 9102-9111.
Ohki, I., N. Shimotake, N. Fujita, J. Jee, T. Ikegami, M. Nakao and M. Shirakawa (2001). "Solution structure of the methyl-CpG binding domain of human MBD1 in complex with methylated DNA." Cell **105**(4): 487-497.
Okano, M., D. W. Bell, D. A. Haber and E. Li (1999). "DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development." Cell **99**(3): 247-257.
Okashita, N., Y. Kumaki, K. Ebi, M. Nishi, Y. Okamoto, M. Nakayama, S. Hashimoto, T. Nakamura, K. Sugasawa, N. Kojima, T. Takada, M. Okano and Y. Seki (2014). "PRDM14 promotes active DNA demethylation through the ten-eleven translocation (TET)-mediated base excision repair pathway in embryonic stem cells." Development **141**(2): 269-280.
Okonechnikov, K., O. Golosova and M. Fursov (2012). "Unipro UGENE: a unified bioinformatics toolkit." Bioinformatics **28**(8): 1166-1167.
Ono, R., T. Taki, T. Taketani, M. Taniwaki, H. Kobayashi and Y. Hayashi (2002). "LCX, leukemia-associated protein with a CXXC domain, is fused to MLL in acute myeloid leukemia with trilineage dysplasia having t(10;11)(q22;q23)." Cancer Res **62**(14): 4075-4080.

Ooi, S. K., C. Qiu, E. Bernstein, K. Li, D. Jia, Z. Yang, H. Erdjument-Bromage, P. Tempst, S. P. Lin, C. D. Allis, X. Cheng and T. H. Bestor (2007). "DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA." Nature **448**(7154): 714-717.
Org, T., F. Chignola, C. Hetenyi, M. Gaetani, A. Rebane, I. Liiv, U. Maran, L. Mollica, M. J. Bottomley, G. Musco and P. Peterson (2008). "The autoimmune regulator PHD finger binds to non-methylated histone H3K4 to activate gene expression." EMBO Rep **9**(4): 370-376.
Oswald, J., S. Engemann, N. Lane, W. Mayer, A. Olek, R. Fundele, W. Dean, W. Reik and J. Walter (2000). "Active demethylation of the paternal genome in the mouse zygote." Curr Biol **10**(8): 475-478.
Ou, J. N., J. Torrisani, A. Unterberger, N. Provencal, K. Shikimi, M. Karimi, T. J. Ekstrom and M. Szyf (2007). "Histone deacetylase inhibitor Trichostatin A induces global and gene-specific DNA demethylation in human cancer cell lines." Biochem Pharmacol **73**(9): 1297-1307.
Ozdemir, A., S. Spicuglia, E. Lasonder, M. Vermeulen, C. Campsteijn, H. G. Stunnenberg and C. Logie (2005). "Characterization of lysine 56 of histone H3 as an acetylation site in Saccharomyces cerevisiae." J Biol Chem **280**(28): 25949-25952.
Papait, R., C. Pistore, U. Grazini, F. Babbio, S. Cogliati, D. Pecoraro, L. Brino, A. L. Morand, A. M. Dechampesme, F. Spada, H. Leonhardt, F. McBlane, P. Oudet and I. M. Bonapace (2008). "The PHD domain of Np95 (mUHRF1) is involved in large-scale reorganization of pericentromeric heterochromatin." Mol Biol Cell **19**(8): 3554-3563.
Papait, R., C. Pistore, D. Negri, D. Pecoraro, L. Cantarini and I. M. Bonapace (2007). "Np95 is implicated in pericentromeric heterochromatin replication and in major satellite silencing." Mol Biol Cell **18**(3): 1098-1106.
Papworth, M., M. Moore, M. Isalan, M. Minczuk, Y. Choo and A. Klug (2003). "Inhibition of herpes simplex virus 1 gene expression by designer zinc-finger transcription factors." Proc Natl Acad Sci U S A **100**(4): 1621-1626.
Parry, L. and A. R. Clarke (2011). "The Roles of the Methyl-CpG Binding Proteins in Cancer." Genes Cancer **2**(6): 618-630.
Pastor, W. A., L. Aravind and A. Rao (2013). "TETonic shift: biological roles of TET proteins in DNA demethylation and transcription." Nat Rev Mol Cell Biol **14**(6): 341-356.
Pavletich, N. P. and C. O. Pabo (1991). "Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 A." Science **252**(5007): 809-817.
Pawson, E. J., B. Duran-Jimenez, R. Surosky, H. E. Brooke, S. K. Spratt, D. R. Tomlinson and N. J. Gardiner (2010). "Engineered zinc finger protein-mediated VEGF-a activation restores deficient VEGF-a in sensory neurons in experimental diabetes." Diabetes **59**(2): 509-518.
Pendino, F., E. Nguyen, I. Jonassen, B. Dysvik, A. Azouz, M. Lanotte, E. Segal-Bendirdjian and J. R. Lillehaug (2009). "Functional involvement of RINF, retinoid-inducible nuclear factor (CXXC5), in normal and tumoral human myelopoiesis." Blood **113**(14): 3172-3181.
Peng, H., G. E. Begg, S. L. Harper, J. R. Friedman, D. W. Speicher and F. J. Rauscher, 3rd (2000). "Biochemical analysis of the Kruppel-associated box (KRAB) transcriptional repression domain." J Biol Chem **275**(24): 18000-18010.
Peng, H., G. E. Begg, D. C. Schultz, J. R. Friedman, D. E. Jensen, D. W. Speicher and F. J. Rauscher, 3rd (2000). "Reconstitution of the KRAB-KAP-1 repressor complex: a model system for defining the molecular anatomy of RING-B box-coiled-coil domain-mediated protein-protein interactions." J Mol Biol **295**(5): 1139-1162.
Peng, J., D. Schwartz, J. E. Elias, C. C. Thoreen, D. Cheng, G. Marsischky, J. Roelofs, D. Finley and S. P. Gygi (2003). "A proteomics approach to understanding protein ubiquitination." Nat Biotechnol **21**(8): 921-926.

Perez, E. E., J. Wang, J. C. Miller, Y. Jouvenot, K. A. Kim, O. Liu, N. Wang, G. Lee, V. V. Bartsevich, Y. L. Lee, D. Y. Guschin, I. Rupniewski, A. J. Waite, C. Carpenito, R. G. Carroll, J. S. Orange, F. D. Urnov, E. J. Rebar, D. Ando, P. D. Gregory, J. L. Riley, M. C. Holmes and C. H. June (2008). "Establishment of HIV-1 resistance in CD4+ T cells by genome editing using zinc-finger nucleases." Nat Biotechnol **26**(7): 808-816.

