Published in final edited form as:

Biochim Biophys Acta Gene Regul Mech. 2020 July ; 1863(7): 194561. doi:10.1016/j.bbagrm.2020.194561.

Insights on the regulation of the MLL/SET1 family histone methyltransferases

Liang Sha[1,*], **Alex Ayoub**[1,*], **Uhn-Soo Cho**[2], **Yali Dou**[1,2,#]

[1]Department of Pathology, University of Michigan, Ann Arbor

[2]Department of Biological Chemistry, University of Michigan, Ann Arbor

Abstract

In eukaryotes, histone H3K4 methylation by the MLL/SET1 family histone methyltransferases is enriched at transcription regulatory elements including gene promoters and enhancers. The level of H3K4 methylation is highly correlated with transcription activation and is one of the most frequently used histone post-translational modifications to predict transcriptional outcome. Recently, it has been shown that rearrangement of the cellular landscape of H3K4 mono-methylation at distal enhancers precedes cell fate transition and is used for identification of novel regulatory elements for development and disease progression. Similarly, broad H3K4 tri-methylation regions have also been used to predict intrinsic tumor suppression properties of regulator regions in a variety of cellular models. Understanding the regulation for how H3K4 methylation is deposited and regulated is of paramount importance. In this review, we will discuss new findings on how the MLL/SET1 family enzymes are regulated on chromatin and their potential functional and regulatory implications.

Keywords

Histone methylation; NCP; Mixed Lineage Leukemia (MLL); SET domain; Chromatin Binding

1. Introduction

Histone lysine methylation is a major post-translational modification (PTM) in eukaryotes. It occurs on the ε-amino group in three discrete states of mono-, di-, and tri-methylation. Since

[#]Correspondence: yalid@med.umich.edu, Tel: (734) 615-1315.
[*]Contributed equally

Author Statement

Liang Sha: Writing original draft preparation for section 3–4. **Alex Ayoub**: Writing- Original draft preparation for section 1–2. **Uhn-Soo Cho**: Writing- Reviewing and Editing. **Yali Dou**: Supervision, Writing- Reviewing and Editing

Declaration of Competing Interest

The authors declare that no conflict of interest exists.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

the first histone lysine methyltransferase (KMT or HMT) was discovered twenty years ago [1], over 60 putative or predicted enzymes have been identified [2]. Among well-characterized KMTs, the highly conserved mixed lineage leukemia (MLL or KMT2) family of proteins is responsible for deposition of the majority of histone 3 lysine 4 (H3K4) methylation in eukaryotes. Complexity of the H3K4 HMTs increases as eukaryotes evolved from single cell organisms to mammals, concomitant with increasing demands for spatial and temporal gene regulation. In yeast, ySET1 (Su(Var)3–9, Enhancer of Zeste, Trithorax), an MLL homolog, is responsible for all H3K4 methylation [3–5]. In *Drosophila melanogaster*, there are three MLL family enzymes, TRX, TRR, and dSET1, responsible for global H3K4 methylation [6, 7]. Each of the three genes (i.e., TRX, TRR and dSET1) are duplicated in mammals, giving rise to MLL1 and MLL2 (KMT2A and 2B), MLL3 and MLL4 (KMT2C and 2D), and SET1A and SET1B (KMT2F and 2G), respectively. Despite general conservation of the catalytic SET domain, each MLL/SET1 protein has non-redundant functions in development and is subject to distinct regulations [6]. Recent cryo-EM structures have revealed how the MLL1, MLL3 and ySET1 complexes bind to the nucleosome core particle (NCP) [8–11]. These studies shed light on distinct features of these MLL complexes on chromatin that may have implications for their respective regulation. Here we will focus on these exciting new discoveries and discuss how chromatin binding by the MLL family proteins is regulated in eukaryotes.

