Published in final edited form as:

Biochim Biophys Acta. 2014 August ; 1839(8): 686–693. doi:10.1016/j.bbagr.2014.04.001.

Towards understanding methyllysine readout[☆]

Catherine A. Musselman^{a,*}, Sepideh Khorasanizadeh^b, and Tatiana G. Kutateladze^c

^aDepartment of Biochemistry, University of Iowa, Iowa City, IA 52242, USA

^bDiabetes and Obesity Research Center, Sanford–Burnham Medical Research Institute, Orlando, FL 32827, USA

^cDepartment of Pharmacology, University of Colorado School of Medicine, Aurora, CO 80045, USA

Abstract

Lysine methylation is the most versatile covalent posttranslational modification (PTM) found in histones and non-histone proteins. Over the past decade a number of methyllysine-specific readers have been discovered and their interactions with histone tails have been structurally and biochemically characterized. More recently innovative experimental approaches have emerged that allow for studying reader interactions in the context of the full nucleosome and nucleosomal arrays. New studies reveal various reader-nucleosome contacts outside the methylated histone tail, thus offering a better model for the association of histone readers to chromatin and broadening our understanding of the functional implications of these interactions. In this review we give a brief overview of the known mechanisms of histone lysine methylation readout, summarize progress recently made in exploring interactions with methylated nucleosomes, and discuss the latest advances in the development of small molecule inhibitors of the methyllysine-specific readers.

Keywords

Histone; Methyllysine; Nucleosome; Histone reader; Inhibitor

1. Introduction

Lysine methylation was initially discovered as a post-translational modification (PTM) on histones in 1964 [1,2]. It is perhaps the most versatile of the histone PTMs and can exist in three states: mono-methyl (me1), di-methyl (me2) and tri-methyl (me3). Lysine residues that are known to be methylated include H3K4, H3K9, H3K27, H3K36, H3K79, H4K20, H1K26, H2BK5, and H2AK36 (Fig. 1). Lysine methylation is a reversible mark, which is placed by lysine methyltransferases (KMTs) and removed by lysine demethylases (KDMs) [3]. Specific lysine methylation patterns are commonly associated with certain chromatin states and genomic elements, and are linked to distinct biological outcomes such as transcription activation or repression [3–6].

[☆]This article is part of a Special Issue entitled: Molecular mechanisms of histone modification function.

© 2014 Published by Elsevier B.V.

*Corresponding author. catherine-musselman@uiowa.edu (C.A. Musselman).

Histone PTMs can directly alter chromatin structure or act as binding targets for nuclear proteins. Little is known about the effect of lysine methylation on chromatin structure itself, however a number of methyllysine-binding domains, or readers of this PTM, have been identified since 2001, when a chromodomain (CD) of HP1 was found to recognize histone H3K9me3 [7–10]. The structural basis of this recognition was characterized shortly thereafter [11,12]. To date, methylated lysines on histone tails appear to be targeted by the largest and most diverse set of readers. This includes ADD (ATRX-DNMT3-DNMT3L), ankyrin, BAH (bromo adjacent homology), chromo-barrel, chromodomain (CD), double chromodomain (DCD), HEAT, MBT (malignant brain tumor), PHD (plant homeodomain), PWWP, SAWADEE, tandem Tudor domain (TTD), Tudor, WD40 and zf-CW (zinc finger CW), see Table 1. The interaction of these domains with their target methyllysine in histone tails has been found to stabilize association of co-factors to chromatin for a variety of functions. Most notably these domains are found in proteins involved in transcriptional regulation, however they have also been characterized in RNA splicing, and DNA replication, recombination and repair (reviewed in [13,14]).

In-depth biochemical and structural studies reveal a well-conserved mechanism of methyllysine recognition. This mechanism and the activities of methyllysine-specific readers have been extensively reviewed [13,14] and thus are only briefly summarized below. We note however that the majority of the structural and mechanistic details have been obtained through studying interactions of the reader domain with a chemically synthesized peptide representing a modified histone tail. Although this approach has helped generate critical information regarding the specificity of recognition, it leaves a large gap in our knowledge of how readers bind to histone PTMs within the framework of the complete nucleosome. It also significantly limits our ability to tackle the question of how reader domains interpret multiple PTMs, which is essential in understanding chromatin function, structure and dynamics.

Alterations in histone methylation patterns are associated with a wide variety of diseases due to aberrant gene expression patterns as well as genomic instability, and a number of examples of methyllysine readers are implicated in disease [15–18]. Thus, investigating the binding of these readers to methylated nucleosomes is not only important for elucidating fundamental mechanisms of chromatin regulation, but also necessary for the design of targeted therapeutics.

In this review we summarize the known methyllysine binding mechanisms, focusing on recent studies that explore the methyllysine readout in the context of the full nucleosome. We discuss determinants of recognition outside the histone tail that mediate or alter affinity and specificity. We also highlight the current efforts in the design of chemical probes and antagonists of methyllysine interactions.

2. Diverse family of methyllysine readers

The family of methyllysine readers is expanding rapidly and currently consists of at least 15 members. Some methyllysine readers, including PHD and zf-CW, show a high degree of sequence specificity, while others, including MBT and WD40, are more promiscuous

though they can select for a certain methylation state of a target lysine. Despite the wide variety of the readers and histone targets, the majority of these domains have comparable binding affinities, with dissociation constants of the complexes being in the high nanomolar to low micromolar range (reviewed in [13,14]).

The methyllysine histone sequences are recognized by a conserved mechanism, in which the side chain of the methylated lysine inserts into an aromatic cage of the reader domain (reviewed in [13,14]). Typically, this cage contains two to four aromatic residues whose aromatic moieties are engaged in cation- π and van der Waals interactions with the methylammonium group of the modified lysine.