Peters, A. H., S. Kubicek, K. Mechtler, R. J. O'Sullivan, A. A. Derijck, L. Perez-Burgos, A. Kohlmaier, S. Opravil, M. Tachibana, Y. Shinkai, J. H. Martens and T. Jenuwein (2003). "Partitioning and plasticity of repressive histone methylation states in mammalian chromatin." Mol Cell **12**(6): 1577-1589.

Pfaffeneder, T., B. Hackner, M. Truss, M. Munzel, M. Muller, C. A. Deiml, C. Hagemeier and T. Carell (2011). "The discovery of 5-formylcytosine in embryonic stem cell DNA." Angew Chem Int Ed Engl **50**(31): 7008-7012.

Pichler, G., A. Jack, P. Wolf and S. B. Hake (2012). "Versatile Toolbox for High Throughput Biochemical and Functional Studies with Fluorescent Fusion Proteins." PLoS ONE **7**(5): e36967.

Pichler, G., P. Wolf, C. S. Schmidt, D. Meilinger, K. Schneider, C. Frauer, K. Fellinger, A. Rottach and H. Leonhardt (2011). "Cooperative DNA and histone binding by Uhrf2 links the two major repressive epigenetic pathways." J Cell Biochem **112**(9): 2585-2593.

Pradhan, M., P. O. Esteve, H. G. Chin, M. Samaranayke, G. D. Kim and S. Pradhan (2008). "CXXC domain of human DNMT1 is essential for enzymatic activity." Biochemistry **47**(38): 10000-10009.

Pradhan, S., A. Bacolla, R. D. Wells and R. J. Roberts (1999). "Recombinant human DNA (cytosine-5) methyltransferase. I. Expression, purification, and comparison of de novo and maintenance methylation." J Biol Chem **274**(46): 33002-33010.

Prendergast, G. C. and E. B. Ziff (1991). "Methylation-sensitive sequence-specific DNA binding by the c-Myc basic region." Science **251**(4990): 186-189.

Prokhortchouk, A., B. Hendrich, H. Jorgensen, A. Ruzov, M. Wilm, G. Georgiev, A. Bird and E. Prokhortchouk (2001). "The p120 catenin partner Kaiso is a DNA methylation-dependent transcriptional repressor." Genes Dev **15**(13): 1613-1618.

Qian, C., S. Li, J. Jakoncic, L. Zeng, M. J. Walsh and M. M. Zhou (2008). "Structure and hemimethylated CpG binding of the SRA domain from human UHRF1." J Biol Chem **283**(50): 34490-34494.

Qin, W., H. Leonhardt and G. Pichler (2011). "Regulation of DNA methyltransferase 1 by interactions and modifications." Nucleus **2**(5): 392-402.

Qin, W., H. Leonhardt and F. Spada (2011). "Usp7 and Uhrf1 control ubiquitination and stability of the maintenance DNA methyltransferase Dnmt1." J Cell Biochem **112**(2): 439-444.

Quenneville, S., G. Verde, A. Corsinotti, A. Kapopoulou, J. Jakobsson, S. Offner, I. Baglivo, P. V. Pedone, G. Grimaldi, A. Riccio and D. Trono (2011). "In embryonic stem cells, ZFP57/KAP1 recognize a methylated hexanucleotide to affect chromatin and DNA methylation of imprinting control regions." Mol Cell **44**(3): 361-372.

Rajakumara, E., Z. Wang, H. Ma, L. Hu, H. Chen, Y. Lin, R. Guo, F. Wu, H. Li, F. Lan, Y. G. Shi, Y. Xu, D. J. Patel and Y. Shi (2011). "PHD finger recognition of unmodified histone H3R2 links UHRF1 to regulation of euchromatic gene expression." Mol Cell **43**(2): 275-284.

Ramsahoye, B. H., D. Biniszkiewicz, F. Lyko, V. Clark, A. P. Bird and R. Jaenisch (2000). "Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a." Proc Natl Acad Sci U S A **97**(10): 5237-5242.

Rangam, G., K. M. Schmitz, A. J. Cobb and S. K. Petersen-Mahrt (2012). "AID enzymatic activity is inversely proportional to the size of cytosine C5 orbital cloud." PLoS One **7**(8): e43279.

Ravid, T. and M. Hochstrasser (2008). "Diversity of degradation signals in the ubiquitin-proteasome system." Nat Rev Mol Cell Biol **9**(9): 679-690.
Rebar, E. J., Y. Huang, R. Hickey, A. K. Nath, D. Meoli, S. Nath, B. Chen, L. Xu, Y. Liang, A. C. Jamieson, L. Zhang, S. K. Spratt, C. C. Case, A. Wolffe and F. J. Giordano (2002). "Induction of angiogenesis in a mouse model using engineered transcription factors." Nat Med **8**(12): 1427-1432.
Reik, W., W. Dean and J. Walter (2001). "Epigenetic reprogramming in mammalian development." Science **293**(5532): 1089-1093.
Reik, W. and J. Walter (2001). "Genomic imprinting: parental influence on the genome." Nat Rev Genet **2**(1): 21-32.
Reynolds, L., C. Ullman, M. Moore, M. Isalan, M. J. West, P. Clapham, A. Klug and Y. Choo (2003). "Repression of the HIV-1 5' LTR promoter and inhibition of HIV-1 replication by using engineered zinc-finger transcription factors." Proc Natl Acad Sci U S A **100**(4): 1615-1620.
Robert, M. F., S. Morin, N. Beaulieu, F. Gauthier, I. C. Chute, A. Barsalou and A. R. MacLeod (2003). "DNMT1 is required to maintain CpG methylation and aberrant gene silencing in human cancer cells." Nat Genet **33**(1): 61-65.
Robertson, K. D., S. Ait-Si-Ali, T. Yokochi, P. A. Wade, P. L. Jones and A. P. Wolffe (2000). "DNMT1 forms a complex with Rb, E2F1 and HDAC1 and represses transcription from E2F-responsive promoters." Nat Genet **25**(3): 338-342.
Rothbart, S. B., B. M. Dickson, M. S. Ong, K. Krajewski, S. Houliston, D. B. Kireev, C. H. Arrowsmith and B. D. Strahl (2013). "Multivalent histone engagement by the linked tandem Tudor and PHD domains of UHRF1 is required for the epigenetic inheritance of DNA methylation." Genes Dev **27**(11): 1288-1298.
Rothbauer, U., K. Zolghadr, S. Muyldermans, A. Schepers, M. C. Cardoso and H. Leonhardt (2008). "A versatile nanotrap for biochemical and functional studies with fluorescent fusion proteins." Mol Cell Proteomics **7**(2): 282-289.
Rothbauer, U., K. Zolghadr, S. Tillib, D. Nowak, L. Schermelleh, A. Gahl, N. Backmann, K. Conrath, S. Muyldermans, M. C. Cardoso and H. Leonhardt (2006). "Targeting and tracing antigens in live cells with fluorescent nanobodies." Nat Methods **3**(11): 887-889.
Rottach, A., C. Frauer, G. Pichler, I. M. Bonapace, F. Spada and H. Leonhardt (2010). "The multi-domain protein Np95 connects DNA methylation and histone modification." Nucleic Acids Res **38**(6): 1796-1804.
Rottach, A., E. Kremmer, D. Nowak, H. Leonhardt and M. C. Cardoso (2008). "Generation and characterization of a rat monoclonal antibody specific for multiple red fluorescent proteins." Hybridoma (Larchmt) **27**(5): 337-343.
Rountree, M. R., K. E. Bachman and S. B. Baylin (2000). "DNMT1 binds HDAC2 and a new co-repressor, DMAP1, to form a complex at replication foci." Nat Genet **25**(3): 269-277.
Ryu, H. W., D. H. Lee, L. Florens, S. K. Swanson, M. P. Washburn and S. H. Kwon (2014). "Analysis of the heterochromatin protein 1 (HP1) interactome in Drosophila." J Proteomics **102**: 137-147.
Saito, M. and F. Ishikawa (2002). "The mCpG-binding domain of human MBD3 does not bind to mCpG but interacts with NuRD/Mi2 components HDAC1 and MTA2." J Biol Chem **277**(38): 35434-35439.
Saitou, M., S. Kagiwada and K. Kurimoto (2012). "Epigenetic reprogramming in mouse pre-implantation development and primordial germ cells." Development **139**(1): 15-31.
Samson, M., F. Libert, B. J. Doranz, J. Rucker, C. Liesnard, C. M. Farber, S. Saragosti, C. Lapoumeroulie, J. Cognaux, C. Forceille, G. Muyldermans, C. Verhofstede, G. Burtonboy, M. Georges, T. Imai, S. Rana, Y. Yi, R. J. Smyth, R. G. Collman, R. W. Doms, G. Vassart

