Structural biology-based insights into combinatorial readout and crosstalk among epigenetic marks[*]

Jiamu Du[a,b,*] and **Dinshaw J. Patel**[a,**]

[a]Structural Biology Program, Memorial Sloan-Kettering Cancer Center, New York, NY 10065, USA

[b]Shanghai Center for Plant Stress Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 201602, China

Abstract

Epigenetic mechanisms control gene regulation by writing, reading and erasing specific epigenetic marks. Within the context of multi-disciplinary approaches applied to investigate epigenetic regulation in diverse systems, structural biology techniques have provided insights at the molecular level of key interactions between upstream regulators and downstream effectors. The early structural efforts focused on studies at the single domain-single mark level have been rapidly extended to research at the multiple domain–multiple mark level, thereby providing additional insights into connections within the complicated epigenetic regulatory network. This review focuses on recent results from structural studies on combinatorial readout and crosstalk among epigenetic marks. It starts with an overview of multiple readout of histone marks associated with both single and dual histone tails, as well as the potential crosstalk between them. Next, this review further expands on the simultaneous readout by epigenetic modules of histone and DNA marks, thereby establishing connections between histone lysine methylation and DNA methylation at the nucleosomal level. Finally, the review discusses the role of pre-existing epigenetic marks in directing the writing/erasing of certain epigenetic marks. This article is part of a Special Issue entitled: Molecular mechanisms of histone modification function.

Keywords

Combinatorial readout; Epigenetic regulation; Histone modification; DNA methylation

1. Introduction

Gene regulatory mechanisms that are unrelated to the DNA genetic code have been classified as epigenetic events, mediated by covalent modifications on histones, DNA and RNA, chromatin remodeling, micro-RNA mediation, and higher-ordered chromatin architectures. At the fundamental level, ~146 bp of DNA wraps around a histone octamer, composed of two copies of the four histone proteins, H2A, H2B, H3, and H4, thereby

---

forming the nucleosome core particle [1]. The nucleosome core particle constitutes the basic building block of chromatin and provides the structural unit and platform for mediating epigenetic regulation events. The flexible N-terminal tails of histones, which extend from the nucleosome core segment, contain a range of site-specific posttranslational modifications, including methylation, acetylation and ubiquitination of lysine, methylation of arginine, and phosphorylation of serine, threonine and tyrosine residues. In addition, some residues located within the nucleosome core can also be modified. The multiple posttranslational modifications on histone tails form specific patterns defined as the 'histone code', with the marks deposited, read and removed by specific enzymes in a sequence- and modification-specific manner [2]. Both the existence and/or absence of certain posttranslational modifications or marks on histone tails provide unique docking sites for specific binding and/or release of downstream effector proteins, resulting in diverse biological functions including transcription regulation, cell cycle control, differentiation and apoptosis. In addition to histone tail modifications, DNA methylation at the 5 position of the cytosine base constitutes another common covalent epigenetic mark, which contributes to gene silencing [3]. The methylated DNA provides DNA methylation-specific binding sites for MBD (methyl-DNA binding domains), SRA (SET and RING-associated domains), and some specific zinc finger-containing proteins [4]. In mammals, DNA methylation mainly occurs at symmetric CG sites and impacts on important biological processes ranging from regulation of gene expression, to genome imprinting and X-chromosome inactivation [3,5]. In plants, DNA methylation is much more diversified and occurs at three sequence contexts: CG, CHG (H refers to A, T, or C) and CHH, and plays an important role in the suppression of both transposable and repetitive DNA sequence elements [3,5].

Although an individual epigenetic mark may have its own downstream effectors or specific roles, a complex epigenetic landscape usually requires that the various epigenetic marks work together, capitalizing on different mechanisms that control the dynamic equilibrium among marks. In 2003, the C. David Allis laboratory introduced the pioneering concept of 'binary switches' and 'modification cassettes' to the field and for the first time uncovered the possible interrelationship among epigenetic marks within a local, closely spaced region [6]. Systematic recent studies have extended this concept to the specific readout of combinations of multiple epigenetic marks, which are classified as multivalent readout and crosstalk among epigenetic marks [7]. The epigenetic marks can work together either in a compatible or mutually exclusive manner, at the single histone tail level or in the context of the nucleosome or even the chromatin level, all of which provide insights into the complex regulation of gene expression by epigenetic mechanisms and deepen our comprehension of the field.

More than 10 families of different histone modification-specific reader modules and three methylated DNA-specific reader modules have been characterized by structural approaches at the histone peptide or isolated DNA level. Several excellent reviews have been written that have provided a comprehensive overview of the molecular mechanisms associated with readout of epigenetic marks [4,8–11]. In this review, we would like to primarily focus on the structural biology of multivalent readout and crosstalk among epigenetic marks. We will discuss the multivalent readout of histone modifications from the perspective of a single histone tail, as well as the multivalent readout of histone modifications from different

histone tails, together with an extension to the concept of multivalent readout at the nucleosome level. Next, we will expand our understanding of readout of histone modifications to DNA methylation marks, which constitute an equally important epigenetic modification. Finally, we would like to elucidate the role of pre-existing epigenetic marks in directing writing/erasing of epigenetic marks in histone modification and DNA methylation pathways. The multivalent readout and crosstalk occur both at the single protein level and at the multiple protein complex level, with this review focused on the former case given the predominance of the published literature to date on such systems.

2. Multivalent readout of histone marks positioned on a single histone tail

The nucleosome core particle constitutes the basic platform for epigenetic regulation. The N-terminal tails of histones, which extend out from the core region, contain the majority of posttranslational modifications compared to the core segment. Thus, a range of epigenetic marks coexist in a sequential context on the same histone tail, with recognition of multiple histone marks on the same tail achieved through specific readout by multiple tandem linked reader modules. With each of the modules recognizing a specific modification, the combination of two or more modules together can usually promote enhanced binding affinity on the basis of cooperativity between modules. Such cooperativity can result in enhanced binding for a given pattern of modifications. Currently, the majority of documented studies are available for paired chromatin-associated modules [12].

The first reported paired chromatin-associated module which exhibited combinatorial effects on two or more histone marks was the double bromodomain of $TAF_{II}250$ (TBP-associated factor 250 kDa), the largest subunit of TFIID (RNA polymerase II transcription factor D) (Fig. 1A) [13]. The double bromodomain of $TAF_{II}250$ has two side-by-side pockets, which serve as acetyllysine-binding pockets, each reflective of a classic bromodomain fold [14]. The isothermal calorimetric titration (ITC) data indicated that the double bromodomain bound to H4K16ac peptide with a dissociation constant (K_D) of 39 μM. Further, the binding is 7 to 28 fold enhanced on recognition of di/tetra-acetyllysine-modified H4 peptides (H4K8ac/K16ac and H4K5ac/K8ac/K12ac/K16ac) [13]. The crystal structure of $TAF_{II}250$ double bromodomain in the free state revealed that the binding pockets of the two bromodomains span a distance of 25 Å, thereby serving as a molecular ruler that spans about seven or more peptide residues (Fig. 1A). This should allow two acetylated-lysines to be simultaneously read by the linked bromodomains, thereby providing the structural basis for enhanced binding on recognition of a multiply acetylated H4 tail at the single peptide level [13]. $TAF_{II}250$ is the subunit of the transcription factor TFIID, with functional implications during initiation of transcription. Hyper-acetylation of histone proteins within their promoter regions represents a hallmark of transcriptionally active genes. The structural studies provide a plausible mechanism whereby hyper-acetylated promoter regions of transcriptionally active genes can successfully recruit $TAF_{II}250$ with high affinity and compete against hypo-acetylated genes with weaker binding affinities for $TAF_{II}250$, resulting in the initiation of target genes. Nevertheless, the underlying mechanism of combinatorial readout at the molecular level remains to be established, given the lack to date of the structure of a multiple acetylated-lysine H4 peptide bound to the double bromodomain of $TAF_{II}250$.

Typically, the PHD (Plant Homeo Domain) finger has been reported to specifically recognize various epigenetic marks, including the H3K4me3 epigenetic mark or its unmodified H3K4me0 counterpart, H3K9me3, and H3K36me3 [15,16]. Nevertheless, the tandem PHD fingers of DPF3b (D4, Zinc and Double PHD Fingers, Family 3b), a BAF (BRG1- or HRBM-associated factors) chromatin remodeling complex associated protein, were reported to recognize acetylated histone H3 and H4 tails, with important functional implications in the initiation of transcription during heart and muscle developments [17]. The DPF3b tandem PHD fingers bound to unmodified H3(1–18) peptide with a K_D of 2.3 μM, which was shown to exhibit an increase in binding affinity by four-fold on acetylation of H3K14 [18]. The NMR (Nuclear Magnetic Resonance) solution structure of DPF3b tandem PHD fingers in complex with the H3(1–20)K14ac peptide revealed the structural mechanism whereby the two PHD fingers cooperatively read the bound H3K14accontaining peptide (Fig. 1B) [18]. The N-terminal *** group of H3 was anchored within a negatively charged pocket of PHD2. The Arg2, Lys4 and Lys9 side chains of the histone peptide insert into the interface between the two PHD finger domains and interact with several acidic residues from PHD2, as well as some residues from PHD1 (Fig. 1B). The specific recognition of acetylated K14 was achieved by projecting its side chain into a surface pocket of PHD1. It is worth noting that further modification of K4, such as methylation or acetylation, resulted in a loss of binding affinity by 15 to 20 fold, indicating that recognition required dual readout of unmodified H3K4 and H3K14ac marks [18]. Thus, the tandem PHD fingers of DPF3b use its PHD2–PHD1 interface and PHD1 finger to accommodate the unmodified H3K4 and acetylated H3K14, respectively (Fig. 1B). An adjacent PHD finger not only generates an additional binding site at the individual domain level, but also creates one more binding site between the two domains. Such recognition largely extends the interaction surface and target selection sites, in the process of achieving higher binding affinity and selectivity. DPF3b functions in the initiation of transcription, a process characterized by hyper-acetylation within the promoter region. The enhanced binding with acetylated H3K14 enabled the chromatin remodeling complex to pause at the pre-initiating locus, while the unmodified H3K4-specific binding allowed the release of the initiated activation sites, thereby revealing the dynamic regulation of transcription by the different histone modification patterns. In another example, similar results have been observed for the double PHD fingers of MOZ (human histone acetyltransferase monocytic leukemia zinc finger protein), which cooperatively recognized unmodified H3R2 by PHD2 and acetylated H3K14 by PHD1 [19], using recognition principles similar to those observed for DPF3b. Although H3K14ac recognition was not observed in the crystal structure of the MOZ complex due to blockage by a bound acetate molecule, the NMR titration and ITC binding data successfully located the H3K14ac binding pocket, with acetylation of H3K14 increasing the binding affinity by about three-fold [19]. The enhanced binding affinity of MOZ for the H3 tail most likely facilitated targeting to the promoter of *HOXA9*, resulting in more acetylation of H3K14 marks, thereby resulting in a positive feedback mechanism and activation of the target gene.