The mono-, di-, or trimethylated state of lysine is selected for by the composition and size of the aromatic cage. A reader prefers mono- and dimethylated lysines over the trimethylated species when a negatively charged residue is also present in the cage. The carboxylic group of an aspartate residue can make additional and favorable hydrogen bonding contacts with di- and mono- but not with trimethylated lysine. A small aromatic cage can also exclude binding of a higher methylation state due to steric hindrance, whereas a larger pocket prefers a higher methylation state as necessary contacts are only possible with the bulkier trimethylammonium group. Beyond caging of the Kme, the mechanism of recognition of surrounding residues often involves hydrogen bonding and hydrophobic contacts that afford sequence specificity and regulation. These contacts also lead to sensitivity to the post-translational modification status of surrounding residues, which may reinforce or disrupt the recognition of a given methyllysine. For example, the phosphorylation of H3S10 inhibits binding of the HP1 CD to H3K9me3 [19] and similarly phosphorylation of T3 decreases association of the CHD1 DCD and the PHD finger of TAF3, and MLL5 and Dido for H3K4me3 [20–23]. Similar types of cross-talk are seen between H3R2 methylation and H3K4me3 [24–27].

Notable variations in the aromatic cage binding mechanism have been reported recently. The ATRX ADD domain and the PHD fingers of CHD4 and TRIM33 are examples of methyllysine readers lacking the aromatic cage [28–33]. The side chain of K9me3 in the H3K9me3-bound ATRX ADD domain inserts between two parts of the domain (a zinc-knuckle and an adjacent PHD finger) and is uniquely coordinated through hydrophobic and cation- π contacts with a single aromatic residue and a set of nonconventional carbon-oxygen hydrogen bonds [28, 29]. The PHD finger of TRIM33 associates with the A1–S10 residues of H3K9me3K14acK18ac peptide [33]. The trimethylammonium group of K9 makes a nonconventional carbon-oxygen hydrogen bond with a carbonyl oxygen of the PHD finger and is also involved in a cation- π interaction with a tryptophan residue. Similarly, cation- π and hydrophobic interactions between a phenylalanine and K9me3 stabilize the CHD4 PHD2–H3K9me3 complex [31,32]. A single tryptophan residue is also seen in the well-defined K4me3-binding cage of the MLL5 PHD finger and, even though there is an aspartate opposite to the tryptophan in the aromatic cage, this PHD finger appears to prefer trimethylated species [22].

3. Methyllysine readout in the context of the nucleosome

It becomes increasingly clear that to fully understand the methyllysine reading mechanism, it is essential to study these interactions in the context of nucleosomes and nucleosome arrays. Methylated nucleosomes can be either generated through purification out of various cellular systems or constructed *in vitro* using recombinant histones and DNA [34]. The latter method allows for greater control of the nucleosome composition, however histones expressed in *Escherichia coli* contain no PTMs. In contrast, those purified out of cellular systems are post-translationally modified but there is little control over the extent and content of their PTMs or associated DNA sequences. Though particular set knockout allows for some manipulation of the levels of methylation in nucleosomes purified from yeast, these approaches are still limited as they do not allow for the necessary control of other histone PTMs. Treatment of the recombinant systems with histone methyltransferases often leads to low levels of methylation and is not always specific. Thus, the challenge has been to generate systems with well-defined and homogenous levels of modification that would allow for investigation of the methyllysine recognition in a more physiologically relevant context. Recently, several such methods have been developed and implemented, and the results of these studies provide tremendous insight into the mechanisms and functional consequence of methyllysine readout.

4. Incorporation of methylation marks into nucleosomes

Over the past decade, three main approaches for the incorporation of methyllysine into recombinant histones have been developed. These are chemical ligation, installation of methyllysine analogues, and genetic installation (Fig. 2). All three methods have their advantages and disadvantages as discussed below.

The semi-synthetic approach incorporates a true modification (*i.e.* not an analogue) at high levels of homogeneity into full-length histones. Native chemical ligation (NCL) [35] allows for coupling of a synthesized modified histone tail peptide carrying a carboxy-terminal thioester to a recombinant histone core containing an amino-terminal cysteine. In the expressed protein ligation (EPL) method, a recombinant histone core with a thioester carboxy-terminus is ligated to a modified histone tail peptide with an amino-terminal cysteine [36–38] (Fig. 2A). There are two drawbacks to this approach. One is the incorporation of a cysteine, though it can be desulfurized with hydrogen/Raney nickel to an alanine. In addition this method can be cost prohibitive if large amounts of the modified histone are needed. These methods have been more extensively used to generate acetylated and phosphorylated histones, but a few studies have utilized them in the investigation of methyllysine binding.

Installation of a methyllysine analogue (MLA) is an alternative approach in which the lysine of interest is mutated to a cysteine residue. Subsequent treatment with an alkylation reagent produces a lysine analogue, which can be mono-, di-, or trimethylated [39] (Fig. 2B). This method yields high levels of homogeneously modified histone and is very cost effective. However, the MLA contains sulfur at the γ position, which leads to a 0.28 Å increase in side-chain length as well as an increased acidity. Analysis of the robustness of the analogues

has yielded mixed results. Overall they appear to be reasonable mimics of methyllysine, demonstrating the ability to be methylated, demethylated and recognized by effector domains [40]. However, there is also evidence that they display lower activity as compared to native methyllysine [41, 42]. Thus, additional controls are necessary to confirm activity of the MLA. This method has become the most popular approach in producing methylated histones.

In the most recently developed approach methylated histones are generated through genetic incorporation in a bacterial system. By using an orthogonal pyrrolysyl-tRNA synthetase and tRNA_{CUA} an N3-tert-butyloxycarbonyl-N3-methyl-L-lysine can be genetically installed at sites of “amber” codons. The tert-butyloxycarbonyl group can then be removed to produce monomethylated lysine [43] (Fig. 2C). This method can be further expanded to generate dimethylated lysine [44]. Though this approach provides homogenous levels of modification, protein yield can be significantly impaired, and it has yet to be extended to trimethylated lysine species.