and M. Parmentier (1996). "Resistance to HIV-1 infection in caucasian individuals bearing mutant alleles of the CCR-5 chemokine receptor gene." Nature **382**(6593): 722-725.
Sanchez, R. and M. M. Zhou (2011). "The PHD finger: a versatile epigenome reader." Trends Biochem Sci **36**(7): 364-372.
Santoro, R., J. Li and I. Grummt (2002). "The nucleolar remodeling complex NoRC mediates heterochromatin formation and silencing of ribosomal gene transcription." Nat Genet **32**(3): 393-396.
Saurin, A. J., K. L. Borden, M. N. Boddy and P. S. Freemont (1996). "Does this have a familiar RING?" Trends Biochem Sci **21**(6): 208-214.
Schermelleh, L., A. Haemmer, F. Spada, N. Rosing, D. Meilinger, U. Rothbauer, M. C. Cardoso and H. Leonhardt (2007). "Dynamics of Dnmt1 interaction with the replication machinery and its role in postreplicative maintenance of DNA methylation." Nucleic Acids Res **35**(13): 4301-4312.
Schermelleh, L., F. Spada, H. P. Easwaran, K. Zolghadr, J. B. Margot, M. C. Cardoso and H. Leonhardt (2005). "Trapped in action: direct visualization of DNA methyltransferase activity in living cells." Nat Methods **2**(10): 751-756.
Schiesser, S., B. Hackner, T. Pfaffeneder, M. Muller, C. Hagemeier, M. Truss and T. Carell (2012). "Mechanism and stem-cell activity of 5-carboxycytosine decarboxylation determined by isotope tracing." Angew Chem Int Ed Engl **51**(26): 6516-6520.
Schindler, U., H. Beckmann and A. R. Cashmore (1993). "HAT3.1, a novel Arabidopsis homeodomain protein containing a conserved cysteine-rich region." Plant J **4**(1): 137-150.
Schneider, K., C. Fuchs, A. Dobay, A. Rottach, W. Qin, P. Wolf, J. M. Alvarez-Castro, M. M. Nalaskowski, E. Kremmer, V. Schmid, H. Leonhardt and L. Schermelleh (2013). "Dissection of cell cycle-dependent dynamics of Dnmt1 by FRAP and diffusion-coupled modeling." Nucleic Acids Res **41**(9): 4860-4876.
Schones, D. E., K. Cui, S. Cuddapah, T. Y. Roh, A. Barski, Z. Wang, G. Wei and K. Zhao (2008). "Dynamic regulation of nucleosome positioning in the human genome." Cell **132**(5): 887-898.
Schultz, D. C., K. Ayyanathan, D. Negorev, G. G. Maul and F. J. Rauscher, 3rd (2002). "SETDB1: a novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins." Genes Dev **16**(8): 919-932.
Schultz, D. C., J. R. Friedman and F. J. Rauscher, 3rd (2001). "Targeting histone deacetylase complexes via KRAB-zinc finger proteins: the PHD and bromodomains of KAP-1 form a cooperative unit that recruits a novel isoform of the Mi-2alpha subunit of NuRD." Genes Dev **15**(4): 428-443.
Senner, C. E. and N. Brockdorff (2009). "Xist gene regulation at the onset of X inactivation." Curr Opin Genet Dev **19**(2): 122-126.
Sharif, J., M. Muto, S. Takebayashi, I. Suetake, A. Iwamatsu, T. A. Endo, J. Shinga, Y. Mizutani-Koseki, T. Toyoda, K. Okamura, S. Tajima, K. Mitsuya, M. Okano and H. Koseki (2007). "The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA." Nature **450**(7171): 908-912.
Shen, L., H. Wu, D. Diep, S. Yamaguchi, A. C. D'Alessio, H. L. Fung, K. Zhang and Y. Zhang (2013). "Genome-wide analysis reveals TET- and TDG-dependent 5-methylcytosine oxidation dynamics." Cell **153**(3): 692-706.
Shi, F. T., H. Kim, W. Lu, Q. He, D. Liu, M. A. Goodell, M. Wan and Z. Songyang (2013). "Ten-eleven translocation 1 (Tet1) is regulated by O-linked N-acetylglucosamine transferase (Ogt) for target gene repression in mouse embryonic stem cells." J Biol Chem **288**(29): 20776-20784.
Shi, X., T. Hong, K. L. Walter, M. Ewalt, E. Michishita, T. Hung, D. Carney, P. Pena, F. Lan, M. R. Kaadige, N. Lacoste, C. Cayrou, F. Davrazou, A. Saha, B. R. Cairns, D. E.