In addition to the tandem readers of the same family of domains, there are additional combinations of different reader modules within individual multi-domain proteins. The PHD finger and bromodomain frequently exist adjacent to each other as PHD finger-

bromodomain (PHD-Bromo) cassettes in several multi-domain proteins. Two recent studies on the combinatorial recognition of histone marks by the PHD-Bromo cassette of TRIM (Tripartite Motif) family proteins have greatly improved our comprehension on the readout of multiple histone marks. The TRIM family protein TRIM24 has a C-terminal paired PHD-Bromo cassette. The PHD finger of TRIM24 can specifically recognize unmodified H3K4 mark using a classic recognition mode (hydrogen bonding to peptide carbonyl and acidic side chains) as revealed by both the binding studies and the crystal structure of the complex (Fig. 1C) [20]. Besides, the PHD finger and bromodomain of TRIM24 physically interact with each other, which raise the possibility that the bromodomain of the PHD-Bromo cassette may recognize another mark within the same histone tail. Indeed, this postulate was confirmed by ITC measurements that identified H3K23ac as the mark recognized by the bromodomain. Further, a crystal structure of TRIM24 PHD Bromo cassette in complex with H3(13–32)K23ac peptide established the molecular mechanism underlying complex formation (Fig. 1C). The combination of unmodified H3K4 and H3K23ac increased the binding of TRIM24 to the H3 tail to a K_D of 0.096 μM, compared with individual binding with K_D of 2.3 μM and 8.8 μM, respectively [20]. This example presents further structural insight into a paired module that recognized two tandem marks within a single histone tail, for which binding studies supported a combinatorial readout mechanism [20]. It is worth noting that combinatorial readout of the two marks was not structurally visualized in a single complex due to the difficulty in getting the crystals of such a complex.

Subsequent structural studies on a related TRIM protein, TRIM33, overcame the above crystallization obstacle and established for the first time how the paired PHD-Bromo cassette is capable of recognizing multiple histone marks on the same histone tail [21]. TRIM33 has a similar domain alignment to TRIM24, with a C-terminal paired PHD-Bromo cassette that adopted a similar ternary architecture [21]. The structure of TRIM33 PHD-Bromo cassette in complex with a H3(1–28)K9me3/K14ac/K18ac/K23ac peptide revealed the principles underlying multiple readout (Fig. 1D). The PHD finger specifically recognized the unmodified K4 using a classic recognition mode and K9me3 through stacking with a tryptophan residue. For acetyllysine recognition, the peptide extended toward the bromodomain with the K18ac side chain inserted into the binding pocket of the linked bromodomain (Fig. 1D) [21]. Thus, only K18ac, but not K14ac nor K23ac, was positioned for correct selection in the context of the bound H3 peptide, indicating that the sequence context and the distance to the unmodified H3K4 and H3K9me3 contribute to the selectivity of Kac recognition by the bromodomain. TRIM33 PHD-Bromo cassette exhibits a K_D of 0.46 μM for the unmodified H3(1–28) peptide (through recognition of unmodified K4 mark). Addition of K9me3 or K18ac mark enhanced the binding affinity by around two-fold with K_D values of 0.2 μM and 0.21 μM, respectively. In sharp contrast, the combination of all the three marks together greatly increased the binding affinity by about 8-fold (K_D of 0.06 μM) [21]. Functionally, the combination of multiple marks greatly enhanced binding affinity for the H3 tail, facilitating the displacement of chromatin compacting factor HP1 (Heterochromatin Protein 1), and allowed nodal response elements access to Smad4–Smad2/3 (Similar to Mothers Against Decapentaplegic), thereby resulting in the activation of relevant genes [21]. This structure–function study was the first to demonstrate multivalent histone mark readout within a single structure unit.

3. Multivalent readout of histone marks positioned on different histone tails

Considering that each nucleosome has two copies of histones H2A, H2B, H3 and H4, and given that nucleosomes are densely distributed along the chromatin, it is conceivable that histone tails from different nucleosomes could be positioned in close proximity, thereby allowing readout of histone marks from more than one nucleosome. The multidomain human BPTF (Bromodomain and PHD Domain Transcription Factor) protein was a well-studied example to illustrate the combinatorial readout of two histone tails at the single nucleosome level [22]. BPTF is the largest subunit of the ATP-dependent chromatin-remodeling complex NURF (Nucleosome Remodeling Factor), which functions in transcription activation [23–25]. BPTF possesses a PHD-Bromo cassette toward the C-terminus of the protein. The PHD finger was identified as a reader of the H3K4me3 mark [26], a conclusion validated by both biochemical binding assays and structure determination of the complex (Fig. 2A) [27]. The crystal structure of BPTF PHD-Bromo cassette in complex with H3(1–15)K4me3 peptide highlighted the molecular mechanism underlying recognition of bound H3 peptide by the PHD finger, with the K4me3 group positioned within an aromatic cage (Fig. 2A) [27]. The bromodomain of the PHD-Bromo cassette was connected to the PHD finger through a short α-helical linker [27]. Such a rigid linker constrained both separation and binding pocket orientations of the PHD finger relative to the bromodomain, thereby maintaining a constrained architecture within the PHD-Bromo cassette (Fig. 2A).

As part of an effort to elucidate the molecular function of the bromodomain within the BPTF PHD-Bromo cassette, peptide arrays containing most known histone acetylation marks were screened, yielding H4K12ac, H4K16ac, and H4K20ac, as potential binding candidates [22]. At the peptide level, the binding between BPTF PHD-Bromo cassette with H3K4me3 peptide was not affected by the addition of saturating amounts of H4K16ac peptide. In a reciprocal experiment, the binding between BPTF PHD-Bromo cassette with H4K16ac peptide was not affected by the addition of saturating amounts of H3K4me3 peptide, indicating the absence of allosteric cooperativity between these two binding sites at the peptide level. In contrast, by using designer nucleosomes, obtained by capitalizing on a histone semi-synthesis approach termed expressed protein ligation [28,29], an in vitro pulldown assay exhibited a 2 to 3-fold enhanced binding between BPTF PHD-Bromo cassette with the H3K4me3/H4K16ac doubly-modified nucleosome compared with the nucleosome containing only the H3K4me3 modification [22]. It was notable that this type of enhancement was only achieved for the H4K16ac mark but not for the H4K12ac or H4K20ac marks, thereby establishing that the position of acetyllysine on H4 played a key role in the simultaneous readout of H3K4me3. Crystal structures of BPTF bromodomain in complexes with H4K16ac peptide revealed that the peptide can be docked in two opposing directions (Fig. 2A), indicative of a reduced selectivity. ChIP-seq (Chromatin Immunoprecipitation-sequencing) data on the localization of the PHD-Bromo cassette revealed that the distribution of the PHD-Bromo cassette generally correlated with the H3K4me3 mark but not at all H3K4me3 sites, with its localization better correlated with loci that were combined with H4 acetylation [22]. This result indicated that the enhancement of binding to nucleosomes by the BPTF bromodomain can modulate the micro-preference of

the distribution of BPTF across the chromatin. This work represented an ideal study case on multivalent readout from a single histone tail to different histone tails.