2. Regulation of the MLL methyltransferase activity on the NCPs

2a. MLL Family Enzymes Reside in a Conserved Core Complex

The MLL/SET1 family enzymes are large proteins with multiple functional domains containing substantial stretches of disordered regions [12, 13]. While they share a highly conserved C-terminal SET domain that confers H3K4 methylation [14], they also have subclass specific domains such as the CxxC and bromodomain for KMT2A/2B, the PHD domains for KMT2A-D, and the RRM domain for KMT2F/G. Biochemical studies show that the catalytic SET domains of the MLL/SET1 family enzymes have low intrinsic enzymatic activity [14, 15]. This activity can be drastically enhanced by interacting with a core complex containing highly conserved WDR5 (WD40 repeat-containing protein 5), RbBP5 (Retinoblastoma binding protein 5), ASH2L (Absent, small, homeotic disks-2-like), and DPY30 (DumPY protein 30) proteins [6, 15, 16]. Among them, WDR5, RbBP5 and ASH2L, together with $MLL1^{SET}$, are sufficient to reconstitute full activity of the MLL1 holo-complex on histone H3 [15, 17]. The core complex also acts as a platform for interacting with transcriptional factors, chromatin remodeling complexes and IncRNAs [6, 18–20], constituting a basic functional unit of the MLL/SET1 complexes.

2b. Co-crystal Structures of the MLL1/MLL3/ySET1 Core Complexes

Biochemical and structural studies have characterized inter-subunit interactions within the MLL/SET1 core complex [21–23]. Recent co-crystal structures of the MLL1, MLL3, and ySET1 complexes delineate detailed architectures of the core complex with or without substrates, i.e., *S*-adenosine-L-methionine (SAM) and histone H3 [24, 25]. Li and colleagues reported the first co-crystal structure of the MLL3 core complex showing that $MLL3^{SET}$ makes extensive interactions with an acidic surface of the RbBP5-ASH2L heterodimer via a

conserved SET-I arginine residue [25]. This interaction is stabilized by two hydrophobic residues in SET-I, which are conserved in KMT2B-G. Interestingly, this interaction is not conserved in the $MLL1^{SET}$ domain and mutating MLL1 residues to MLL3-like sequences (N3861I/Q3867L, $MLL1^{IL}$) stabilizes RbBP5-ASH2L binding and circumvents WDR5 requirement for MLL1 activation [25]. Furthermore, RbBP5-ASH2L association with $MLL3^{SET}$ and $MLL1^{IL}$ reduces SET-I flexibility, allowing for stable substrate binding [25]. This study provides a structural basis for regulation of the MLL1/3 activities by its minimal core components [15, 17, 26] as well as the unique requirement of WDR5 for the MLL1 complex [27]. Divergent SET domain sequences confer distinct biochemical properties for the MLL/SET1 family HMTs, despite overall structural similarity [28–30]. The co-crystal structure of the ySET1 complex by Hsu and colleagues shows that ySET1 contains a unique glycine-centered motif (GI/NR)G(V/I/C/SS) that acts as a 'hinge' to control substrate access to the ySET1 catalytic site [24], rendering a naturally inactive state. Distinct primary sequences of the MLL3/4 SET domain also confer precise regulation of substrate state specificity [24, 31]. Targeting unique biochemical properties of individual MLL/SET1 complex has led to development of the MLL1-specific inhibitors that show good efficacy in cancer treatment as well as embryonic stem cell reprograming [27, 32–34]. Such approaches can be envisioned to specifically target other MLL family enzymes as we learn more about their unique features in the future.