Ayer, T. G. Kutateladze, Y. Shi, J. Cote, K. F. Chua and O. Gozani (2006). "ING2 PHD domain links histone H3 lysine 4 methylation to active gene repression." Nature **442**(7098): 96-99.
Sims, J. J., F. Scavone, E. M. Cooper, L. A. Kane, R. J. Youle, J. D. Boeke and R. E. Cohen (2012). "Polyubiquitin-sensor proteins reveal localization and linkage-type dependence of cellular ubiquitin signaling." Nat Methods **9**(3): 303-309.
Sjolund, A. B., A. G. Senejani and J. B. Sweasy (2013). "MBD4 and TDG: multifaceted DNA glycosylases with ever expanding biological roles." Mutat Res **743-744**: 12-25.
Song, C. X., K. E. Szulwach, Q. Dai, Y. Fu, S. Q. Mao, L. Lin, C. Street, Y. Li, M. Poidevin, H. Wu, J. Gao, P. Liu, L. Li, G. L. Xu, P. Jin and C. He (2013). "Genome-wide profiling of 5-formylcytosine reveals its roles in epigenetic priming." Cell **153**(3): 678-691.
Song, J., O. Rechkoblit, T. H. Bestor and D. J. Patel (2011). "Structure of DNMT1-DNA complex reveals a role for autoinhibition in maintenance DNA methylation." Science **331**(6020): 1036-1040.
Song, J., M. Teplova, S. Ishibe-Murakami and D. J. Patel (2012). "Structure-based mechanistic insights into DNMT1-mediated maintenance DNA methylation." Science **335**(6069): 709-712.
Spada, F., A. Haemmer, D. Kuch, U. Rothbauer, L. Schermelleh, E. Kremmer, T. Carell, G. Langst and H. Leonhardt (2007). "DNMT1 but not its interaction with the replication machinery is required for maintenance of DNA methylation in human cells." J Cell Biol **176**(5): 565-571.
Spada, F., U. Rothbauer, K. Zolghadr, L. Schermelleh and H. Leonhardt (2006). "Regulation of DNA methyltransferase 1." Adv Enzyme Regul **46**: 224-234.
Spruijt, C. G., F. Gnerlich, A. H. Smits, T. Pfaffeneder, P. W. Jansen, C. Bauer, M. Munzel, M. Wagner, M. Muller, F. Khan, H. C. Eberl, A. Mensinga, A. B. Brinkman, K. Lephikov, U. Muller, J. Walter, R. Boelens, H. van Ingen, H. Leonhardt, T. Carell and M. Vermeulen (2013). "Dynamic readers for 5-(hydroxy)methylcytosine and its oxidized derivatives." Cell **152**(5): 1146-1159.
Sripathy, S. P., J. Stevens and D. C. Schultz (2006). "The KAP1 corepressor functions to coordinate the assembly of de novo HP1-demarcated microenvironments of heterochromatin required for KRAB zinc finger protein-mediated transcriptional repression." Mol Cell Biol **26**(22): 8623-8638.
Stivers, J. T. and Y. L. Jiang (2003). "A mechanistic perspective on the chemistry of DNA repair glycosylases." Chem Rev **103**(7): 2729-2759.
Strohner, R., A. Nemeth, P. Jansa, U. Hofmann-Rohrer, R. Santoro, G. Langst and I. Grummt (2001). "NoRC--a novel member of mammalian ISWI-containing chromatin remodeling machines." Embo j **20**(17): 4892-4900.
Strohner, R., A. Nemeth, K. P. Nightingale, I. Grummt, P. B. Becker and G. Langst (2004). "Recruitment of the nucleolar remodeling complex NoRC establishes ribosomal DNA silencing in chromatin." Mol Cell Biol **24**(4): 1791-1798.
Suetake, I., F. Shinozaki, J. Miyagawa, H. Takeshima and S. Tajima (2004). "DNMT3L stimulates the DNA methylation activity of Dnmt3a and Dnmt3b through a direct interaction." J Biol Chem **279**(26): 27816-27823.
Sun, L., A. Liu and K. Georgopoulos (1996). "Zinc finger-mediated protein interactions modulate Ikaros activity, a molecular control of lymphocyte development." Embo j **15**(19): 5358-5369.
Suzuki, M. M. and A. Bird (2008). "DNA methylation landscapes: provocative insights from epigenomics." Nat Rev Genet **9**(6): 465-476.
Syeda, F., R. L. Fagan, M. Wean, G. V. Avvakumov, J. R. Walker, S. Xue, S. Dhe-Paganon and C. Brenner (2011). "The replication focus targeting sequence (RFTS) domain is a DNA-competitive inhibitor of Dnmt1." J Biol Chem **286**(17): 15344-15351.