Another reported example of multivalent readout of multiple histone tails emerged from studies of DNA methylation in *Arabidopsis thaliana*. In plants, non-CG DNA methylation was shown to be abundant and mainly maintained by a family of plant specific DNA methyltransferases, such as CMT3 (chromomethylase3) [3,30]. The multi-domain protein CMT3 is composed of an N-terminal BAH (Bromo-Adjacent Homology) domain, a C-terminal DNA methyltransferase domain, together with a chromodomain embedded inside the DNA methyltransferase domain [31]. As DNA methylation in plants has been highly correlated with the histone H3K9me2 modification [32], and the chromodomain has been shown to be a well-studied histonemethylated-lysine modification recognition module [33], it appeared conceivable that the chromodomain of CMT3 most likely participated in the binding of H3K9me2 mark. Indeed, genome wide ChIP-seq data established that CMT3 co-localized with the H3K9me2 mark [34]. The direct binding between CMT3 and H3K9me2 was further confirmed by ITC and in vitro pull-down assays [34–36]. A surprising observation that emerged from the ITC binding experiments was that both CMT3 and its maize homolog ZMET2 (*Zea* methyltransferase2) yielded a stoichiometric value of around 2 upon titration with H3K9me2 peptides [34,36], indicative of CMT3/ZMET2 containing two H3K9me2 binding sites. By ITC titrations with individual domains of ZMET2, both the chromodomain and the BAH domain were shown to independently target H3K9me2 marks [34]. A crystal structure of nearly full-length ZMET2 in complex with H3(1–15)K9me2 peptide in which the peptide bound to the chromodomain revealed that the ZMET2 chromodomain used a classic aromatic cage-binding mode to accommodate the H3K9me2 mark (Fig. 2B) [34]. In another crystal form, the crystal structure of ZMET2 bound to H3(1–32)K9me2 peptide established that the BAH domain of ZMET2 also used a classic aromatic cage to accommodate the H3K9me2 mark (Fig. 2B) [34]. In an effort to establish that both chromodomain and BAH domain recognition of H3K9me2 exhibited functional relevance, triple mutations of the aromatic cage residues of either the chromodomain or the BAH domain when introduced into a *cmt3* mutant *Arabidopsis* strain, resulted in a complete loss of H3K9me2 binding capacity. Sequencing of CMT3-controlled chromatin *loci* revealed that both mutants failed to restore CHG DNA methylation, suggesting that recognition of H3K9me2 by both domains is essential for its DNA methylation activity in vivo [34]. A plausible explanation for the requirement of simultaneous recognition of H3K9me2 by both domains could be that CMT3 has relatively low catalytic activity compared with other reported DNA methyltransferases such as Dnmt1 and Dnmt3a [3,37,38], so that two histone binding sites are required to direct the enzyme to chromatin regions enriched in H3K9me2 silencing marks, in the process enabling the enzyme to approach the nucleosomal DNA so as to increase the accessibility to its substrate. A modeling study provided support for this conclusion by showing that upon superposing a ZMET2 molecule onto the mononucleosome with its chromo and BAH domains targeted to the two H3 tails, its active methyl donor site could be positioned with directionality toward the nucleosomal DNA [34]. This study represented the first structural report of one protein using two different domains simultaneously to recognize two H3 tails projecting from a nucleosome, thereby shedding light on epigenetic regulation at the nucleosomal level. Given that H3K9me2 marks are

densely enriched within heterochromatin regions, it is also conceivable that the CMT3 protein could simultaneously recognize two H3K9me2 tails from adjacent nucleosomes, thereby providing a pathway for spreading of CHG DNA methylation across the heterochromatin region [34]. In addition to ZMET2, the CHD4 (Chromodomain-helicase-DNA-binding protein 4) and CHD5 exhibit similar properties. The CHD4 protein possesses two PHD fingers toward its N-terminus connected by a short linker. Both PHD fingers can recognize H3 tails with the PHD1 preferably binding to unmodified H3 tail and PHD2 slightly preferring unmodified H3K4 plus H3K9me3 [39]. This linked double PHD module allows the protein to simultaneously recognize two H3 tails of a mononucleosome or from adjacent nucleosomes. The CHD5 displays similar potential to use two tandem PHD fingers to recognize two H3 tails of nucleosome [40].

4. Multivalent readout of histone modifications in combination with DNA

DNA methylation at the C-5 position of the cytosine base is the most important DNA modification, a feature conserved from bacteria to higher eukaryotes. In bacteria, DNA methylation confers protection to the host genome against endonucleases, while invading foreign DNA lacking this epigenetic mark becomes susceptible to cleavage. In higher eukaryotes, DNA methylation acts as an important epigenetic marker and has been associated with a range of important biological process such as gene repression, X-chromosome inactivation, genomic imprinting and several human diseases [41–43]. Generally, the DNA methylation epigenetic mark is representative of gene silencing or repression [3,30]. DNA methylation and histone modification are not two independent systems, but rather exhibit extensive crosstalk between their marks [44]. Therefore, some proteins are likely to be involved in both recognition of histone and DNA mediated by their respective marks. A well-established example includes the multi-domain protein UHRF1 (Ubiquitin-like, containing PHD and RING finger domains 1) implicated as a co-factor of the maintenance DNA methyltransferase Dnmt1, in the process targeting Dnmt1 to hemimethylated replication forks [45,46]. UHRF1 possesses multiple domains, in which the N-terminal tandem Tudor and PHD finger domains and the C-terminal SRA domain are of great interest toward deciphering their role in epigenetic regulation (Fig. 3A). The SRA domain of UHRF1 can recognize hemimethylated CpG DNA [47]. Crystal structures of UHRF1 SRA domain in complex with hemimethylated DNA revealed that the SRA domain used two loops to specifically interact with the DNA and that the 5-methylcytosine was flipped out and subsequently accommodated within a pocket in the SRA domain, with stabilization associated with planar stacking contacts, hydrogen bonds and specific hydrophobic interactions involving the methyl group of 5mC (Fig. 3B) [48–50]. On the other end, structural studies established that the UHRF1 PHD finger recognized the unmodified H3 N-terminal tail, with recognition associated predominantly with unmodified R2 through a network of intermolecular hydrogen bonding interactions (Fig. 3C) [51–53]. The tandem Tudor domain of UHRF1 binds H3K9me3 by a classic aromatic cage recognition mode as revealed by both X-ray crystallography and NMR structures [54]. Two further studies of the UHRF1 tandem Tudor-PHD finger cassette in complex with H3K9me3 peptide revealed that the PHD finger can combinatorialy enhance the binding of H3K9me3 to the tandem Tudor domain (Fig. 3C) [55,56]. DNA methylation was enhanced within the

nucleosomal region relative to the linker region [57]. Thus, it is conceivable that the SRA domain most likely binds to hemimethylated CpG DNA at the nucleosomal level, which raises the possibility that UHRF1 might use its three domains (tandem Tudor, PHD finger and SRA) to coordinately interact with the nucleosome. A structure determination of the entire UHRF1 protein complexed with the nucleosome might reveal the combinatorial interaction between histone and DNA.

In addition to H3K9 methylation, H3K27me3 was also reported to be associated with de nono DNA methylation [58–60]. Different from normal cells, the polycomb complex of cancer cells can both establish the H3K27me3 mark and recruit the DNA methyltransferase, with the latter resulting in de novo DNA methylation at certain regions [59]. However, the underlying molecular mechanism remains unclear. Further, molecular functional studies could provide approaches for undertaking structural studies.

Examples also include of co-recognition by effector proteins of histone marks and unmodified DNA. The only such example with structural characterization to date is the MSL3 (Male-Specific Lethal-3) chromodomain in complex with the H4K20me1 peptide and DNA (Fig. 3D) [61]. The MSL3 chromodomain bound to the GA-rich MSL recognition element DNA with affinity K_D of 0.4 μM [62]. In the absence of DNA, MSL3 did not show any binding with tested histone marks. In sharp contrast, in the presence of its target DNA, MSL3 selectively recognized H4K20me1 peptide with a K_D of 10 μM, reflective of a DNA-dependent interaction [61]. In the crystal structure of MSL3 in complex with DNA and H4K20me1 peptide, the chromodomain forms extensive interactions with the minor groove and the sugar-phosphate backbone of the DNA. In addition, the H4K20me1 peptide interacts with both the MSL3 protein and the DNA. The H4K20me1 mark was accommodated within an aromatic cage formed by four aromatic residues of the chromodomain. Besides, H4H18 and H4R19 of the peptide are positioned within the DNA minor groove with recognition mediated by hydrogen bonding interactions (Fig. 3D), indicating that the DNA contributes to the sequence specific recognition of H4. In this case, the structure established for the first time that DNA can serve as a cofactor to assist in the specific recognition of an epigenetic histone mark. Recently, it was reported that the PWWP (Pro-Trp-Trp-Pro) domain of LEDGF/p75 (Lens epithelium-derived growth factor) can specifically recognize the H3K36me3mark, as well as non-specifically bind to DNA, which were both monitored by NMR [63]. The LEDGF PWWP domain showed mM level binding affinity for the H3K36me3 peptide and about 1.5 μM binding affinity for DNA. In sharp contrast, it binds to the H3K36me3-containing mononucleosome with a much higher affinity of about 48 nM, indicative of the combinatorial impact of the combination of two weak binding factors together achieving a higher binding affinity [63].

5. Epigenetic modification-directed generation of additional epigenetic marks

Epigenetic mechanisms control multiple gene regulation systems. The generation or elimination of a certain epigenetic mark usually requires some cellular signal to direct and trigger the onset of the reaction. This type of regulation works at the epigenetic level, resulting in epigenetic-mark based feedback and crosstalk between epigenetic marks.

5.1. Readout of histone marks directs writing/erasing of additional histone marks

The writer and eraser modules of epigenetic marks are often controlled and targeted by some pre-existing marks as reflected by the frequent coexistence of writer/eraser module with reader module within a single protein. Several writer proteins also contain reader modules within the same protein, with the potential for targeting the written product. For instance, MLL1 (Mixed Lineage Leukemia 1) can use its SET [Su(var)3–9, Enhancer-of-zeste and Trithorax] domain to deposit the H3K4me3 mark, with this mark in turn subsequently read by its third PHD finger [64–67]. Similarly, the H3K9me3 methyltransferases G9a and GLP (G9a Like Protein) have SET domains for deposition of the H3K9me mark, which in turn use their ankyrin repeat domains to specifically recognize the H3K9me mark [68,69], suggesting a self-reinforcing feedback loop mechanism. Nevertheless, the lack of direct interaction between the reader and writer modules makes the structure-based feedback mechanism beyond current reach for explanation in molecular terms.