5. Contacts outside the methylated histone tail

It has been noticed that interactions between some readers and their methylated targets appear to be quite weak or non-specific when investigated using histone peptide substrates, and K_d values in the millimolar range have been reported [45,46]. However this could be rationalized if contacts outside the histone tail are also necessary for the binding affinities to reach a physiologically relevant range. An early example is the MSL3 CD, which is a reader that efficiently recognizes double strand DNA and relies on its interaction with DNA to selectively interact with the H4K20me1 peptide [47]. Another example is the PSIP1/LEDGF PWWP domain which not only recognizes H3K36me3 but also binds double stranded DNA (dsDNA) [48,49]. Binding to the individual components is relatively weak (millimolar K_d s for the methylated peptide and ~150 μ M for free dsDNA), however the bimodal interaction increases the avidity of PWWP for an H3K36me3-containing nucleosome to ~10 [4] fold. Interestingly, though the interaction is clearly energetically more driven by contacts with the DNA, the specificity for the methylated nucleosome is apparent, as interaction with the unmethylated nucleosome is significantly weaker. Such a dramatic increase in affinity suggests that there are likely additional structural rearrangements induced in the context of the nucleosome that also enhance binding.

The Iswb1 remodeling complex subunit Ioc4 binds H3K36me3-nucleosomes but not unmodified nucleosomes. Importantly, the Ioc4–H3K36me3 interaction cannot be detected when tested with methylated peptides alone [50,51], and a similar result is observed for the association of the Pdp1 PWWP domain with H4K20me3-nucleosomes [52]. A recent study of the PHF1 Tudor domain also reveals multivalent contacts with H3K36me3 and nucleosomal DNA [53]. In contrast to the PWWP domain this association is far more driven by interaction with the methylated mark itself [53].

For some interactions, regions adjacent to the canonical reader are also necessary. The specific interaction of the SCLM2 MBT repeats with the H2AK36me1-nucleosomes is not detected by electromobility shift assays, however it can be monitored once a DNA-binding

region adjacent to the MBT repeats is included [54]. Likewise the WD40 domain of EED only binds H3K27me3-nucleosomes well when an N-terminal H3-binding region is present [55,56]. Weak interactions with nucleosomal DNA have also been seen between the hinge region and chromo-shadow domain of HP1 α [57], HP1 β [58] and the *Schizosaccharomyces pombe* HP1 homolog Swi6, important in the H3K9me3 targeting by the protein's CD [59]. Together these studies demonstrate that for some readers contacts outside the histone tail are essential in the readout of methyl marks (Fig. 3A).

6. Multivalent interactions

Many chromatin-associating proteins contain more than one reader domain and/or exist in a complex with additional readers. This opens up the possibility of multivalent interactions that could lend specificity and affinity required for targeting specific regions of chromatin. The combinatorial readout of PTMs by multiple domains has long been proposed [60,61] but is difficult to characterize. Though peptides with multiple PTMs can be very useful in studying the readout of multiple PTMs, it is likely that the inter-tail orientation afforded by the nucleosome structure and inter-nucleosome orientation will play a significant role in the mechanism and functional consequence of these interactions.

Several recent studies have demonstrated the power of multivalent interactions in the context of the nucleosome. For example, though the bromodomain of BPTF recognizes various acetylated histone H4 tail peptides, the PHD-bromodomain cassette of BPTF specifically recognizes the H3K4me3/H4K16ac pattern of PTMs present on the same nucleosome [62]. In contrast, a trans-nucleosomal multivalent interaction is proposed for the Rpd3s histone deacetylase complex, with the CD of the Eaf3 subunit binding to H3K36me3 on one nucleosome and a PHD finger of the Rco1 subunit binding to H3K36me3 on an adjacent nucleosome [63] (Fig. 3A). The TFIID complex binds stronger to nucleosomes containing both H3K4me3 and either H3K9ac or H3K14ac marks through the TAF3 PHD finger and TAF1 bromodomains, than to singly modified nucleosomes. This interaction can be further augmented by incorporation of a consensus TATA sequence in the modified nucleosomes [64].

7. Towards a chromatin model

It is now well established that methyllysine can be recognized by a number of domains, however the functional implications of these associations beyond simple targeting remain poorly understood. A few recent reports have examined the mechanisms and consequences of the methyllysine readout in a manner that leads to more relevant model of functional outcomes in the context of chromatin. These include two prominent examples delineating heterochromatin formation in the yeast *S. pombe* where Swi6 (HP1 homolog) and Clr4 (Suv39 homolog) proteins coordinate their functions along nucleosomes and use their CDs to read H3K9me3.

Canzio and coworkers found that the combined action of recognition of H3K9me3 by the CD of Swi6 and Swi6 dimerization via its chromo-shadow domain (CSD) can bridge nucleosomes, inducing an inter-nucleosome structure that is essential in the formation of heterochromatin [65]. Built upon the initial discovery, this research team further

demonstrated that binding of Swi6 to methylated nucleosomes drives a switch from an auto-inhibited state to a spreading-competent state [59]. The auto-inhibited state is a dimer in which the methyllysine binding site in one Swi6 molecule is blocked through interaction with a histone-mimicking loop (consisting of an ARK motif) of the CD of the associated Swi6 molecule. In contrast to what has been observed for Swi6 the interaction of the mammalian HP1 with the H3K9me3-nucleosome appears overall to be very dynamic [58], indicating that additional factors are needed for spreading. One potential mode of regulation may be through post-translational modifications of the HP1 protein itself to signal formation of a spreading-competent HP1. In support of this mechanism, the histone-mimicking loop in a human HP1 CD contains a lysine that can become methylated [66], thus potentially stabilizing the interaction between two HP1 CDs.