Szczepek, M. (2007). "Structure-based redesign of the dimerization interface reduces the toxicity of zinc-finger nucleases." Nature Biotech. **25**: 786-793.
Szwagierczak, A., S. Bultmann, C. S. Schmidt, F. Spada and H. Leonhardt (2010). "Sensitive enzymatic quantification of 5-hydroxymethylcytosine in genomic DNA." Nucleic Acids Res **38**(19): e181.
Tachibana, M., K. Sugimoto, M. Nozaki, J. Ueda, T. Ohta, M. Ohki, M. Fukuda, N. Takeda, H. Niida, H. Kato and Y. Shinkai (2002). "G9a histone methyltransferase plays a dominant role in euchromatic histone H3 lysine 9 methylation and is essential for early embryogenesis." Genes Dev **16**(14): 1779-1791.
Tahiliani, M., K. P. Koh, Y. Shen, W. A. Pastor, H. Bandukwala, Y. Brudno, S. Agarwal, L. M. Iyer, D. R. Liu, L. Aravind and A. Rao (2009). "Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1." Science **324**(5929): 930-935.
Takeshita, K., I. Suetake, E. Yamashita, M. Suga, H. Narita, A. Nakagawa and S. Tajima (2011). "Structural insight into maintenance methylation by mouse DNA methyltransferase 1 (Dnmt1)." Proc Natl Acad Sci U S A **108**(22): 9055-9059.
Taverna, S. D., S. Ilin, R. S. Rogers, J. C. Tanny, H. Lavender, H. Li, L. Baker, J. Boyle, L. P. Blair, B. T. Chait, D. J. Patel, J. D. Aitchison, A. J. Tackett and C. D. Allis (2006). "Yng1 PHD finger binding to H3 trimethylated at K4 promotes NuA3 HAT activity at K14 of H3 and transcription at a subset of targeted ORFs." Mol Cell **24**(5): 785-796.
Thiru, A., D. Nietlispach, H. R. Mott, M. Okuwaki, D. Lyon, P. R. Nielsen, M. Hirshberg, A. Verreault, N. V. Murzina and E. D. Laue (2004). "Structural basis of HP1/PXVXL motif peptide interactions and HP1 localisation to heterochromatin." Embo j **23**(3): 489-499.
Torres, M. A. and W. J. Nelson (2000). "Colocalization and redistribution of dishevelled and actin during Wnt-induced mesenchymal morphogenesis." J Cell Biol **149**(7): 1433-1442.
Tsai, W. W., Z. Wang, T. T. Yiu, K. C. Akdemir, W. Xia, S. Winter, C. Y. Tsai, X. Shi, D. Schwarzer, W. Plunkett, B. Aronow, O. Gozani, W. Fischle, M. C. Hung, D. J. Patel and M. C. Barton (2010). "TRIM24 links a non-canonical histone signature to breast cancer." Nature **468**(7326): 927-932.
Tsukamoto, T., N. Hashiguchi, S. M. Janicki, T. Tumbar, A. S. Belmont and D. L. Spector (2000). "Visualization of gene activity in living cells." Nat Cell Biol **2**(12): 871-878.
Tumbar, T., G. Sudlow and A. S. Belmont (1999). "Large-scale chromatin unfolding and remodeling induced by VP16 acidic activation domain." J Cell Biol **145**(7): 1341-1354.
Udeshi, N. D., P. Mertins, T. Svinkina and S. A. Carr (2013). "Large-scale identification of ubiquitination sites by mass spectrometry." Nat Protoc **8**(10): 1950-1960.
Urnov, F. D., E. J. Rebar, M. C. Holmes, H. S. Zhang and P. D. Gregory (2010). "Genome editing with engineered zinc finger nucleases." Nat Rev Genet **11**(9): 636-646.
van Wijk, S. J., E. Fiskin, M. Putyrski, F. Pampaloni, J. Hou, P. Wild, T. Kensche, H. E. Grecco, P. Bastiaens and I. Dikic (2012). "Fluorescence-based sensors to monitor localization and functions of linear and K63-linked ubiquitin chains in cells." Mol Cell **47**(5): 797-809.
Vanamee, E. S., S. Santagata and A. K. Aggarwal (2001). "FokI requires two specific DNA sites for cleavage." J. Mol. Biol. **309**: 69-78.
Vance, K. W. and C. P. Ponting (2014). "Transcriptional regulatory functions of nuclear long noncoding RNAs." Trends Genet **30**(8): 348-355.
Vella, P., A. Scelfo, S. Jammula, F. Chiacchiera, K. Williams, A. Cuomo, A. Roberto, J. Christensen, T. Bonaldi, K. Helin and D. Pasini (2013). "Tet proteins connect the O-linked N-acetylglucosamine transferase Ogt to chromatin in embryonic stem cells." Mol Cell **49**(4): 645-656.

Voo, K. S., D. L. Carlone, B. M. Jacobsen, A. Flodin and D. G. Skalnik (2000). "Cloning of a mammalian transcriptional activator that binds unmethylated CpG motifs and shares a CXXC domain with DNA methyltransferase, human trithorax, and methyl-CpG binding domain protein 1." Mol Cell Biol **20**(6): 2108-2121.
Waddington, C. H. (1942, 2012). "The epigenotype. 1942." Int J Epidemiol **41**(1): 10-13.
Waddington, C. H. (1968). "Towards a theoretical biology." Nature **218**(5141): 525-527.
Wallrath, L. L. (1998). "Unfolding the mysteries of heterochromatin." Curr. Opin. Genet. Dev. **8**: 147-153.
Wang, B., D. N. M. Jones, B. P. Kaine and M. A. Weiss (1998). "High-resolution structure of an archaeal zinc ribbon defines a general architectural motif in eukaryotic RNA polymerases." Structure **6**(5): 555-569.
Wang, C., J. Shen, Z. Yang, P. Chen, B. Zhao, W. Hu, W. Lan, X. Tong, H. Wu, G. Li and C. Cao (2011). "Structural basis for site-specific reading of unmodified R2 of histone H3 tail by UHRF1 PHD finger." Cell Res **21**(9): 1379-1382.
Wang, H., L. Wang, H. Erdjument-Bromage, M. Vidal, P. Tempst, R. S. Jones and Y. Zhang (2004). "Role of histone H2A ubiquitination in Polycomb silencing." Nature **431**(7010): 873-878.
Wang, Z., C. Zang, J. A. Rosenfeld, D. E. Schones, A. Barski, S. Cuddapah, K. Cui, T. Y. Roh, W. Peng, M. Q. Zhang and K. Zhao (2008). "Combinatorial patterns of histone acetylations and methylations in the human genome." Nat Genet **40**(7): 897-903.
Watanabe, D., I. Suetake, T. Tada and S. Tajima (2002). "Stage- and cell-specific expression of Dnmt3a and Dnmt3b during embryogenesis." Mech Dev **118**(1-2): 187-190.
Wiedemann, S. M., S. N. Mildner, C. Bonisch, L. Israel, A. Maiser, S. Matheisl, T. Straub, R. Merkl, H. Leonhardt, E. Kremmer, L. Schermelleh and S. B. Hake (2010). "Identification and characterization of two novel primate-specific histone H3 variants, H3.X and H3.Y." J Cell Biol **190**(5): 777-791.
Williams, K., J. Christensen, M. T. Pedersen, J. V. Johansen, P. A. Cloos, J. Rappsilber and K. Helin (2011). "TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity." Nature **473**(7347): 343-348.
Wolfe, S. A., L. Nekludova and C. O. Pabo (2000). "DNA recognition by Cys2His2 zinc finger proteins." Annu. Rev. Biophys. Biomol. Struct. **29**: 183-212.
Wong, E., K. Yang, M. Kuraguchi, U. Werling, E. Avdievich, K. Fan, M. Fazzari, B. Jin, A. M. Brown, M. Lipkin and W. Edelmann (2002). "Mbd4 inactivation increases Cright-arrowT transition mutations and promotes gastrointestinal tumor formation." Proc Natl Acad Sci U S A **99**(23): 14937-14942.
Workman, J. L. and R. E. Kingston (1998). "Alteration of nucleosome structure as a mechanism of transcriptional regulation." Annu Rev Biochem **67**: 545-579.
Wu, H. and Y. Zhang (2011). "Tet1 and 5-hydroxymethylation: a genome-wide view in mouse embryonic stem cells." Cell Cycle **10**(15): 2428-2436.
Wu, H. and Y. Zhang (2014). "Reversing DNA methylation: mechanisms, genomics, and biological functions." Cell **156**(1-2): 45-68.
Wysocka, J., T. Swigut, H. Xiao, T. A. Milne, S. Y. Kwon, J. Landry, M. Kauer, A. J. Tackett, B. T. Chait, P. Badenhorst, C. Wu and C. D. Allis (2006). "A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling." Nature **442**(7098): 86-90.
Xiao, H., A. Pearson, B. Coulombe, R. Truant, S. Zhang, J. L. Regier, S. J. Triezenberg, D. Reinberg, O. Flores, C. J. Ingles and et al. (1994). "Binding of basal transcription factor TFIIH to the acidic activation domains of VP16 and p53." Mol Cell Biol **14**(10): 7013-7024.
Xie, S., J. Jakoncic and C. Qian (2012). "UHRF1 double tudor domain and the adjacent PHD finger act together to recognize K9me3-containing histone H3 tail." J Mol Biol **415**(2): 318-328.