Notably, well-characterized systems do exist, such as PHF8 (PHD finger protein 8) and KIAA1718 (also known as JHDM1D) histone lysine demethylases (KDMs) [70], as well as the KIAA1718 ortholog from *Caenorhabditis elegans*, ceKDM7A (Fig. 4) [71,72]. H3K4me3 is an activating mark, while H3K9me2 is a repressive mark. Both PHF8 and KIAA1718 harbor an N-terminal PHD finger, which targets the H3K4me3 mark and a C-terminal Jumonji domain, which can demethylate H3K9me2 or H3K27me2, linking the activating mark to the removal of a repressive mark within a single protein context. In the structure of PHF8 in complex with a H3(1–24)K4me3/K9me2 peptide and NOG (N-Oxalylglycine) cofactor, the PHF8 adopted a bent conformation such that the K4me3 binding pocket within the PHD finger was positioned in close proximity to the demethylase active site which accommodated the K9me2 mark (Fig. 4A) [70]. Thus, binding of H3K4me3 enhanced the accessibility of the Jumonji domain to the H3K9me2 mark, resulting in a 12-fold increase in enzymatic activity as revealed by activity assays [70]. By contrast, KIAA1718 and its ortholog ceKDM7A adopted an extended conformation, where by the K4me3 binding site on the PHD finger was positioned far away from the demethylase active site on the Jumonji domain (Fig. 4B) [70,72]. This increased separation restricted the accessibility of the KIAA1718/ceKDM7A Jumonji domain to the H3K9me2 mark when the H3K4me3 on the same peptide docked into the PHD finger-binding pocket, because the length between H3K4me3 and H3K9me2 on the same peptide was significantly shorter than the distance between the PHD finger binding pocket and the active site on the Jumonji domain (Fig. 4B) [70,72], which is also revealed by a significant loss in enzymatic activity for the H3K4me3/K9me2 peptide compared with the H3K9me2 peptide [70,72]. By contrast, the H3K27me2 demethylase activity of KIAA1718/ceKDM7A was enhanced on inclusion of the H3K4me3 mark in the same peptide because the distance between the H3K4me3 and H3K27me2 marks is long enough for them to occupy the PHD and Jumonji domain pockets simultaneously (Fig. 4B) [70]. These two structures have provided novel examples of how epigenetic reader modules can both up-regulate and/or downregulate another epigenetic module through contributions associated with steric effects.

5.2. Histone mark-directed DNA methylation

Both in animals and plants, DNA methylation constitutes a gene silencing or repression signal, which can be correlated with repressive histone H3K9me2/3 and unmodified H3K4 marks [3,30,44]. Thus, both establishment and maintenance of DNA methylation are directed by repressive histone modification marks. In mammals, establishment of DNA methylation is governed by the de novo DNA methyltransferase Dnmt3a/b together with their inactive counterpart Dnmt3L. Both Dnmt3a/b and Dnmt3L possess an N-terminal ADD (ATRX–DNMT3– DNMT3L) domain, which recognizes unmodified H3K4, and structures of the complex of Dnmt3a–Dnmt3L in the apo state have been reported in the literature [73,74]. The C-terminal catalytic domains of Dnmt3a and Dnmt3L form a functional hetero-tetramer with a Dnmt3a dimer in the middle flanked by C-terminal domains of Dnmt3L on either side [37]. Thus, binding of histone tails by the ADD domains of Dnmt3a and Dnmt3L could recruit the Dnmt3a–Dnmt3L tetramer to unmodified H3K4-containing nucleosomal loci, with subsequent generation of DNA methylation marks [75]. It has been reported that unmodified H3K4 tails can allosterically activate de novo Dnmt3a–Dnmt3L DNA methyltransferase activity by 8-fold and that such regulation is reversibly correlated with H3K4 methylation, providing a potential crosstalk regulation mechanism [76]. However, there is as yet no reported structure of the Dnmt3a–Dnmt3L tetramer in complex with bound histone peptide. The postulated regulatory mechanism remains to be validated at the structural level.

In plants, the DNA methyltransferase CMT3/ZMET2 can be targeted by dual recognition of H3K9me2 marks, as discussed above [34]. Interestingly, modeling studies indicated that both Dnmt3a–Dnmt3L and ZMET2 can be modeled onto the mono-nucleosome, such that the two histone-binding domains can target a pair of H3 tails extending out from the nucleosome, while the catalytic sites are positioned toward the nucleosomal DNA [34,75], suggesting that structural studies at the nucleosomal level could shed light on histone dependent regulation of DNA methylation.

5.3. DNA methylation-directed histone lysine methylation

In addition to histone mark-directed DNA methylation, DNA methylation also impacts on regulating histone modifications. In plants, non-CG DNA methylation has been shown to be highly correlated with H3K9me2 modification through a reinforcing loop between DNA methyltransferase CMT3 and histone methyltransferase KYP (KRYPTONITE), also known as SUVH4 [SU(var)3–9 homologue4] and its close homologs SUVH5/6 [3]. The KYP protein has an N-terminal SRA domain, which preferentially targets methylated CHH and CHG DNA, while its C-terminal SET domain is a H3K9 methyltransferase [77,78]. Thus, the KYP protein can link recognition of methylated DNA with methylation of histone H3K9. The eventual structure determination of KYP with bound methylated DNA and H3 peptide should shed light on the coupling of these two types of epigenetic marks.

6. Future challenges and opportunities

Starting from the first structural characterization of a histone modification reader, the P/CAF (P300/CBP-associated factor) bromodomain, and its recognition of acetyllysine marks [79],

biochemical and structural studies over the subsequent 15 years have built up a solid knowledge base related to recognition and regulation of specific epigenetic marks by various single domain reader modules. However, epigenetic regulation constitutes a more complicated system with interacting components, whereby the potential exists for multiple factors to function together, with interrelated regulation of individual components. Thus, research on structural studies of epigenetic regulation at the single domain level readout of isolated histone marks has been expanded in recent years to structural studies on the combinatorial readout by multi-domain proteins of multiple marks. As the basic functional unit, the nucleosome core particle has been crystallized both by itself and in complex with other proteins [80], thereby forming an ideal platform for further investigation of multivalent readout of epigenetic marks and the crosstalk between them by structural biology approaches. To date, crystal structures of complexes of several nucleosome–core particles have been reported, including those bound by LANA (Latency-Associated Nuclear Antigen), RCC1 (Regulator of chromosome condensation 1) and Sir3 (Silent information regulator 3) BAH domain [81–85]. However, no structure of a complex between the nucleosome core particle containing specific epigenetic modification(s) together with bound effector protein(s) has been reported to date. Further advances will require access to milligram scale production of purified reconstituted nucleosomes harboring various epigenetic marks and their combinations, which is achievable using either available intein-based coupling technology and/or the genetic installation of modified amino acids by genetic code expansion developed in the former case at both mononucleosome and nucleosome array levels (reviewed in [86]). Thus, we anticipate that expansion of ongoing structural studies currently underway on multivalent readout of combinations of histone marks together with extension of ongoing efforts from the histone peptide to the nucleosome level has the potential for providing new details of the crosstalk between epigenetic marks at the nucleosomal level. An additional challenge would involve the preparation and incorporation of methylated DNA into assembled nucleosomes, so as to further our current comprehension of multivalent readout by both histone and DNA marks and the crosstalk between them.

Acknowledgment

References

1. Luger K, Mader AW, Richmond RK, Sargent DF, Richmond TJ. Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature. 1997; 389:251–260. [PubMed: 9305837]

2. Jenuwein T, Allis CD. Translating the histone code. Science. 2001; 293:1074–1080. [PubMed: 11498575]

3. Law JA, Jacobsen SE. Establishing, maintaining and modifying DNA methylation patterns in plants and animals. Nat. Rev. Genet. 2010; 11:204–220. [PubMed: 20142834]

4. Liu Y, Zhang X, Blumenthal RM, Cheng X. A common mode of recognition for methylated CpG. Trends Biochem. Sci. 2013; 38:177–183. [PubMed: 23352388]

5. Feng S, Jacobsen SE, Reik W. Epigenetic reprogramming in plant and animal development. Science. 2010; 330:622–627. [PubMed: 21030646]

6. Fischle W, Wang Y, Allis CD. Binary switches and modification cassettes in histone biology and beyond. Nature. 2003; 425:475–479. [PubMed: 14523437]

7. Ruthenburg AJ, Li H, Patel DJ, Allis CD. Multivalent engagement of chromatin modifications by linked binding modules. Nat. Rev. Mol. Cell Biol. 2007; 8:983–994. [PubMed: 18037899]

8. Kouzarides T. Chromatin modifications and their function. Cell. 2007; 128:693–705. [PubMed: 17320507]

9. Taverna SD, Li H, Ruthenburg AJ, Allis CD, Patel DJ. How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. Nat. Struct. Mol. Biol. 2007; 14:1025–1040. [PubMed: 17984965]

10. Musselman CA, Lalonde ME, Cote J, Kutateladze TG. Perceiving the epigenetic landscape through histone readers. Nat. Struct. Mol. Biol. 2012; 19:1218–1227. [PubMed: 23211769]

11. Patel DJ, Wang Z. Readout of epigenetic modifications. Annu. Rev. Biochem. 2013; 82:81–118. [PubMed: 23642229]

12. Wang Z, Patel DJ. Combinatorial readout of dual histone modifications by paired chromatin-associated modules. J. Biol. Chem. 2011; 286:18363–18368. [PubMed: 21454653]

13. Jacobson RH, Ladurner AG, King DS, Tjian R. Structure and function of a human TAFII250 double bromodomain module. Science. 2000; 288:1422–1425. [PubMed: 10827952]

14. Sanchez R, Zhou MM. The role of human bromodomains in chromatin biology and gene transcription. Curr. Opin. Drug Discov. Devel. 2009; 12:659–665.