2c. Binding of the MLL1 Core Complex on the NCP

It has been shown that the MLL1 core complex has much higher activity and processivity on the NCP as compared to other substrates (i.e., H3 peptide or recombinant H3) [35]. Cryo-EM structures of the MLL1, MLL3 and ySET1 complexes with the NCP have shed light on the underlying mechanisms [8–10]. These studies, for the first time, reveal how the MLL/SET1 complexes engage the H3 substrate in a more physiological context and, importantly, highlight divergent regulation of the MLL family enzymes on chromatin. Two laboratories (including ours) reported the cryo-EM structures of the MLL1-NCP complex [8, 10]. Both structures show a dynamic interaction between the MLL1 core complex and the NCP. We show that the MLL1 core complex (EMDB: 20512, PDB: 6PWV) binds at the edge of the NCP via RbBP5 and ASH2L [8]. In this conformation, the MLL1 complex anchors on the NCP through RbBP5 and ASH2L interactions at DNA superhelix (SHL) 1.5 and SHL 7, respectively [8]. This positions $MLL1^{SET}$ at the nucleosome dyad, facilitating access of both H3 tails to the catalytic site. The RbBP5-NCP interface constitutes a conserved Quad-R motif in RbBP5 that interacts with DNA at SHL1.5 as well as an I-loop emanating from the RbBP5 WD40 repeats that interact with histone H4. Mutating two of the arginine residues in Quad-R or deleting the I-loop is sufficient to inhibit H3K4 tri-methylation on the NCP [8]. Interestingly, two of the Quad-R residues (R220Q and R272W) in RbBP5 are mutated in a variety of cancers [36, 37]. It would be interesting to examine whether they contribute to tumorigenesis by disrupting MLL1 binding to chromatin. Comparing to the RbBP5-NCP interaction, the ASH2L-NCP interface is more dynamic and involves a highly conserved $_{205}$-KRK-$_{207}$ motif at the N-terminus of ASH2L [8]. This binding conformation is also observed by Xue and colleagues (EMDB:0694, PDB: 6KIX) [10]. Interestingly, Xue and colleagues have reported a second conformation of the MLL1-NCP interaction in the majority of their cryo-EM particles [10]. In this case, $MLL1^{SET}$ binds across the nucleosome-disc, in close

proximity to the C-terminal helical region of H2A while RbBP5 binds to DNA SHL2 on the NCP. In this conformation, an arginine anchor (Arg3821) in $MLL1^{SET}$ makes contacts with asparagine 72 (Asp72) of H2A. Several hydrophobic residues in RbBP5 (Loop 1, Leu248 and Val249) are close to αC of H2B. Electrostatic contacts between RbBP5 (Loop 2, Glu296) and K79 of H3 is also observed [10]. Importance of these residues in MLL1 activity on the NCP has not been reported [10]. We have recently found that mutating residues involved in the second MLL1-NCP conformation does not affect overall MLL1 activity on the NCP (unpublished observation). It is likely that the MLL1 complex binds to the NCP in multiple conformations with distinct intrinsic activities. The highly dynamic interactions between the MLL1 complex and chromatin allows for loci-specific fine-tuning of H3K4me levels and state *in vivo*. Determining the functionally active conformation in cells and whether transition between binding conformations is regulated are important future directions.

2d. Binding of the ySET1 Core Complex on the NCP

Comparing to the MLL1-NCP complex, the cryo-EM structure of the ySET1-NCP complex shows a similar conformation as the off-dyad binding mode of the MLL1-NCP complex [9–11]. The ySET1 complex maintains the overall architecture upon NCP interaction with slight rotation of three components [24]. The SET domain of ySET1 bisects the nucleosome surface, with a conserved arginine residue (R856) in the arginine-rich motif (ARM) anchoring near the acidic patch of the NCP. The ySET1 complex engages nucleosome DNA at SHL2 through Spp1 (CFP1 homolog) and Swd1 (RbBP5 homolog) and at SHL6 through Bre2 (ASH2L homolog), respectively [9]. Importantly, the ARM of ySET1 adopts an α-helical conformation upon interacting with the NCP, which blocks H3 access to the ySET1 active site [9, 24]. This induced α-helical conformation is stabilized by electrostatic interactions between the SET domain and H2A as well as the hydrophobic interactions with αC of H2B [9, 24]. The loss of these anchoring motifs severely attenuate all H3K4me activity *in vitro* and *in vivo* [9, 11, 38]. It is worth noting is that the ARM motif is unique to ySET1 and hSET1A/SET1B proteins, therefore, its function in NCP-induced autoinhibition is not conserved for MLL1-4. Divergent SET domain sequences and different binding modes on the NCP imply that different MLL/SET1 family enzymes may be subject to distinct regulations on chromatin.