Xu, C., C. Bian, R. Lam, A. Dong and J. Min (2011). "The structural basis for selective binding of non-methylated CpG islands by the CFP1 CXXC domain." Nat Commun **2**: 227.
Xu, F., K. Zhang and M. Grunstein (2005). "Acetylation in histone H3 globular domain regulates gene expression in yeast." Cell **121**(3): 375-385.
Xu, Y., F. Wu, L. Tan, L. Kong, L. Xiong, J. Deng, A. J. Barbera, L. Zheng, H. Zhang, S. Huang, J. Min, T. Nicholson, T. Chen, G. Xu, Y. Shi, K. Zhang and Y. G. Shi (2011). "Genome-wide regulation of 5hmC, 5mC, and gene expression by Tet1 hydroxylase in mouse embryonic stem cells." Mol Cell **42**(4): 451-464.
Xu, Y., C. Xu, A. Kato, W. Tempel, J. G. Abreu, C. Bian, Y. Hu, D. Hu, B. Zhao, T. Cerovina, J. Diao, F. Wu, H. H. He, Q. Cui, E. Clark, C. Ma, A. Barbara, G. J. Veenstra, G. Xu, U. B. Kaiser, X. S. Liu, S. P. Sugrue, X. He, J. Min, Y. Kato and Y. G. Shi (2012). "Tet3 CXXC domain and dioxygenase activity cooperatively regulate key genes for Xenopus eye and neural development." Cell **151**(6): 1200-1213.
Yan, W., C. Smith and L. Cheng (2013). "Expanded activity of dimer nucleases by combining ZFN and TALEN for genome editing." Sci Rep **3**: 2376.
Yang, H., Y. Liu, F. Bai, J. Y. Zhang, S. H. Ma, J. Liu, Z. D. Xu, H. G. Zhu, Z. Q. Ling, D. Ye, K. L. Guan and Y. Xiong (2013). "Tumor development is associated with decrease of TET gene expression and 5-methylcytosine hydroxylation." Oncogene **32**(5): 663-669.
Yang, M., W. S. May and T. Ito (1999). "JAZ requires the double-stranded RNA-binding zinc finger motifs for nuclear localization." J Biol Chem **274**(39): 27399-27406.
Yoon, H. G., D. W. Chan, A. B. Reynolds, J. Qin and J. Wong (2003). "N-CoR mediates DNA methylation-dependent repression through a methyl CpG binding protein Kaiso." Mol Cell **12**(3): 723-734.
Yu, Z., P. A. Genest, B. ter Riet, K. Sweeney, C. DiPaolo, R. Kieft, E. Christodoulou, A. Perrakis, J. M. Simmons, R. P. Hausinger, H. G. van Luenen, D. J. Rigden, R. Sabatini and P. Borst (2007). "The protein that binds to DNA base J in trypanosomatids has features of a thymidine hydroxylase." Nucleic Acids Res **35**(7): 2107-2115.
Yun, M., J. Wu, J. L. Workman and B. Li (2011). "Readers of histone modifications." Cell Res **21**(4): 564-578.
Zeng, L., Q. Zhang, S. Li, A. N. Plotnikov, M. J. Walsh and M. M. Zhou (2010). "Mechanism and regulation of acetylated histone binding by the tandem PHD finger of DPF3b." Nature **466**(7303): 258-262.
Zhang, G. and S. Pradhan (2014). "Mammalian epigenetic mechanisms." IUBMB Life **66**(4): 240-256.
Zhang, H., X. Zhang, E. Clark, M. Mulcahey, S. Huang and Y. G. Shi (2010). "TET1 is a DNA-binding protein that modulates DNA methylation and gene transcription via hydroxylation of 5-methylcytosine." Cell Res **20**(12): 1390-1393.
Zhang, J., Q. Gao, P. Li, X. Liu, Y. Jia, W. Wu, J. Li, S. Dong, H. Koseki and J. Wong (2011). "S phase-dependent interaction with DNMT1 dictates the role of UHRF1 but not UHRF2 in DNA methylation maintenance." Cell Res **21**(12): 1723-1739.
Zhang, L., X. Lu, J. Lu, H. Liang, Q. Dai, G. L. Xu, C. Luo, H. Jiang and C. He (2012). "Thymine DNA glycosylase specifically recognizes 5-carboxylcytosine-modified DNA." Nat Chem Biol **8**(4): 328-330.
Zhang, Q., X. Liu, W. Gao, P. Li, J. Hou, J. Li and J. Wong (2014). "Differential regulation of the ten-eleven translocation (TET) family of dioxygenases by O-linked beta-N-acetylglucosamine transferase (OGT)." J Biol Chem **289**(9): 5986-5996.
Zhang, X., J. Yazaki, A. Sundaresan, S. Cokus, S. W. Chan, H. Chen, I. R. Henderson, P. Shinn, M. Pellegrini, S. E. Jacobsen and J. R. Ecker (2006). "Genome-wide high-resolution mapping and functional analysis of DNA methylation in arabidopsis." Cell **126**(6): 1189-1201.

Zhou, T., J. Xiong, M. Wang, N. Yang, J. Wong, B. Zhu and R. M. Xu (2014). "Structural Basis for Hydroxymethylcytosine Recognition by the SRA Domain of UHRF2." Mol Cell **54**(5): 879-886.
Zhou, Y., R. Santoro and I. Grummt (2002). "The chromatin remodeling complex NoRC targets HDAC1 to the ribosomal gene promoter and represses RNA polymerase I transcription." Embo j **21**(17): 4632-4640.
Zhu, H., V. Joliot and R. Prywes (1994). "Role of transcription factor TFIIF in serum response factor-activated transcription." J Biol Chem **269**(5): 3489-3497.
Zolghadr, K., O. Mortusewicz, U. Rothbauer, R. Kleinhans, H. Goehler, E. E. Wanker, M. C. Cardoso and H. Leonhardt (2008). "A fluorescent two-hybrid assay for direct visualization of protein interactions in living cells." Mol Cell Proteomics **7**(11): 2279-2287.