15. Li Y, Li H. Many keys to push: diversifying the 'readership' of plant homeodomain fingers. Acta Biochim. Biophys. Sin. (Shanghai). 2012; 44:28–39. [PubMed: 22194011]

16. Musselman CA, Kutateladze TG. Handpicking epigenetic marks with PHD fingers. Nucleic Acids Res. 2011; 39:9061–9071. [PubMed: 21813457]

17. Lange M, Kaynak B, Forster UB, Tonjes M, Fischer JJ, Grimm C, Schlesinger J, Just S, Dunkel I, Krueger T, Mebus S, Lehrach H, Lurz R, Gobom J, Rottbauer W, Abdelilah-Seyfried S, Sperling S. Regulation of muscle development by DPF3, a novel histone acetylation and methylation reader of the BAF chromatin remodeling complex. Genes Dev. 2008; 22:2370–2384. [PubMed: 18765789]

18. Zeng L, Zhang Q, Li S, Plotnikov AN, Walsh MJ, Zhou MM. Mechanism and regulation of acetylated histone binding by the tandem PHD finger of DPF3b. Nature. 2010; 466:258–262. [PubMed: 20613843]

19. Qiu Y, Liu L, Zhao C, Han C, Li F, Zhang J, Wang Y, Li G, Mei Y, Wu M, Wu J, Shi Y. Combinatorial readout of unmodified H3R2 and acetylated H3K14 by the tandem PHD finger of MOZ reveals a regulatory mechanism for HOXA9 transcription. Genes Dev. 2012; 26:1376–1391. [PubMed: 22713874]

20. Tsai WW, Wang Z, Yiu TT, Akdemir KC, Xia W, Winter S, Tsai CY, Shi X, Schwarzer D, Plunkett W, Aronow B, Gozani O, Fischle W, Hung MC, Patel DJ, Barton MC. TRIM24 links a non-canonical histone signature to breast cancer. Nature. 2010; 468:927–932. [PubMed: 21164480]

21. Xi Q, Wang Z, Zaromytidou AI, Zhang XH, Chow-Tsang LF, Liu JX, Kim H, Barlas A, Manova-Todorova K, Kaartinen V, Studer L, Mark W, Patel DJ, Massague J. A poised chromatin platform for TGF-beta access to master regulators. Cell. 2011; 147:1511–1524. [PubMed: 22196728]

22. Ruthenburg AJ, Li H, Milne TA, Dewell S, McGinty RK, Yuen M, Ueberheide B, Dou Y, Muir TW, Patel DJ, Allis CD. Recognition of a mononucleosomal histone modification pattern by BPTF via multivalent interactions. Cell. 2011; 145:692–706. [PubMed: 21596426]

23. Tsukiyama T, Wu C. Purification and properties of an ATP-dependent nucleosome remodeling factor. Cell. 1995; 83:1011–1020. [PubMed: 8521501]

24. Xiao H, Sandaltzopoulos R, Wang HM, Hamiche A, Ranallo R, Lee KM, Fu D, Wu C. Dual functions of largest NURF subunit NURF301 in nucleosome sliding and transcription factor interactions. Mol. Cell. 2001; 8:531–543. [PubMed: 11583616]

25. Narlikar GJ, Fan HY, Kingston RE. Cooperation between complexes that regulate chromatin structure and transcription. Cell. 2002; 108:475–487. [PubMed: 11909519]

26. Wysocka J, Swigut T, Xiao H, Milne TA, Kwon SY, Landry J, Kauer M, Tackett AJ, Chait BT, Badenhorst P, Wu C, Allis CD. A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. Nature. 2006; 442:86–90. [PubMed: 16728976]

27. Li H, Ilin S, Wang W, Duncan EM, Wysocka J, Allis CD, Patel DJ. Molecular basis for site-specific read-out of histone H3K4me3 by the BPTF PHD finger of NURF. Nature. 2006; 442:91–95. [PubMed: 16728978]

28. Muir TW. Semisynthesis of proteins by expressed protein ligation. Annu. Rev. Biochem. 2003; 72:249–289. [PubMed: 12626339]

29. Shogren-Knaak MA, Peterson CL. Creating designer histones by native chemical ligation. Methods Enzymol. 2004; 375:62–76. [PubMed: 14870659]

30. He XJ, Chen T, Zhu JK. Regulation and function of DNA methylation in plants and animals. Cell Res. 2011; 21:442–465. [PubMed: 21321601]

31. Lindroth AM, Cao X, Jackson JP, Zilberman D, McCallum CM, Henikoff S, Jacobsen SE. Requirement of CHROMOMETHYLASE3 for maintenance of CpXpG methylation. Science. 2001; 292:2077–2080. [PubMed: 11349138]

32. Bernatavichute YV, Zhang X, Cokus S, Pellegrini M, Jacobsen SE. Genome-wide association of histone H3 lysine nine methylation with CHG DNA methylation in *Arabidopsis thaliana*. PLoS One. 2008; 3:e3156. [PubMed: 18776934]

33. Blus BJ, Wiggins K, Khorasanizadeh S. Epigenetic virtues of chromodomains. Crit. Rev. Biochem. Mol. Biol. 2011; 46:507–526. [PubMed: 22023491]

34. Du J, Zhong X, Bernatavichute YV, Stroud H, Feng S, Caro E, Vashisht AA, Terragni J, Chin HG, Tu A, Hetzel J, Wohlschlegel JA, Pradhan S, Patel DJ, Jacobsen SE. Dual binding of chromomethylase domains to H3K9me2-containing nucleosomes directs DNA methylation in plants. Cell. 2012; 151:167–180. [PubMed: 23021223]

35. Lindroth AM, Shultis D, Jasencakova Z, Fuchs J, Johnson L, Schubert D, Patnaik D, Pradhan S, Goodrich J, Schubert I, Jenuwein T, Khorasanizadeh S, Jacobsen SE. Dual histone H3 methylation marks at lysines 9 and 27 required for interaction with CHROMOMETHYLASE3. EMBO J. 2004; 23:4286–4296. [PubMed: 15457214]

36. Stroud H, Do T, Du J, Zhong X, Feng S, Johnson L, Patel DJ, Jacobsen SE. Non-CG methylation patterns shape the epigenetic landscape in *Arabidopsis*. Nat. Struct. Mol. Biol. 2014; 21:64–72. [PubMed: 24336224]

37. Jia D, Jurkowska RZ, Zhang X, Jeltsch A, Cheng X. Structure of Dnmt3a bound to Dnmt3L suggests a model for de novo DNA methylation. Nature. 2007; 449:248–251. [PubMed: 17713477]

38. Song J, Rechkoblit O, Bestor TH, Patel DJ. Structure of DNMT1–DNA complex reveals a role for autoinhibition in maintenance DNA methylation. Science. 2011; 331:1036–1040. [PubMed: 21163962]

39. Mansfield RE, Musselman CA, Kwan AH, Oliver SS, Garske AL, Davrazou F, Denu JM, Kutateladze TG, Mackay JP. Plant homeodomain (PHD) fingers of CHD4 are histone H3-binding modules with preference for unmodified H3K4 and methylated H3K9. J. Biol. Chem. 2011; 286:11779–11791. [PubMed: 21278251]

40. Oliver SS, Musselman CA, Srinivasan R, Svaren JP, Kutateladze TG, Denu JM. Multivalent recognition of histone tails by the PHD fingers of CHD5. Biochemistry. 2012; 51:6534–6544. [PubMed: 22834704]

41. Baylin SB, Herman JG, Graff JR, Vertino PM, Issa JP. Alterations in DNA methylation: a fundamental aspect of neoplasia. Adv. Cancer Res. 1998; 72:141–196. [PubMed: 9338076]

42. Jones PA, Gonzalgo ML. Altered DNA methylation and genome instability: a new pathway to cancer? Proc. Natl. Acad. Sci. U. S. A. 1997; 94:2103–2105. [PubMed: 9122155]

43. Laird PW, Jaenisch R. The role of DNA methylation in cancer genetic and epigenetics. Annu. Rev. Genet. 1996; 30:441–464. [PubMed: 8982461]

44. Cedar H, Bergman Y. Linking DNA methylation and histone modification: patterns and paradigms. Nat. Rev. Genet. 2009; 10:295–304. [PubMed: 19308066]

45. Bostick M, Kim JK, Esteve PO, Clark A, Pradhan S, Jacobsen SE. UHRF1 plays a role in maintaining DNA methylation in mammalian cells. Science. 2007; 317:1760–1764. [PubMed: 17673620]

46. Sharif J, Muto M, Takebayashi S, Suetake I, Iwamatsu A, Endo TA, Shinga J, Mizutani-Koseki Y, Toyoda T, Okamura K, Tajima S, Mitsuya K, Okano M, Koseki H. The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. Nature. 2007; 450:908–912. [PubMed: 17994007]

47. Unoki M, Nishidate T, Nakamura Y. ICBP90, an E2F-1 target, recruits HDAC1 and binds to methyl-CpG through its SRA domain. Oncogene. 2004; 23:7601–7610. [PubMed: 15361834]

48. Arita K, Ariyoshi M, Tochio H, Nakamura Y, Shirakawa M. Recognition of hemimethylated DNA by the SRA protein UHRF1 by a base-flipping mechanism. Nature. 2008; 455:818–821. [PubMed: 18772891]

49. Avvakumov GV, Walker JR, Xue S, Li Y, Duan S, Bronner C, Arrowsmith CH, Dhe-Paganon S. Structural basis for recognition of hemi-methylated DNA by the SRA domain of human UHRF1. Nature. 2008; 455:822–825. [PubMed: 18772889]

50. Hashimoto H, Horton JR, Zhang X, Bostick M, Jacobsen SE, Cheng X. The SRA domain of UHRF1 flips 5-methylcytosine out of the DNA helix. Nature. 2008; 455:826–829. [PubMed: 18772888]

51. Rajakumara E, Wang Z, Ma H, Hu L, Chen H, Lin Y, Guo R, Wu F, Li H, Lan F, Shi YG, Xu Y, Patel DJ, Shi Y. PHD finger recognition of unmodified histone H3R2 links UHRF1 to regulation of euchromatic gene expression. Mol. Cell. 2011; 43:275–284. [PubMed: 21777816]