2e. Distinct regulation of the ySET1 and MLL1/3 activity by H2BK120ub

Regulation of H3K4me3 by H2B ubiquitylation (H2BK120ub) is one of the best described histone cross-talks *in vivo.* In yeast, H2BK120ub, a highly prevalent mark [39, 40], is a prerequisite for global H3K4me3 [38–40]. While a previous biochemical study identified ARM of ySET1 and Spp1 are essential factors for this regulation [41], the cryo-EM structure of the ySET1-H2BK120ubNCP complex reveals fascinating molecular details underlying this regulation [9]. The H2BK120ub mediated trans-tail regulation occurs intra-nucleosomally, similar to that of DOT1L [42]. It shows that H2BK120ub is able to alleviate ARM-mediated auto-inhibition to enhance ySET1 activity on the NCP [9, 11]. Ubiquitin attached to H2BK120 makes direct contact with ARM and effectively moves ARM away to allow H3 access to the catalytic site [9]. It also stabilizes the SET domain on the NCP without altering overall binding affinity [9]. Notably, H2BK120ub-mediated trans-tail

regulation is relatively modest in the absence of Spp1, which augments H2BK120ub-dependent stimulation, in part by suppressing basal activity of the ySET1 complex on the unmodified NCP [9]. However, the role of Spp1 in ySET1 regulation on the NCP is probably more complicated. Previous studies show that Spp1 increases ySET1 activity on the unmodified NCP [41] and genetic deletion of Spp1 leads to 40% reduction of global H3K4me *in vivo* [43]. The discrepancy could be due to presence of other components in previous studies, i.e., Swd2 and Shg1, that significantly repress ySET1 activity on the H2BK120ub-NCP but are not included in the structure studies. Alternatively, the extended nSET region in the ySET1 may confer additional functional interplays with Spp1 that remain to be characterized [41]. The mammalian homolog of Spp1, CFP1, is important for activation of the hSET1 complex on H2BK120ub-NCP [41]. It would be important to examine the conserved and divergent functions of CFP1 in the hSET1 complex for H2BK120ub-dependent regulation.

Given the unique autoinhibitory role of ARM in ySET1 and lack of conservation of ARM in MLL1-4 homologs, it raises the question of whether H2BK120ub regulates other MLL family enzymes? Unlike yeast, decoupling of H2BK120ub and H3K4me3 has been widely described in mammals *and Tetrahymena thermophila* [44, 45]. Furthermore, it has been shown that the MLL1 complex has very high methylation activity on the unmodified NCP [8, 17, 46]. H2BK120ub does not regulate activity of the MLL3 complex and has only a modest effect on the MLL1 activity *in vitro* [46]. The cryo-EM structure of the MLL1/3 core complexes bound to H2BK120ubNCP (MLL1 EMDB:9999, PDB: 6KIV; MLL3 EMDB: 0693, PDB: 6KIW) seems to support a different role of H2BK120ub in regulation of the MLL1/3 activity. Although the MLL1/3 complexes overlay well with that of ySET1-H2BK120ub-NCP, key interactions between ySET1-H2BK120ub-NCP are not conserved in the MLL1/3-H2BK120ub-NCP complexes [9, 10]. The N-terminal region of the $MLL1^{SET}$ domain interacts directly with neither ubiquitin nor the “acidic patch” on the NCP. Instead, the ubiquitin module exhibits dynamic binding to multiple different surfaces near RbBP5 [10]. These results suggest that H2BK120ub is probably less important for regulating the MLL1/3 complexes in higher eukaryotes. Alternatively, it may regulate MLL activities through proteins that are not fully characterized in the structure.

3. Multi-Valent Chromatin Interactions for MLL1

While the structure and biochemical studies show extensive interactions between the MLL/SET1 core complex and the NCP, the MLL/SET1 family enzymes also contain multiple other chromatin-interacting domains that are capable of recognizing specific patterns of histone and DNA modifications. These interactions may contribute to specific distribution of H3K4 methylation at transcriptionally-active gene promoters and distal regulatory enhancers [47] as well as colocalization of H3K4 methylation with other prominent co-transcriptional marks such as H3 acetylation, H3K79 and H3K36 methylation [47–50]. Close correlation of H3K4 methylation has also been established with hypo-methylated DNA [51–54]. We will briefly review the current understanding of the function and regulation of the chromatin binding domains in MLL1.