The H3K9 methyltransferase Suv39, which together with Swi6 mediates chromatin spread in *S. pombe*, was also investigated in the context of methylated nucleosomes [67]. Binding of the CD of Suv39 to pre-existing H3K9me marks stimulates catalytic activity of Suv39 on an adjacent nucleosome. Furthermore, the combined action of Swi6 and Suv39 relies on the ability of the chromodomains of these two proteins to distinguish the methylation states of H3K9, and thus reduces competition. While the CD of Suv39 prefers the trimethylated state of the lysine, the Swi6 CD binds both H3K9me3 and more prevalent H3K9me2.

Several functional studies have implicated recognition of methylated histones in the regulation of chromatin acting enzymatic complexes, in some instances mediating crosstalk between PTMs. For instance the H3 acetyltransferase activity of the HBO1 complex is increased upon binding of the ING4 subunit's PHD finger to H3K4me3 [68]. Recognition of H3K27me3 by the EED WD40 domain enhances methyltransferase activity of the H3K27-specific PRC2 complex [69], whereas binding of the PHF1 Tudor domain to H3K36me3 inhibits this activity [70]. The latter interaction has been shown to alter the structural dynamics of the H3K36me3-containing nucleosome, stabilizing a more open form of the particle, in which the nucleosomal DNA is more accessible to DNA-binding regulatory proteins [53].

8. Chemical probes and inhibitors

In the human proteome, there are several dozens of predicted methyllysine reader domains, and many of them are expected to play roles in disease. Therefore, identification of chemical small molecule or peptide probes for methyllysine readers has become a priority both for interrogating their biological functions, as well as facilitating the design of targeted therapeutics [71–74]. As described below recent advances have indicated the feasibility of discovering diverse tools for interrogating the biological functions of methyllysine readers. Whether these could be effective *in vivo*, and be applicable to any targeted therapy remains unknown.

The features of the recognition pockets of 46 independent methyllysine readers were evaluated for the ability to bind with high affinity to passively absorb small molecules, and a predictive “druggability” score was proposed [75]. The study concluded that selected methyllysine readers within each class display enough variations in volume, enclosure and

hydrophobicity of their binding pockets to suggest feasibility of chemical probe discovery. Indeed using a newly developed HaloTag assay, Wagner and co-workers identified three inhibitors of the JARID1A PHD3 finger association with H3K4me3. Chemical modification of one of these inhibitors led to a 10-fold increase in potency, and interestingly the antagonistic action was found to be independent of the methyllysine binding aromatic pocket itself [76].

Members of the MBT family that bind to methyllysine via a cavity insertion mechanism have been particularly amenable to chemical probe discovery. A high throughput screening method was introduced based on Amplified Luminescence Proximity Homogeneous Assay (AlphaScreen) technology and validated for screening inhibitors of low-affinity peptide–protein interactions with application to MBT containing proteins [77,78]. This method proved reliable in miniaturized and automated library screening resulting in Z' score of 0.87 and identifying a group of inhibitors with half maximal inhibitory concentration (IC_{50}) in the nanomolar range [78]. In addition a virtual screening strategy has identified distinct chemotypes as potential MBT antagonists from a large database of commercially available compounds [79].

A structure-based strategy of ligand search for the MBT family helped to design peptide-like and related small molecule structures, such as nicotinamide derivatives which bind tighter than the methylated peptide [80]. The selectivity for L3MBTL1 was confirmed by comparing the IC_{50} in a panel of various readers. A co-crystal structure of L3MBTL1 with UNC669, which binds with a K_d of 5 μ M, shows that the pyrrolidiny ring acts as a methylated lysine mimic that inserts in the methyllysine binding pocket and is stabilized by hydrogen bonding, salt bridges and hydrophobic interactions [80].

Subsequent synthetic chemistry efforts led to the development of a high-affinity, high-quality probe for another MBT family member. The recent report elegantly links *in vitro* potency and selectivity with cell-based function of the L3MBTL3 protein [81]. A highly potent inhibitor for L3MBTL3 was identified using dibasic small molecules that contain two Kme mimics. Mutagenesis analysis and ITC measurements confirmed that compound UNC1215 binds L3MBTL3 robustly (K_d = 120 nM) within the original predicted cavity. A co-crystal structure of L3MBTL3 with UNC1215 reveals that two UNC1215 molecules bind an L3MBTL3 dimer and effectively disable all L3MBTL3 surfaces necessary for methylated peptide binding (Fig. 4). Cell-based assays demonstrate that the UNC1215 readily penetrates cells and colocalizes with L3MBTL3 in the nucleus [81].

UNC1215 was also utilized to identify novel interacting partners of L3MBTL3 [81] through mass spectrometry analysis of L3MBTL3 associated proteins which were pulled down from cells in the presence or absence of the chemical probe. The authors found that BCLAF1, once methylated at K445, interacts with L3MBTL3 and this interaction is inhibited by UNC1215 [81]. Improvement of the L3MBTL3 chemical probe (UNC1215) by designing novel analogues and evaluating their structure activity relationships has led to the identification of another high-quality chemical probe UNC1679 with improved selectivity against the closely related L3MBTL1 protein [82].

Very recently, peptide chromodomain antagonists with ~200 nM potency were reported to inhibit chromobox homolog 7 (CBX7) CD interaction with H3K27me3, suggesting that suitable potency and selectivity can be achieved for the inhibition of chromodomains [83]. Furthermore, an alternative mode of inhibition has recently been examined that capitalizes on neutralizing the signaling of a methyllysine. Supramolecular host compounds that can cage histone methyllysines are now available. For example, the calixarene-based supramolecular host can associate with methylated lysine in chromatin to selectively disrupt binding of the CHD4 PHD finger and HP1 γ CD to H3K9me3 [84].