52. Hu L, Li Z, Wang P, Lin Y, Xu Y. Crystal structure of PHD domain of UHRF1 and insights into recognition of unmodified histone H3 arginine residue 2. Cell Res. 2011; 21:1374–1378. [PubMed: 21808300]

53. Wang C, Shen J, Yang Z, Chen P, Zhao B, Hu W, Lan W, Tong X, Wu H, Li G, Cao C. Structural basis for site-specific reading of unmodified R2 of histone H3 tail by UHRF1 PHD finger. Cell Res. 2011; 21:1379–1382. [PubMed: 21808299]

54. Nady N, Lemak A, Walker JR, Avvakumov GV, Kareta MS, Achour M, Xue S, Duan S, Allali-Hassani A, Zuo X, Wang YX, Bronner C, Chedin F, Arrowsmith CH, Dhe-Paganon S. Recognition of multivalent histone states associated with heterochromatin by UHRF1 protein. J. Biol. Chem. 2011; 286:24300–24311. [PubMed: 21489993]

55. Arita K, Isogai S, Oda T, Unoki M, Sugita K, Sekiyama N, Kuwata K, Hamamoto R, Tochio H, Sato M, Ariyoshi M, Shirakawa M. 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. 2012; 109:12950–12955. [PubMed: 22837395]

56. Cheng J, Yang Y, Fang J, Xiao J, Zhu T, Chen F, Wang P, Li Z, Yang H, Xu Y. 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. 2013; 288:1329–1339. [PubMed: 23161542]

57. Chodavarapu RK, Feng S, Bernatavichute YV, Chen PY, Stroud H, Yu Y, Hetzel JA, Kuo F, Kim J, Cokus SJ, Casero D, Bernal M, Huijser P, Clark AT, Kramer U, Merchant SS, Zhang X, Jacobsen SE, Pellegrini M. Relationship between nucleosome positioning and DNA methylation. Nature. 2010; 466:388–392. [PubMed: 20512117]

58. Ohm JE, McGarvey KM, Yu X, Cheng L, Schuebel KE, Cope L, Mohammad HP, Chen W, Daniel VC, Yu W, Berman DM, Jenuwein T, Pruitt K, Sharkis SJ, Watkins DN, Herman JG, Baylin SB. A stem cell-like chromatin pattern may predispose tumor suppressor genes to DNA hypermethylation and heritable silencing. Nat. Genet. 2007; 39:237–242. [PubMed: 17211412]

59. Schlesinger Y, Straussman R, Keshet I, Farkash S, Hecht M, Zimmerman J, Eden E, Yakhini Z, Ben-Shushan E, Reubinoff BE, Bergman Y, Simon I, Cedar H. Polycomb-mediated methylation on Lys27 of histone H3 pre-marks genes for de novo methylation in cancer. Nat. Genet. 2007; 39:232–236. [PubMed: 17200670]

60. Widschwendter M, Fiegl H, Egle D, Mueller-Holzner E, Spizzo G, Marth C, Weisenberger DJ, Campan M, Young J, Jacobs I, Laird PW. Epigenetic stem cell signature in cancer. Nat. Genet. 2007; 39:157–158. [PubMed: 17200673]

61. Kim D, Blus BJ, Chandra V, Huang P, Rastinejad F, Khorasanizadeh S. Corecognition of DNA and a methylated histone tail by the MSL3 chromodomain. Nat. Struct. Mol. Biol. 2010; 17:1027–1029. [PubMed: 20657587]

62. Alekseyenko AA, Peng S, Larschan E, Gorchakov AA, Lee OK, Kharchenko P, McGrath SD, Wang CI, Mardis ER, Park PJ, Kuroda MI. A sequence motif within chromatin entry sites directs MSL establishment on the Drosophila X chromosome. Cell. 2008; 134:599–609. [PubMed: 18724933]

63. Eidahl JO, Crowe BL, North JA, McKee CJ, Shkriabai N, Feng L, Plumb M, Graham RL, Gorelick RJ, Hess S, Poirier MG, Foster MP, Kvaratskhelia M. Structural basis for high-affinity binding of LEDGF PWWP to mononucleosomes. Nucleic Acids Res. 2013; 41:3924–3936. [PubMed: 23396443]

64. Chang PY, Hom RA, Musselman CA, Zhu L, Kuo A, Gozani O, Kutateladze TG, Cleary ML. Binding of the MLL PHD3 finger to histone H3K4me3 is required for MLL-dependent gene transcription. J. Mol. Biol. 2010; 400:137–144. [PubMed: 20452361]

65. Grow EJ, Wysocka J. Flipping MLL1's switch one proline at a time. Cell. 2010; 141:1108–1110. [PubMed: 20602992]

66. Wang Z, Song J, Milne TA, Wang GG, Li H, Allis CD, Patel DJ. Pro isomerization in MLL1 PHD3-bromo cassette connects H3K4me readout to CyP33 and HDAC-mediated repression. Cell. 2010; 141:1183–1194. [PubMed: 20541251]

67. Southall SM, Wong PS, Odho Z, Roe SM, Wilson JR. Structural basis for the requirement of additional factors for MLL1 SET domain activity and recognition of epigenetic marks. Mol. Cell. 2009; 33:181–191. [PubMed: 19187761]

68. Collins RE, Northrop JP, Horton JR, Lee DY, Zhang X, Stallcup MR, Cheng X. The ankyrin repeats of G9a and GLP histone methyltransferases are mono- and dimethyllysine binding modules. Nat. Struct. Mol. Biol. 2008; 15:245–250. [PubMed: 18264113]

69. Wu H, Min J, Lunin VV, Antoshenko T, Dombrovski L, Zeng H, Allali-Hassani A, Campagna-Slater V, Vedadi M, Arrowsmith CH, Plotnikov AN, Schapira M. Structural biology of human H3K9 methyltransferases. PLoS One. 2010; 5:e8570. [PubMed: 20084102]

70. Horton JR, Upadhyay AK, Qi HH, Zhang X, Shi Y, Cheng X. Enzymatic and structural insights for substrate specificity of a family of Jumonji histone lysine demethylases. Nat. Struct. Mol. Biol. 2010; 17:38–43. [PubMed: 20023638]

71. Lin H, Wang Y, Wang Y, Tian F, Pu P, Yu Y, Mao H, Yang Y, Wang P, Hu L, Lin Y, Liu Y, Xu Y, Chen CD. Coordinated regulation of active and repressive histone methylations by a dual-specificity histone demethylase ceKDM7A from *Caenorhabditis elegans*. Cell Res. 2010; 20:899–907. [PubMed: 20567262]

72. Yang Y, Hu L, Wang P, Hou H, Lin Y, Liu Y, Li Z, Gong R, Feng X, Zhou L, Zhang W, Dong Y, Yang H, Lin H, Wang Y, Chen CD, Xu Y. Structural insights into a dual-specificity histone demethylase ceKDM7A from *Caenorhabditis elegans*. Cell Res. 2010; 20:886–898. [PubMed: 20567261]

73. Ooi SK, Qiu C, Bernstein E, Li K, Jia D, Yang Z, Erdjument-Bromage H, Tempst P, Lin SP, Allis CD, Cheng X, Bestor TH. DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. Nature. 2007; 448:714–717. [PubMed: 17687327]

74. Otani J, Nankumo T, Arita K, Inamoto S, Ariyoshi M, Shirakawa M. Structural basis for recognition of H3K4methylation status by the DNA methyltransferase 3A ATRX–DNMT3–DNMT3L domain. EMBO Rep. 2009; 10:1235–1241. [PubMed: 19834512]

75. Cheng X, Blumenthal RM. Mammalian DNA methyltransferases: a structural perspective. Structure. 2008; 16:341–350. [PubMed: 18334209]

76. Li BZ, Huang Z, Cui QY, Song XH, Du L, Jeltsch A, Chen P, Li G, Li E, Xu GL. Histone tails regulate DNA methylation by allosterically activating de novo methyltransferase. Cell Res. 2011; 21:1172–1181. [PubMed: 21606950]

77. Johnson LM, Bostick M, Zhang X, Kraft E, Henderson I, Callis J, Jacobsen SE. The SRAmethyl-cytosine-binding domain links DNA and histonemethylation. Curr. Biol. 2007; 17:379–384. [PubMed: 17239600]

78. Jackson JP, Lindroth AM, Cao X, Jacobsen SE. Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase. Nature. 2002; 416:556–560. [PubMed: 11898023]

79. Dhalluin C, Carlson JE, Zeng L, He C, Aggarwal AK, Zhou MM. Structure and ligand of a histone acetyltransferase bromodomain. Nature. 1999; 399:491–496. [PubMed: 10365964]

80. Tan S, Davey CA. Nucleosome structural studies. Curr. Opin. Struct. Biol. 2011; 21:128–136. [PubMed: 21176878]

81. Barbera AJ, Chodaparambil JV, Kelley-Clarke B, Joukov V, Walter JC, Luger K, Kaye KM. The nucleosomal surface as a docking station for Kaposi's sarcoma herpes-virus LANA. Science. 2006; 311:856–861. [PubMed: 16469929]

82. Makde RD, England JR, Yennawar HP, Tan S. Structure of RCC1 chromatin factor bound to the nucleosome core particle. Nature. 2010; 467:562–566. [PubMed: 20739938]

83. Armache KJ, Garlick JD, Canzio D, Narlikar GJ, Kingston RE. Structural basis of silencing: Sir3 BAH domain in complex with a nucleosome at 3.0 A resolution. Science. 2011; 334:977–982. [PubMed: 22096199]

84. Yang D, Fang Q, Wang M, Ren R, Wang H, He M, Sun Y, Yang N, Xu RM. Nalphaacetylated Sir3 stabilizes the conformation of a nucleosome-binding loop in the BAH domain. Nat. Struct. Mol. Biol. 2013; 20:1116–1118. [PubMed: 23934152]