3a. MLL1 binds to H3K36me2 loci through interactions with LEDGF/MENIN.

MLL1 forms a tripartite complex with tumor suppressors MENIN and LEDGF/p75 [55, 56]. Chromatin binding for each protein is mutually dependent. Deletion of MENIN or LEDGF/p75 significantly reduces MLL1 recruitment to the target genes (i.e., *HoxA9,* $p27^{kip1}$ and $p18^{ink4c}$) [56–59]. Reciprocally, MLL1 also plays a critical role in supporting MENIN function *in vivo* [57]. Since LEDGF specifically binds to H3K36me2 [60], it is able to recruit the MLL1 complex to genomic regions enriched for H3K36me2. In support of this, the histone H3K36 methyltransferase Ash1L co-localizes with MLL1 and LEDGF and is required for transcriptional activation [60, 61]. Similarly, H3K36 demethylase KDM2A promotes dissociation of MLL1 and LEDGF from chromatin and functionally antagonistic to both MLL1 and ASH1L in leukemic transformation [60]. While MENIN is a stable component of the MLL2 (KMT2B) complex [55, 62], it remains to be determined whether LEDGF is a *bona fide* component of the MLL2 complex. Furthermore, since MENIN and LEDGF/p75 interact with the oncogenic MLL1 fusion proteins [63], rationally targeting the MLL1-MENIN or LEDGF/p75-MLL1-MENIN interactions has shown great efficacy in blocking MLL1-rearranged leukemia [63–66].

3b. CxxC domain recruits MLL1 to unmethylated DNA CpG islands.

MLL1 contains a CxxC domain that is retained in the MLL1 fusion proteins after chromosomal rearrangement. The MLL1 CxxC domain binds to unmethylated DNA [53, 67]. Structural studies show that the MLL1 CxxC domain makes rigid contacts with DNA nucleobases [68]. The methyl group on cytosine creates a steric clash to the CxxC binding pocket [68]. However, the CxxC mutant deficient in DNA binding does not affect MLL-AF9 recruitment to chromatin with high levels of DNA methylation, suggesting the CxxC domain is not a key contributor to overall binding affinity of MLL1 to chromatin. Instead, it acts to passively regulate DNA methylation by excluding the binding of other CpG-binding proteins [54, 68]. In addition to interacting with DNA, the CxxC domain is also able to interact with the polymerase associating factor (PAF) elongation complex (PAF1C) in mammals [69–71]. This interaction involves a key arginine residue (R1153) of MLL1 that is not important for DNA binding [72]. The R1153 residue is not conserved in MLL2, which has alanine in its place. The R1153A mutant abolishes the interaction between MLL1 CxxC domain and PAF1C [69] and leads to reduced recruitment of MLL-AF9 to HOX targets and attenuation of leukemic transformation [69–71]. These studies demonstrate that the CxxC domains in MLL1 and MLL2 have distinct functions, in part due to differential PAF1C interactions. The CxxC domain is not conserved in KMT2C, 2D, 2F or 2G. However, the function of CxxC domain is likely partially conserved through CFP1 [73, 74], a stable component of the hSET1 complexes. While CFP1 (and human SET1) does not interact with PAF1C [69], the CxxC domain of CFP1 is able to bind non-methylated CpG [75]. Distinct from that of MLL1, CFP1 is causally linked to *de novo* establishment of H3K4me3 at non-methylated CpG in mammalian cells [74, 76–78].

3c. Multifaceted functions of the PHD fingers in MLL1.