5.2 Abbreviations

2-OG	2-oxoglutarate
5caC	5-carboxycytosine
5fC	5-formylcytosine
5hmC	5-hydroxymethylcytosine
5hmU	5-hydroxyuracil
5mC	5-methylcytosine
5'RACE	Rapid amplification of cDNA 5’ ends
AID	Activation induced cytidine deaminase
AIRE	Autoimmune regulator
ALS	Amyotrophic lateral sclerosis
APOBEC	Apolipoprotein B mRNA editing enzyme catalytic polypepetide
APS	Ammonium peroxodisulfate
BAH	Bromo-adjacent homology domains
BER	Base excision repair
BPTF	Bromodomain PHD finger transcription factor
BSA	Bovine serum albumin
CBX	Chromobox homolog
CCR5	C-C chemokine receptor type 5
CFP1	CXXC finger protein 1
CGBP	CpG-binding protein
Ch	Red fluorescent protein Cherry
CHD4	Chromodomain helicase DNA binding protein 4
ChIP	Chromatin immunoprecipitation
Co-IP	Co-immunoprecipitation
CpG	Cytosine and guanine dinucleotides
DAPI	4’,6-diamidino-2-phenylindol
DMSO	Dimethylsulfoxide
DNA	Deoxyribonucleic acid
DNaseI	Deoxyribonuclease I
DNMT1	DNA methyltransferase 1
DNMT3L	DNA methyltransferase 3-like protein
DSB	Double-strand break
DSBH	Double-stranded β-helix
dsRNA	Double stranded RNA
DVL	Disheveled
EGF	Epidermal growth factor

EGFP	Enhanced green fluorescent protein
EGR1	Early growth response protein 1
ESCs	Embryonic stem cells
F3H	Fluorescent three hybrid assay
FACS	Fluorescence activated cell sorting
FBS	Fetal bovine serum
FRAP	Fluorescence recovery after photobleaching
Gapdh	Glyceraldehyde phosphate dehydrogenase
GBM	Glioblastoma multiforme
GBP	GFP binding protein
GDNF	Glial cell line-derived neurotrophic factor
GFP	Green fluorescent protein
H3K18ac	Histone H3 acetyl lysine 18
H3K27me3	Trimethylated lysine 27 of histone H3
H3K4me0	Unmodified lysine 4 of histone H3
H3K4me3	Trimethylated lysine 4 of histone H3
H3K56	Lysine 56 of histone H3
H3K9	H3 at lysine 9
H3K9me3	Trimethylated lysine 9 of histone H3
H3K9MT	H3K9 methyltransferase
H3R2	Unmodified arginine 2 of histone H3
HAT3	Histone acetyltransferase 3
HDACs	Histone deacetylases
HEK293T	Human embryonic kidney 293T
His	Histidine
HIV	Human immunodeficiency virus
HMT-MBD	Histone methyltransferase
HP1	Heterochromatin protein 1
hPRC1L	Human polycomb repressive complex 1-like
HSPCs	Hematopoietic stem/progenitor cells
HSV	Herpes simplex virus
IKZF1	DNA-binding protein Ikaros family zinc finger protein 1
IPTG	Isopropyl-β-D-thiogalactopyranoside
JBP1	Base J-binding protein 1
KAP1	KRAB-associated protein 1
KBS	KAISO binding sequence
KDM2	Lysine-specific demethylase 2
KG repeats	Glycine-lysine repeats
KRAB	Kruppel-associated box

lacI	Lac repressor protein
*lac*O	*Lac* operator
LIF	Leukemia inhibitory factor
MBD	Methyl-CpG-binding domain
MBDs	Methyl-CpG binding domain proteins
MBP	Methyl-CpG binding protein
MeCP2	Methyl CpG binding protein 2
MLL1	Mixed lineage leukemia protein 1
N-CoR	Nuclear receptor corepressor
NEM	N-ethylmaleimide
NSCs	Neural stem cells
NTR	N-terminal regulatory region
NuRD	Nucleosome remodeling and deacetylase complex
ORF	Open reading frame
PAGE	Polyacrylamide gel electrophoresis
PBD	Proliferating cell nuclear antigen (PCNA) binding domain
PBS	Phosphate buffered saline
PCAF	P300/CBP-associated factor
PcG	Polycomb-group protein
PCNA	Proliferating cell nuclear antigen
PCNP	PEST proteolytic signal-containing nuclear protein
PCR	Polymerase chain reaction
PD	Parkinson's disease
PDB	Protein data bank
PEI	Polyethylenimine
PGC	Primordial germ cells
PHD	Plant homeo domain
PI	Propidium iodide
PIRH2	P53-induced protein with a RING-H2 domain
PML	Promyelocytic leukemia body
PMSF	Phenylmethylsulfonylflouride
Pol II	DNA polymerase II
PTMs	Posttranslational modifications
qPCR	Real-time PCR
RAD6	Radiation sensitivity protein 6
RAG2	V(D)J recombination-activating protein 2
RBCC	RING-B box-coiled-coil
RFP	Red fluorescent protein
RING	Really interesting new gene

RNA	Ribonucleic acid
SDS	Sodium dodecyl sulfate
SEM	Standard error of the mean
SETDB1	SET domain bifurcated 1 protein
Ski	Sloan-Kettering institute
SMUG1	Single-strand selective monofunctional uracil DNA glycosylase
SRA	SET and RING associated domain
SUMO	Small ubiquitin like modifier
TAD	Transcription activation domain
TALENs	TALE-based nucleases
TALEs	Transcription activator-like effectors
TCA	Trichloroacetic acid
TCF	T-cell factor (in Wnt pathway)
TDG	Thymine-DNA glycosylase
TET1	Ten-eleven translocation methylcytosine dioxygenase 1
TFIIA	Transcription factor TFIIA
TFIIIA	Transcription factor IIIA
TMAC	Tetramethylammonium chloride
TRDMT1	tRNA aspartic acid methyltransferase 1
TRIM24	Tripartite motif-containing protein 24
TS	Heterochromatin targeting sequence
TTD	Tandem tudor domain
Uba	Ubiquitin association domain
UBL	Ubiquitin-like domain
UDG	Uracil DNA glycosylase
UHRF1	Ubiquitin-like PHD and RING finger domain-containing protein 1
UHRF2	Ubiquitin-like PHD and RING finger domain-containing protein 2
UiFC	Ubiquitination-induced fluorescence complementation assay
UIM	Ubiquitin interacting motif
USP7	Ubiquitin-specific processing protease 7
VEGF-A	Vascular endothelial growth factor A
VP16	Herpes simplex virus protein vmw65
wt	Wild type
YY1	Yin yang 1 protein
CXXC	CXXC type zinc finger
ZFNs	Zinc finger nucleases
ZFPs	Zinc finger proteins
ZFP-TFs	Zinc finger transcription factors
Zn	Zinc

5.3 Declaration

Eidesstattliche Erklärung

Ich versichere hiermit an Eides statt, dass die vorgelegte Dissertation von mir selbständig und ohne unerlaubte Hilfe angefertigt ist.