85. Wang F, Li G, Altaf M, Lu C, Currie MA, Johnson A, Moazed D. Heterochromatin protein Sir3 induces contacts between the amino terminus of histone H4 and nucleosomal DNA. Proc. Natl. Acad. Sci. U. S. A. 2013; 110:8495–8500. [PubMed: 23650358]

86. Chatterjee C, Muir TW. Chemical approaches for studying histone modifications. J. Biol. Chem. 2010; 285:11045–11050. [PubMed: 20147749]
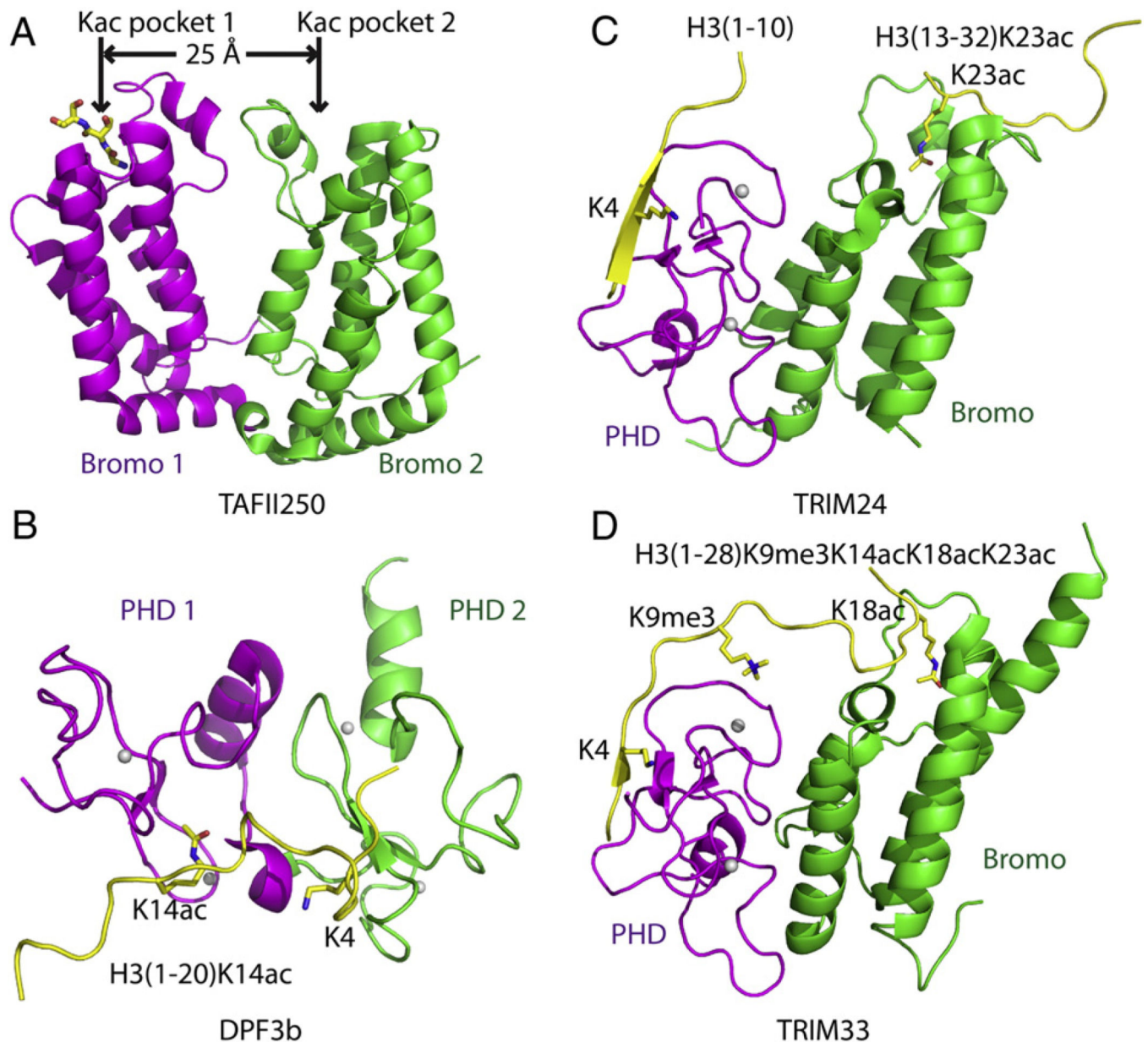
Fig. 1.
Structural basis for multivalent readout of histone marks from a single histone tail. (A) Ribbon representation of the crystal structure of TAF$_{II}$250 double bromodomain (PDB code: 1EQF) with bromodomain 1 colored in magenta and bromodomain 2 in green. The N-terminus of a symmetry related protein inserts into the acetyllysine (Kac) binding pocket of bromodomain 1. The distance between two Kac binding pockets of the two bromodomains is 25 Å. (B) Ribbon representation of the solution NMR structure of DPF3b double PHD fingers in complex with H3(1–20)K14ac peptide (PDB code: 2KWJ)with PHD1 finger colored in magenta, the PHD2 finger in green, and the bound peptide in yellow. The zinc ions are shown as silver balls. The specific residues recognized on the H3 peptide, including K4 and K14ac, are highlighted in stick representations. (C) Ribbon-representation model of

TRIM24 PHD-Bromo cassette simultaneously recognizing unmodified H3K4 and H3K23ac following superposition of the crystal structures of TRIM24 PHD-Bromo-H3(1–10) complex (PDB code: 3O37) and TRIM24 PHD-Bromo-H3(13–32)K23ac complex (PDB code: 3O34). The PHD finger, bromodomain and bound peptides are colored in magenta, green and yellow, respectively. The unmodified H3K4 and H3K23ac are highlighted in stick representations. (D) Ribbon representation of the crystal structure of TRIM33 PHD-Bromo cassette in complex with H3(1–28)K9me3/K14ac/K18ac/K23ac peptide (PDB code: 3U5O and 3U5P) with the PHD finger, bromodomain, and bound peptide colored in magenta, green and yellow, respectively. The unmodified H3K4, H3K9me3 and H3K18ac, which are specifically recognized, are highlighted in stick representations.
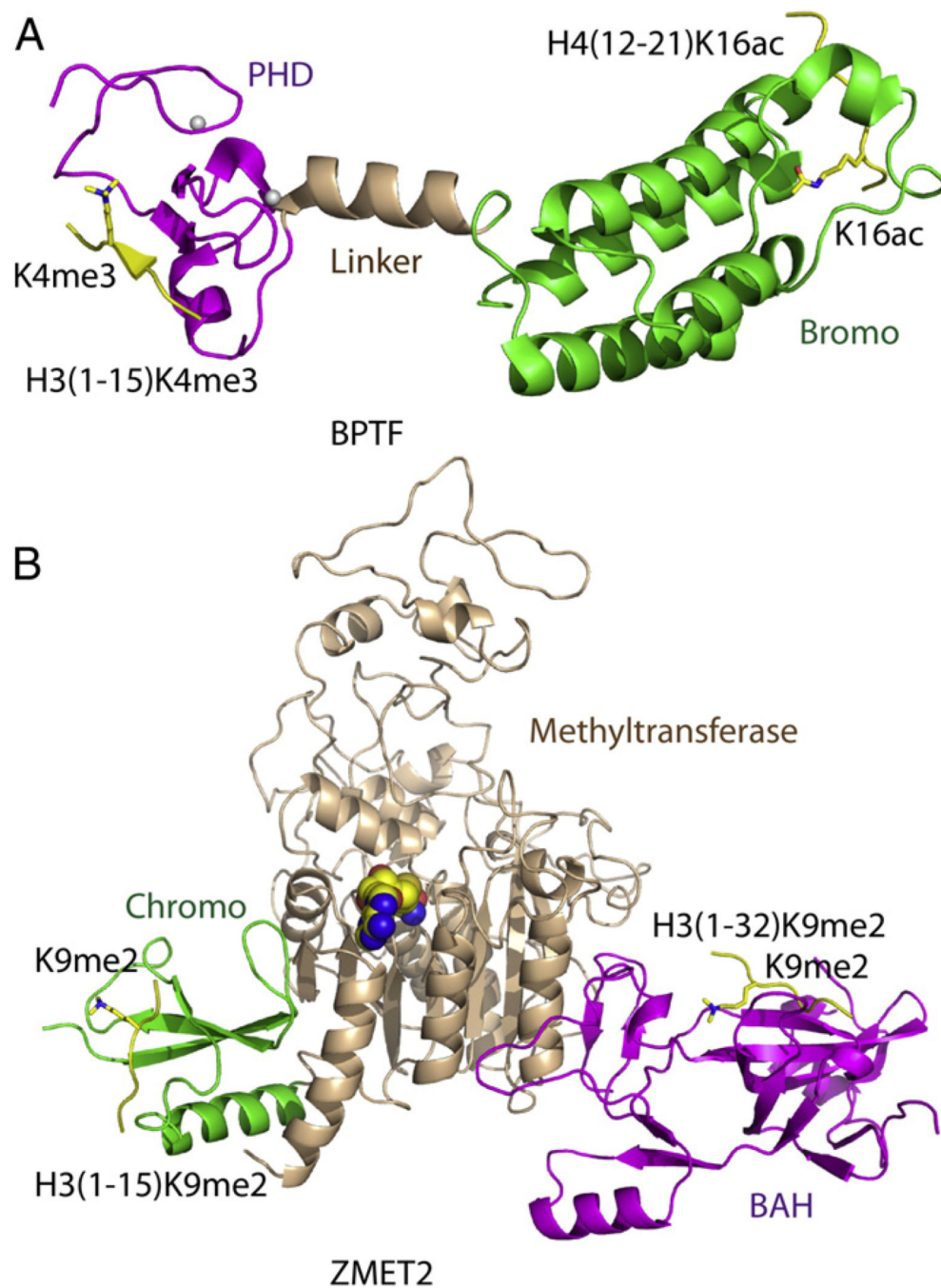
Fig. 2.
Structural basis for multivalent readout of multiple histone tails. (A) Ribbon-representation of a model of the BPTF PHD-Bromo cassette simultaneously recognizing H3K4me3 and H4K16ac following superposition of the crystal structures of BPTF PHDBromo-H3(1–15)K4me3 complex (PDB code: 2F6J) and BPTF PHD-Bromo-H4(12–21) K16ac complex (PDB code: 3QZS). The PHD finger, the bromodomain, the linker region and the bound peptides are colored in magenta, green, wheat and yellow, respectively. TheH3K4me3 and H4K16ac are highlighted in stick representations. (B) Ribbon representation of a model of