MLL1 contains four Plant Homeodomains (PHD) and a bromodomain immediately C-terminal to the CxxC domain. The PHD fingers are also present in other MLL family

enzymes [6]. PHD fingers, together with the CxxC domain, are essential for recruitment of MLL1 to gene targets on chromatin [69, 71]. Specifically, PHD3 is able to recognize di- and tri-methylated H3K4 (H3K4me2/3) [69], contributing to spreading of H3K4 tri-methylation through the coupled 'writer-reader' regulation [47, 50]. Mutation of PHD3 attenuates chromatin recruitment of MLL1 and expression of MLL1 target genes [79]. Interestingly, PHD3 of MLL1 also interacts Cyp33, which is required for HDAC-dependent gene repression [80, 81]. Cyp33 contains a peptidyl prolyl isomerase (PPI) domain on the C-terminus [81, 82]. It induces isomerization of the proline 1629 (P1629) in the MLL1 PHD3-bromodomain, allowing PHD3 to directly interact with its RNA recognition motif (RRM). Binding of PHD3 to H3K4me2/3 and Cyp33 RRM are mutually exclusive. Cyp33 overexpression dramatically decreases H3K4me3 at MLL1 target genes [82], enabling Cyp33 to act as a regulatory switch for gene regulation [82]. In MLL1-rearranged leukemia, PHD3 is not present in the fusion proteins. Loss of PHD3 and the Cyp33-mediated repression potentially leads to constitutive activation of HOXA-mediated leukemia program for malignant transformation [83, 84]. Notably, Cyp33 overexpression can also inhibit leukemia independent of MLL1 rearrangement [83], likely through a different mechanism. Cyp33 interacts specifically with the PHD3 domain of MLL1, but not that of MLL2, despite over 70% sequence homology [81]. Functions of PHD domains in other MLL family enzymes for recognition of H4K16 acetylation and protein degradation have also been reported [85, 86].

3d. Functional interplay between MLL1/SET1 and CBP

MLL1 has a conserved trans-activation domain (TAD) that interacts with CREB-binding protein (CBP) [87]. A solution structure of a ternary complex for the activation domain of transcription factor c-Myb, MLL1 TAD and CBP kinase-inducible domain-interacting domain (KIX) has been reported [88]. Binding of MLL1 TAD stabilizes the binary interaction between c-Myb and CBP through conformational changes in the disordered regions of the KIX domain [88]. MLL1 TAD binding also facilitates interaction between phosphorylated CREB and CBP [87]. The MLL1 TAD-mediated transactivation is largely suppressed by co-expression of adenovirus $E1A_{12S}$, a competitive inhibitor of CBP, or by MLL1 TAD mutants deficient in CBP binding [87]. Interestingly, CBP seems to dictate MLL1 recruitment to either E2F1-mediated early stage pro-survival genes or late stage pro-apoptotic genes in a hepatocellular carcinoma mouse model [89]. The interaction between MLL1 and CBP is evolutionarily conserved. In *Drosophila*, TRX resides in a stable complex with dCBP, which cooperates with TRX in homeotic gene regulation [90]. Similarly, p300/CBP also interacts with the mammalian SET1 complex [91]. Tang and colleagues have elegantly demonstrated that both p300 and the SET1 complex are required for efficient p53-dependent transcription from a reconstituted chromatin template *in vitro* [91]. Although p300 is sufficient to initiate transcription on chromatin in a p53-dependent manner, recruitment of the SET1 complex by p300 enhances H3K4 methylation and further activates transcription in a p300 dose-dependent manner [91]. Knockdown p300 by siRNA leads to global down regulation of H3K4me3 [91]. Given autoinhibition of the SET1 activity on the NCP [9], it remains to be examined how p300/CBP is able to alleviate NCP-induced autoinhibition of SET1, independent of H2BK120ub. It would be interesting to test whether

p300 binding leads to conformational changes in the SET1 complex to alleviate autoinhibition, similar to that of H2BK120ub.

4. Closing Remarks

Spatial and temporal regulation of gene expression is crucial for normal development and proper cellular response to environmental cues. The MLL/SET1 family proteins, through histone H3K4 methylation, play essential roles in regulating specific gene program in cells. As revealed in the recent cryo-EM structures of the MLL1-NCP complex [8, 10], stabilization of the MLL1 complex on chromatin likely enhances the higher H3K4 methylation states, which is pivotal for transcription activation. These studies also highlight divergent regulation of the H3K4 methylation activity within the MLL/SET1 family enzymes, which is further supported by distinct functions of multiple chromatin-interacting domains for each of the MLL/SET proteins. Future studies on how conservation, or lack thereof, of each distinct functional domain contributes to unique functions of the MLL/SET1 family proteins are warranted. Furthermore, it is also important to study how different chromatin-interacting domains within each MLL/SET1 protein coordinate with each other in executing chromatin-based functions. Understanding the complexity of the MLL/SET1 regulation will shed light on how the cellular epigenetic landscape can be modulated in response to cell fate transition during normal and aberrant biological processes.