München, den ....

(unterschift)

Erklärung

Hiermit erkläre ich, *

☐ dass die Dissertation nicht ganz oder in wesentlichen Teilen einer anderen Prüfungskommission vorgelegt worden ist

☐ dass ich mich anderweitig einer Doktorprüfung ohne Erfolg nicht unterzogen habe.

☐ dass ich mich mit Erfolg der Doktorprüfung im Hauptfach.....

und in den Nebenfächern.....

bei der Fakultät für..... der .....

(Hochschule/ Universität)

unterzogen habe

☐ dass ich ohne Erfolg versucht habe, eine Dissertation einzureichen oder mich der Doktorprüfung zu unterziehen.

München, den..... .....

(unterschrift)

*) Nichtzutreffendes streichen

5.4 Acknowledgements

First and foremost I offer my sincerest gratitude to my supervisor Prof. Dr. Heinrich Leonhardt for all your guidance, patience and support during these years. I consider myself very lucky for being able to work with a very considerate and encouraging professor like you. I also would like to thank you for allowing me the space and freedom I needed to work. The experiences during these four years have been very influential not only for my scientific thinking but also for my personal development.

Second, I am very grateful to Dr. Fabio Spada for his supervision during the first two years. Thank you for your brilliant ideas, motivating discussions and appreciation of my work.

Also, I would like to express my heartfelt gratitude to Dr. Weihua Qin for your nice supervision during these four years. Without your support and active participation in every step of the process, my PhD may never have been completed. Not only have you been generous with your time and expertise, but also did share with me your wide scientific knowledge, enlightening ideas and inspiring discussions.

Besides, I would like to thank the Chinese Scholarship Council for the four years' financial support for my PhD study. Without it, I would not have had the opportunity to pursue my PhD in such a nice group.

As well, I would like to thank Wen for always taking time to help me with microscopy work as well as for your support during my thesis writing. Also, I would like to thank Christina for helping me with EMSA assays and for her nice suggestions on my thesis writing. In addition, I would like to thank Weihua, Wen, Katrin, Daniel, Cindy, Pin and Congdi for the nice office atmosphere that we had and have. I am sure I will miss you a lot.

Furthermore, I would like to thank Anja for the nice organization of our daily lab life, and Susanne, I want to say thank you, especially for helping me settle down in Munich and for always taking your time to answer all my questions.

Moreover, I would like to thank Mengxi, Patricia, Martha and Stephanie for the nice collaborations. Thanks to Andrea for your nice suggestions and discussions on the generation of antibodies. Thanks to Daniela and Sebastian for your help with FACS sorting.

5 Annex

Thanks to Christine and Carina for the introduction on qPCR and DNA binding assay. I also would like to thank Irina and Congdi for helping me with microscope operation and histological immunostaining. Thanks to Katharina and Chris for the introduction to northern blotting assays and the “T7 assay”. Thanks also go to Jonas, Bijan and Ola for the instructions on protein expression and purification. Besides, I would like to thank all other current and former members of the Leonhardt group, Boris, Sylvia, Hartmann, Yolanda, Tobi, Udo, Elisabeth, Kamila, Andreas, Jürgen, and Garwin for being very nice colleagues.

I am deeply indebted to my family, especially my Dad and Mom who always have faith in me and encourage me in every decision I made in my life. Their love and support led me to this place. Thank you for everything. Besides, I want to thank all my sisters for being a wonderful family as well as being good friends. Last but not least, I would like to thank my boyfriend Siwei for his love, unconditional support and for helping me getting through hard times.

6 Curriculum Vitae

Curriculum Vitae

Nan Liu

Telephone: +49 89 2180 74227

liu.nan@biologie.uni-muenchen.de

Personal information

Date of Birth	01.28.1983
Place of birth	Jilin, China
Nationality	Chinese
Current work address	Ludwig Maximilians University (LMU), Munich
	Department of Biology II, Biocenter
	Großhaderner Str. 2, 82152, Planegg-Martinsried, Germany

Education

10/2010 – Present	PhD student in Biology,
	Ludwig Maximilians University, Munich, Germany
	Supervisor: Prof. Dr. Heinrich Leonhardt
09/2007-07/2009	Master of Science in Biomedical Engineering,
	School of Pharmaceutical Sciences, Jilin University, Jilin, China
09/2002-07/2007	Bachelor of Science and Medicine in Bioengineering (Medical),
	School of Pharmaceutical Sciences, Jilin University, Jilin, China

7 Publications

7 Publications

Liu N*, Wang M*, Deng W, Schmidt CS, Qin W, Leonhardt H and Spada F (2013). Intrinsic and extrinsic connections of Tet3 Dioxygenase with CXXC zinc finger modules. PLoS One, 2013 May 14;8(5):e62755. (*Equally contributed to this work)

Yan H*, Liu N*, Zhao Z, Zhang X, Xu H, Shao B, Yan W (2012). Expression and purification of human TAT-p53 fusion protein in Pichia pastoris and its influence on HepG2 cell apoptosis. Biotechnol Lett, 2012 Jul;34(7):1217-23. (*Equally contributed to this work)

Weihua Qin, Patricia Wolf, Nan Liu, Stephanie Link, Federica La Mastra, Ignasi Forné, Karin Fellinger, Garwin Pichler, Martha Smets, Fabio Spada, Ian Marc Bonapace, Axel Imhof and Heinrich Leonhardt. DNMT1 ubiquitin interacting motif UIM is required for maintenance of DNA methylation by binding to UHRF1-dependent H3 ubiquitination. (Submitted)

Qin, Weihua; Liu, Nan; Leonhardt, Heinrich (2014), Uhrf1 controls cell cycle dependent stability of heterochromatin protein Cbx1 by polyubiquitination. (In preparation)