9. Conclusion

Altogether these innovative approaches are moving us closer to a true chromatin model of the translation of histone lysine methylation. The generation of nucleosomes with specific patterns of modifications has been key in revealing the roles of these interactions in recruiting of co-factors, inducing and stabilizing chromatin structure, regulating the activity of chromatin modifying enzymes, and mediating PTM cross-talk through multivalent interactions. In this vein, the ability to produce nucleosomes with defined patterns of modifications will be essential in studies of molecular mechanisms and in the discovery of novel readers.

Disruption of patterns of histone modifications is associated with a number of diseases, and there is tremendous therapeutic potential in targeting histone modification pathways. Thus a mechanistic understanding of these processes is essential in the design of chemical probes and targeted therapeutics. As discussed above some progress has been made in the design of antagonists of the PHD and MBT family of readers. One antagonist in particular, UNC1215, which targets L3MBTL1, demonstrates the power of these small molecules both in the discovery of new methyllysine targets (especially non-histone targets) and in the disruption of their association with chromatin.

There is still much to be learned about the mechanisms of methyllysine recognition and the functional consequence of methyllysine readout. The development of new technologies and biochemical approaches to study these interactions promises that great headway will be made in the near future.

Acknowledgments

Research in the TKG laboratory is supported by grants from the NIH, GM096863 and GM101664. Research in the CAM laboratory is supported by grant IRG-77-004-34 from the American Cancer Society, administered through The Holden Comprehensive Cancer Center at The University of Iowa.

References

1. Murray K. The occurrence of epsilon-N-methyl lysine in histones. *Biochemistry*. 1964; 3:10–15. [PubMed: 14114491]
2. Allfrey VG, Faulkner R, Mirsky AE. Acetylation and methylation of histones and their possible role in the regulation of RNA synthesis. *Proc Natl Acad Sci U S A*. 1964; 51:786–794. [PubMed: 14172992]
3. Black JC, Van Rechem C, Whetstine JR. Histone lysine methylation dynamics: establishment, regulation, and biological impact. *Mol Cell*. 2012

4. Shilatifard A. Chromatin modifications by methylation and ubiquitination: implications in the regulation of gene expression. 2006; 75:243–269. <http://dx.doi.org/10.1146/annurev-biochem-072711-165700>.
5. Kouzarides T. Chromatin modifications and their function. *Cell*. 2007
6. Barski A, et al. High-resolution profiling of histone methylations in the human genome. *Cell*. 2007; 129:823–837. [PubMed: 17512414]
7. Bannister AJ, et al. Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature*. 2001; 410:120–124. [PubMed: 11242054]
8. Lachner M, O'Carroll D, Rea S, Mechtler K, Jenuwein T. Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature*. 2001; 410:116–120. [PubMed: 11242053]
9. Jacobs SA, et al. Specificity of the HP1 chromo domain for the methylated N-terminus of histone H3. *EMBO J*. 2001; 20:5232–5241. [PubMed: 11566886]
10. Nakayama JI. Role of histone H3 lysine 9 methylation in epigenetic control of heterochromatin assembly. *Science*. 2001; 292:110–113. [PubMed: 11283354]
11. Jacobs SA, Khorasanizadeh S. Structure of HP1 chromodomain bound to a lysine 9-methylated histone H3 tail. *Science*. 2002
12. Nielsen PR, et al. Structure of the HP1 chromodomain bound to histone H3 methylated at lysine 9. *Nature*. 2002; 416:103–107. [PubMed: 11882902]
13. Patel DJ, Wang Z. Readout of epigenetic modifications. 2013; 82:81–118. <http://dx.doi.org/10.1146/annurev-biochem-072711-165700>.
14. Musselman CA, Lalonde ME, Côté J, Kutateladze TG. Perceiving the epigenetic landscape through histone readers. *Nat Struct Mol Biol*. 2012; 19:1218–1227. [PubMed: 23211769]
15. Portela A, Esteller M. Epigenetic modifications and human disease. *Nat Biotechnol*. 2010; 28:1057–1068. [PubMed: 20944598]
16. Ellis, L.; Atadja, PW.; Johnstone, RW. Epigenetics in cancer: Targeting chromatin modifications. 2009.
17. Vallaster M, Vallaster CD, Wu SM. Epigenetic mechanisms in cardiac development and disease. *Acta Biochim Biophys Sin (Shanghai)*. 2012; 44:92–102. [PubMed: 22194017]
18. Chi P, Allis CD, Wang GG. Covalent histone modifications — miswritten, misinterpreted and mis-erased in human cancers. *Nat Rev Cancer*. 2010; 10:457–469. [PubMed: 20574448]
19. Fischle W, et al. Regulation of HP1–chromatin binding by histone H3 methylation and phosphorylation. *Nat Cell Biol*. 2005; 438:1116–1122.
20. Flanagan JF, et al. Double chromodomains cooperate to recognize the methylated histone H3 tail. *Nature*. 2005; 438:1181–1185. [PubMed: 16372014]
21. Varier RA, et al. A phospho/methyl switch at histone H3 regulates TFIID association with mitotic chromosomes. *EMBO J*. 2010; 29:3967–3978. [PubMed: 20953165]
22. Ali M, et al. Molecular basis for chromatin binding and regulation of MLL5. *Proc Natl Acad Sci U S A*. 2013; 110:11296–11301. [PubMed: 23798402]
23. Gatchalian J, Fütterer A, Rothbart SB, Tong Q. Dido3 PHD modulates cell differentiation and division. *Cell Rep*. 2013
24. Vermeulen M, et al. Selective anchoring of TFIID to nucleosomes by trimethylation of histone H3 lysine 4. *Cell*. 2007; 131:58–69. [PubMed: 17884155]
25. Ramón-Maiques S, et al. The plant homeodomain finger of RAG2 recognizes histone H3 methylated at both lysine-4 and arginine-2. *Proc Natl Acad Sci U S A*. 2007
26. Chignola F, et al. 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*. 2009; 37:2951–2961. [PubMed: 19293276]
27. Yuan CC, Matthews A, Jin Y, Chen CF. Histone H3R2 symmetric dimethylation and histone H3K4 trimethylation are tightly correlated in eukaryotic genomes. *Cell Rep*. 2012
28. Iwase S, et al. ATRX ADD domain links an atypical histone methylation recognition mechanism to human mental-retardation syndrome. *Nat Struct Mol Biol*. 2011; 18:769–776. [PubMed: 21666679]