ZMET2 simultaneously recognizing two H3K9me2 peptides with its BAH domain and chromodomain by superposition of the crystal structures of ZMET2-H3(1–15)K9me2 complex (PDB code: 4FT2) and ZMET2-H3(1–32)K9me2 complex (PDB code: 4FT4). The BAH domain, the methyltransferase, the chromodomain and the bound peptides are colored in magenta, wheat, green and yellow, respectively. The two H3K9me2 marks are highlighted in stick representations.
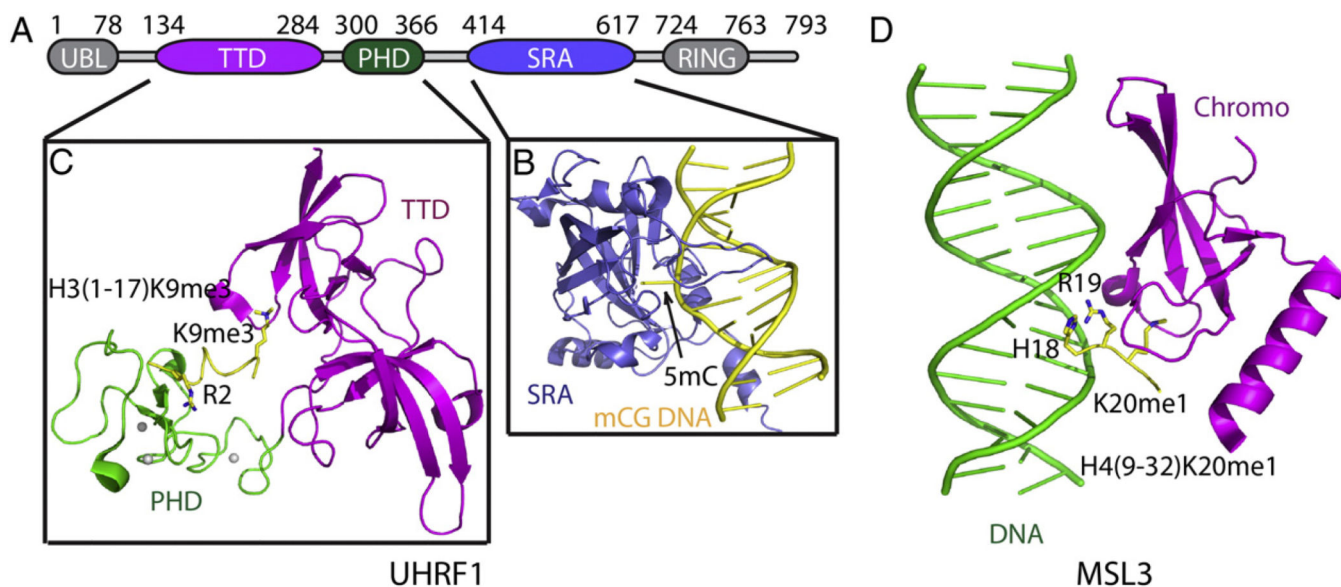
Fig. 3.
Structural basis for multivalent readout of histone and DNA marks. (A) The domain architecture of multiple domain protein UHRF1, which can recognize both histone marks and hemimethylated CpG DNA. (B) Ribbon representation of the structure of the UHRF1 SRA domain in complex with a hemimethylated CpG DNA with the SRA domain (PDB code: 3CLZ) and the DNA colored in slate and yellow, respectively. The 5mC base is flipped out of the DNA helix and indicated by an arrow. (C) Ribbon representation of the crystal structure of UHRF1 tandem Tudor domain (TTD)-PHD cassette in complex with the H3(1–17)K9me3 peptide (PDB code: 4GY5). The TTD domain, the PHD finger domain and the bound peptide are colored in magenta, green and yellow, respectively. The specifically recognized unmodified H3R2 and H3K9me3 residues are highlighted in stickmodel. (D) Ribbon representation of the crystal structure of MSL3 chromodomain in complex with a DNA duplex and H4(9–32)K20me1 peptide (PDB code: 3OA6). The chromodomain, DNA and peptide are colored in magenta, green and yellow, respectively. The peptide residues H18 and R19, which interact with DNA, and H20me1, which is specifically recognized by the chromodomain, are highlighted in stick representations.
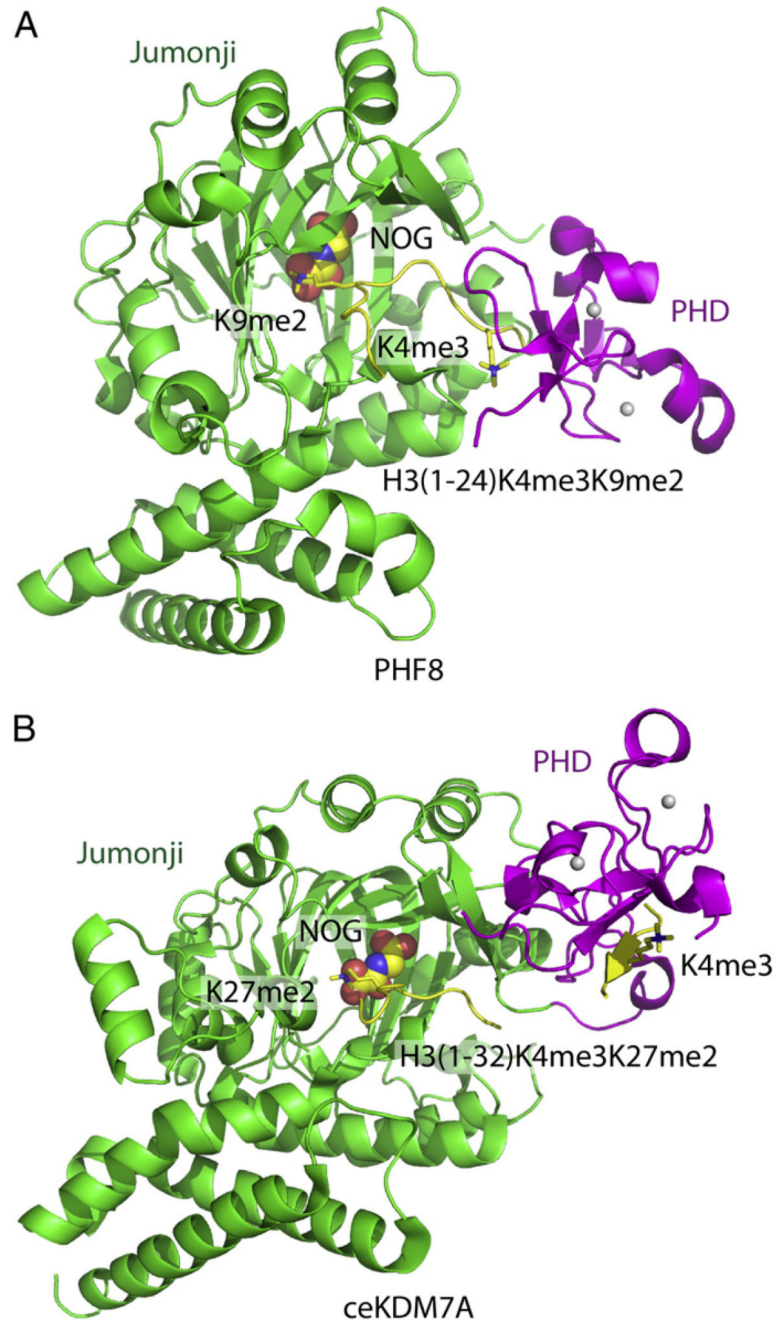
Fig. 4.
Structural basis for recognition by histone modification-directed histone modification enzyme. (A) Ribbon representation of the crystal structure of PHF8 in complex with H3(1–24)K4me3/K9me2 peptide (PDB code: 3KV4). The PHD finger, Jumonji domain and the bound peptide are colored in magenta, green and yellow, respectively. The NOG cofactor is shown in a space-filling representation. The H3K4m3mark, which is specifically recognized by the PHD finger, and the H3K9me2 mark, which specifically inserts into the active site of the Jumonji domain, are highlighted in stick representations. The PHF8 enzyme shows a

bent conformation upon K4me3 binding into the PHD pocket, with the distance between the PHD pocket and Jumonji active site short enough to allow the H3K9me2 mark to simultaneously insert into the Jumonji active site. (B) Ribbon representation of the crystal structure of ceKDM7A in complex with H3(1–32)K4me3/K27me2 peptide (PDB code: 3N9P). The PHD finger, Jumonji domain and the bound peptide are colored in magenta, green and yellow, respectively. The NOG cofactor is shown in a space filling representation. The H3K4me3 mark, which is specifically recognized by the PHD finger, and the H3K27me2 mark, which specifically inserts into the active site of the Jumonji domain, are highlighted in stick representation. The ceKDM7A enzyme shows an extended conformation upon K4me3 binding into the PHD pocket, with the distance between the PHD pocket and Jumonji active site being too long to allow simultaneous recognition of H3K9me2, but should allow simultaneous recognition of H3K27me2 mark by inserting it into the Jumonji active site.