29. Eustermann S, et al. Combinatorial readout of histone H3 modifications specifies localization of ATRX to heterochromatin. *Nat Struct Mol Biol.* 2011; 18:777–782. [PubMed: 21666677]
30. Dhayalan A, et al. The ATRX-ADD domain binds to H3 tail peptides and reads the combined methylation state of K4 and K9. *Hum Mol Genet.* 2011; 20:2195–2203. [PubMed: 21421568]
31. Musselman CA, et al. Binding of the CHD4 PHD2 finger to histone H3 is modulated by covalent modifications. *Biochem J.* 2009; 423:179–187. [PubMed: 19624289]
32. Mansfield RE, et al. 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]
33. Xi Q, et al. A poised chromatin platform for TGF- β access to master regulators. *Cell.* 2011; 147:1511–1524. [PubMed: 22196728]
34. Bao Y, Chakravarthy S, Muthurajan UM, Luger K. Reconstitution of nucleosome core particles from recombinant histones and DNA. *Methods Enzymol.* 2003
35. Dawson PE, Muir TW, Clark-Lewis I, Kent SB. Synthesis of proteins by native chemical ligation. *Science.* 1994; 266:776–779. [PubMed: 7973629]
36. Muir TW, Sondhi D, Cole PA. Expressed protein ligation: a general method for protein engineering. *Proc Natl Acad Sci U S A.* 1998
37. Shogren-Knaak MA, Fry CJ, Peterson CL. A native peptide ligation strategy for deciphering nucleosomal histone modifications. *J Biol Chem.* 2003; 278:15744–15748. [PubMed: 12595522]
38. He S, et al. Facile synthesis of site-specifically acetylated and methylated histone proteins: reagents for evaluation of the histone code hypothesis. *Proc Natl Acad Sci U S A.* 2003
39. Simon MD, Chu F, Racki LR, de la Cruz CC. The site-specific installation of methyl-lysine analogs into recombinant histones. *Cell.* 2007
40. Jia G, et al. A systematic evaluation of the compatibility of histones containing methyl-lysine analogues with biochemical reactions. *Cell Res.* 2009; 19:1217–1220. [PubMed: 19770845]
41. Seeliger D, et al. Quantitative assessment of protein interaction with methyl-lysine analogues by hybrid computational and experimental approaches. *ACS Chem Biol.* 2012; 7:150–154. [PubMed: 21991995]
42. Krishnan S, Collazo E, Ortiz-Tello PA, Trievel RC. Purification and assay protocols for obtaining highly active jumonji C demethylases. *Anal Biochem.* 2012; 420:48–53. [PubMed: 21925481]
43. Nguyen DP, Garcia Alai MM, Kapadnis PB, Neumann H, Chin JW. Genetically encoding N ϵ -methyl-L-lysine in recombinant histones. *J Am Chem Soc.* 2009; 131:14194–14195. [PubMed: 19772323]
44. Nguyen DP, Alai M, Virdee S, Chin JW. Genetically directing ϵ -N, N-dimethyl-L-lysine in recombinant histones. *Chem Biol.* 2010
45. Xu C, Cui G, Botuyan MV, Mer G. Structural basis for the recognition of methylated histone H3K36 by the Eaf3 subunit of histone deacetylase complex Rpd3S. *Structure.* 2008
46. Vezzoli A, et al. Molecular basis of histone H3K36me3 recognition by the PWWP domain of Brpf1. *Nat Struct Mol Biol.* 2010; 17:617–619. [PubMed: 20400950]
47. Kim D, et al. Corecognition of DNA and a methylated histone tail by the MSL3 chromodomain. *Nat Struct Mol Biol.* 2010; 17:1027–1029. [PubMed: 20657587]
48. Eidahl JO, et al. Structural basis for high-affinity binding of LEDGF PWWP to mononucleosomes. *Nucleic Acids Res.* 2013
49. van Nuland R, et al. Nucleosomal DNA binding drives the recognition of H3K36-methylated nucleosomes by the PSIP1-PWWP domain. *Epigenetics Chromatin.* 2013; 6:12. [PubMed: 23656834]
50. Smolle M, Venkatesh S, Gogol MM, Li H. Chromatin remodelers Isw1 and Chd1 maintain chromatin structure during transcription by preventing histone exchange: nature structural & molecular biology: nature publishing group. *Nat Struct Mol Biol.* 2012
51. Maltby VE, et al. Histone H3 lysine 36 methylation targets the Isw1b remodeling complex to chromatin. *Mol Cell Biol.* 2012; 32:3479–3485. [PubMed: 22751925]
52. Qiu Y, et al. Solution structure of the Pdp1 PWWP domain reveals its unique binding sites for methylated H4K20 and DNA. *Biochem J.* 2012; 442:527–538. [PubMed: 22150589]

53. Musselman CA, et al. Binding of PHF1 Tudor to H3K36me3 enhances nucleosome accessibility. *Nat Commun.* 2013; 4:2969. [PubMed: 24352064]
54. Nady N, Krichevsky L, Zhong N, Duan S. Histone recognition by human malignant brain tumor domains. *J Mol Biol.* 2012
55. Margueron R, Reinberg D. Chromatin structure and the inheritance of epigenetic information. *Nat Rev Genet.* 2010; 11:285–296. [PubMed: 20300089]
56. Tie F, Stratton CA, Kurzhals RL, Harte PJ. The N terminus of *Drosophila* ESC binds directly to histone H3 and is required for E(Z)-dependent trimethylation of H3 lysine 27. *Mol Cell Biol.* 2007; 27:2014–2026. [PubMed: 17210640]
57. Mishima Y, et al. Hinge and chromoshadow of HP1 α participate in recognition of K9 methylated histone H3 in nucleosomes. *J Mol Biol.* 2013; 425:54–70. [PubMed: 23142645]
58. Munari F, et al. Methylation of lysine 9 in histone H3 directs alternative modes of highly dynamic interaction of heterochromatin protein hHP1 β with the nucleosome. *J Biol Chem.* 2012
59. Canzio D, et al. A conformational switch in HP1 releases auto-inhibition to drive heterochromatin assembly. *Nature.* 2013; 496:377–381. [PubMed: 23485968]
60. Strahl BD, Allis CD. The language of covalent histone modifications. *Nature.* 2000; 403:41–45. [PubMed: 10638745]
61. Ruthenburg AJ, Li H, Patel DJ, David Allis C. Multivalent engagement of chromatin modifications by linked binding modules. *Nat Rev Mol Cell Biol.* 2007; 8:983–994. [PubMed: 18037899]
62. Ruthenburg AJ, et al. Recognition of a mononucleosomal histone modification pattern by BPTF via multivalent interactions. *Cell.* 2011; 145:692–706. [PubMed: 21596426]
63. Huh JW, Wu J, Lee CH, Yun M, Gilada D. Multivalent di-nucleosome recognition enables the Rpd3S histone deacetylase complex to tolerate decreased H3K36 methylation levels. *EMBO J.* 2012
64. van Nuland R, et al. Multivalent engagement of TFIID to nucleosomes. *PLoS ONE.* 2013; 8:e73495. [PubMed: 24039962]
65. Canzio D, et al. Chromodomain-mediated oligomerization of HP1 suggests a nucleosome-bridging mechanism for heterochromatin assembly. *Mol Cell.* 2011; 41:67–81. [PubMed: 21211724]
66. LeRoy G, et al. A quantitative atlas of histone modification signatures from human cancer cells. *Epigenetics Chromatin.* 2013; 6:20. [PubMed: 23826629]
67. Al-Sady B, Madhani HD, Narlikar GJ. Division of labor between the chromodomains of HP1 and Suv39 methylase enables coordination of heterochromatin spread. *Mol Cell.* 2013
68. Hung T, Binda O, Champagne KS, Kuo AJ, Johnson K. ING4 mediates crosstalk between histone H3 K4 trimethylation and H3 acetylation to attenuate cellular transformation. *Mol Cell.* 2009
69. Margueron R, et al. Role of the polycomb protein EED in the propagation of repressive histone marks. *Nature.* 2009; 461:762–767. [PubMed: 19767730]
70. Musselman CA, et al. Molecular basis for H3K36me3 recognition by the Tudor domain of PHF1. *Nat Struct Mol Biol.* 2012; 19:1266–1272. [PubMed: 23142980]
71. Arrowsmith CH, Bountra C, Fish PV, Lee K. Epigenetic protein families: a new frontier for drug discovery: article. *Nat Rev Drug Discov.* 2012
72. James LI, Frye SV. Targeting chromatin readers. *Clin Pharmacol Ther.* 2013; 93:312–314. [PubMed: 23403847]
73. Wagner, T.; Robaa, D.; Sippl, W.; Jung, M. Mind the methyl: methyllysine binding proteins in epigenetic regulation. *ChemMedChem.* 2014. <http://dx.doi.org/10.1002/cmdc.201300422>
74. Wigle TJ, Copeland RA. Drugging the human methylome: an emerging modality for reversible control of aberrant gene transcription. *Curr Opin Chem Biol.* 2013
75. Santiago C, Nguyen K, Schapira M. Druggability of methyl-lysine binding sites. *J Comput Aided Mol Des.* 2011; 25:1171–1178. [PubMed: 22146969]
76. Wagner EK, Nath N, Flemming R, Feltenberger JB, Denu JM. Identification and characterization of small molecule inhibitors of a plant homeodomain finger. *Biochemistry.* 2012; 51:8293–8306. [PubMed: 22994852]
77. Quinn AM, et al. A homogeneous method for investigation of methylation-dependent protein–protein interactions in epigenetics. *Nucleic Acids Res.* 2010; 38:e11. [PubMed: 19897549]

78. Wigle TJ, et al. Screening for inhibitors of low-affinity epigenetic peptide–protein interactions: an AlphaScreen-based assay for antagonists of methyl-lysine binding proteins. *J Biomol Screen*. 2010; 15:62–71. [PubMed: 20008125]
79. Kireev D, et al. Identification of non-peptide malignant brain tumor (MBT) repeat antagonists by virtual screening of commercially available compounds. *J Med Chem*. 2010; 53:7625–7631. [PubMed: 20931980]
80. Herold JM, et al. Small-molecule ligands of methyl-lysine binding proteins. *J Med Chem*. 2011; 54:2504–2511. [PubMed: 21417280]
81. James LI, et al. Discovery of a chemical probe for the L3MBTL3 methyllysine reader domain. *Nat Chem Biol*. 2013; 9:184–191. [PubMed: 23292653]
82. James LI, et al. Small-molecule ligands of methyl-lysine binding proteins: optimization of selectivity for L3MBTL3. *J Med Chem*. 2013; 56:7358–7371. [PubMed: 24040942]
83. Simhadri, C., et al. Chromodomain antagonists that target the polycomb-group methyllysine reader protein chromobox homolog 7 (CBX7). *J Med Chem*. 2014. <http://dx.doi.org/10.1021/jm401487x>
84. Allen, H., et al. Inhibition of histone binding by supramolecular hosts. *Biochem J*. 2014. <http://dx.doi.org/10.1042/BJ20140145>

Fig. 1. Sites of histone methylation. A mono-nucleosome (histones in salmon, DNA in gray) with sites of lysine methylation (red ovals) denoted. For clarity sites are only denoted on one of the two copies of each histone. Not shown is the K26 methylation site on the linker histone H1.

Fig. 2. Methods for installing methyllysine on histone proteins. (A) Native chemical ligation (NCL) and expressed protein ligation (EPL) strategies for the generation of semisynthetic modified histones [37,38]. Peptides with methyl groups (denoted by Me) are generated synthetically. For NCL the peptide is labeled with a thioester at the carboxy-terminus and ligated to a protein containing an amino-terminal cysteine residue (Cys). In EPL a recombinant protein with a carboxy-terminal thioester is captured and ligated to a peptide with an amino-terminal Cys. Both reactions can be further processed using hydrogen/Raney nickel to convert the Cys to an alanine (Ala). (B) Strategy for the installation of methyllysine analogues on recombinant histones [39]. The lysine of interest (Lys) is mutated to a cysteine residue (Cys). Purified mutant protein is treated with the alkylating agent of choice to produce a

mono-methyl (K_{Cme1}), di-methyl (K_{Cme2}) or tri-methyl (K_{Cme3}) lysine analogue. Unmodified lysine can also be produced (not shown). Methyl groups are denoted by Me. (C) Method for the genetic installation of mono-methyllysine histones in an *E. coli* system. Cells are transformed with plasmids for the orthogonal pyrrolysyl-tRNA synthetase/tRNA_{CUA} as well as the histone containing plasmid with an amber codon (UAG) at the desired site of mono-methylation and proteins expressed in the presence of N3-tert-butyloxycarbonyl-N3-methyl-L-lysine. Proteins are purified using standard methods and then deprotected to reveal the mono-methyllysine [43]. This method can be extended to incorporate di-methyllysine [44] (not shown).

Fig. 3. Functional consequence of methyllysine readout. Studies with methylated nucleosomes have revealed mechanisms and functional consequences of methyllysine readout, including co-factor targeting, roles in chromatin structure and chromatin enzyme regulation. (A) Some reader domains require contacts outside the histone tail for robust targeting, whereas others utilize multivalent mechanisms. Shown are the PWWP domain of LEDGF/PISP1 [48,49] (left) and Tudor domain of PHF1 [53] (middle), which have been found to concomitantly recognize H3K36me3 and nucleosomal DNA. PWWP makes electrostatic contacts with the DNA from a highly basic patch (shown in blue) on the protein surface whereas Tudor largely utilizes hydrophobic contacts (shown in green). Also shown is the potential model for multivalent targeting of the Rpd3s HDAC complex to adjacent H3K36me3-nucleosomes through interactions of the Eaf3 CD and the Rco1 PHD (right) [63]. (B) The recognition of H3K9me3 by Swi6 has been shown to be essential in the formation and spread of heterochromatin. In a dimer of Swi6 an ARK histone mimic loop (red) associates with the

adjacent CD of Swi6 locking the protein into an autoinhibited state (left) which is in equilibrium with an open and histone binding competent state (middle) from which the free CDs bind to H3K9me3 and the ARK loop makes contacts with nucleosomal DNA. Subsequent oligomerization of Swi6 mediated through the chromoshadow domain (CSD) induces heterochromatin formation and spread [59]. (C) Reader domain recognition of methyllysine has been shown to regulate the activity of chromatin enzymes. The Suv39 CD recognition of H3K9me3 increases the K9-methyltransferase activity of the enzyme by orienting the active site to an adjacent nucleosome (left) [67]. The H3 acetyltransferase activity of the HBO1 complex is upregulated through recognition of H3K4me3 by the PHD finger of ING4 (right) [68]. Colored ovals denote H3K36me3 (red), H3K9me3 (yellow) and H3K4me3 (cyan), acetyl groups are denoted as blue squares.

Fig. 4.

L3MBTL1 in complex with UNC1215. Structure of the L3MBTL1 3-MBT repeat in complex with the UNC1215 antagonist (PDB ID: 4FL6). (A) One of the 3-MBT repeats is shown as a surface representation with the aromatic cage residues colored in pink. The two UNC1215 antagonists are shown in yellow. (B) The antagonist functions by capturing a dimer state of L3MBTL1, binding in the methyllysine pocket of one molecule and bridging it with the adjacent binding pocket of the second molecule. This mechanism effectively blocks all binding regions from interaction.

Table 1

Methyllysine readers and their histone targets.

Methyllysine reader	Methyllysine target
ADD	H3K9me3
Ankyrin	H3K9me2/1
BAH	H4K20me2, H3K9me2
Chromo barrel	H3K36me3/2, H4K20me1
Chromodomain (CD)	H3K9me3/2, H3K27me3/2
Double chromodomain (DCD)	H3K4me3/2/1
HEAT	H4K20me1
MBT	H3Kme1/2, H4Kme1/2
PHD	H3K4me3/2, H3K9me3
PWWP	H3K36me3, H4K20me1/3, H3K79me3
SAWADEE	H3K9me1/2/3
Tandem Tudor domain (TTD)	H3K4me3, H3K9me3, H4K20me2
Tudor	H3K36me3
WD40	H3K27me3, H3K9me3
zf-CW	H3K4me3