Acknowledgments

Funding

This work is supported by Michigan Institute for Clinical and Health Research (MICHR) Postdoctoral Translational Scholar Program (PTSP) Fellowship to L.S, University of Michigan Rackham MERIT Fellowship to A.A and NIGMS grant (GM082856) to Y.D.

Bibliography

1. Rea S, et al., Regulation of chromatin structure by site-speci®c histone H3 methyltransferases. NATURE, 2000 406.
2. Binda O, On your histone mark, SET, methylate! Epigenetics, 2013 8(5): p. 457–463. [PubMed: 23625014]
3. Roguev A, The Saccharomyces cerevisiae Set1 complex includes an Ash2 homologue and methylates histone 3 lysine 4. The EMBO Journal, 2001 20: p. 7137–7148. [PubMed: 11742990]
4. Briggs SD, et al., Histone H3 lysine 4 methylation is mediated by Set1 and required for cell growth and rDNA silencing in Saccharomyces cerevisiae. Genes Dev, 2001 15(24): p. 3286–95. [PubMed: 11751634]
5. Nagy PL, et al., A trithorax-group complex purified from Saccharomyces cerevisiae is required for methylation of histone H3. Proc Natl Acad Sci U S A, 2002 99(1): p. 90–4. [PubMed: 11752412]
6. Rao RC and Dou Y, Hijacked in cancer: the KMT2 (MLL) family of methyltransferases. Nature Reviews Cancer, 2015 15: p. 334. [PubMed: 25998713]
7. Morgan MA and Shilatifard A, Drosophila sets its sights on cancer: Trr/MLL3/4 COMPASS-like complexes in development and disease. Molecular and Cellular Biology, 2013 33(9): p. 1698–1701. [PubMed: 23459940]
8. Park SH, et al., Cryo-EM structure of the human Mixed Lineage Leukemia-1 complex bound to the nucleosome. bioRxiv, 2019: p. 737478.
9. Hsu PL, et al., Structural Basis of H2B Ubiquitination-Dependent H3K4 Methylation by COMPASS. Mol Cell, 2019 76(5): p. 712–732 e4. [PubMed: 31733991]

10. Xue H, et al., Structural basis of nucleosome recognition and modification by MLL methyltransferases. Nature, 2019 573(7774): p. 445–449. [PubMed: 31485071]
11. Worden EJ, Zhang X, and Wolberger C, Structural basis for COMPASS recognition of an H2B-ubiquitinated nucleosome. Elife, 2020 9.
12. Lazar T, et al., Intrinsic protein disorder in histone lysine methylation. Biology Direct, 2016 11(1): p. 30. [PubMed: 27356874]
13. Szabó B, et al., Disordered Regions of Mixed Lineage Leukemia 4 (MLL4) Protein Are Capable of RNA Binding. International Journal of Molecular Sciences, 2018 19(11): p. 3478.
14. Milne TA, et al., MLL Targets SET Domain Methyltransferase Activity to Hox Gene Promoters. Molecular Cell, 2002 10(5): p. 1107–1117. [PubMed: 12453418]
15. Dou Y, et al., Regulation of MLL1 H3K4 methyltransferase activity by its core components. Nature Structural & Molecular Biology, 2006 13: p. 713–719.
16. Cho YW, et al., PTIP associates with MLL3- and MLL4-containing histone H3 lysine 4 methyltransferase complex. J Biol Chem, 2007 282(28): p. 20395–406. [PubMed: 17500065]
17. Patel A, et al., On the mechanism of multiple lysine methylation by the human mixed lineage leukemia protein-1 (MLL1) core complex. J Biol Chem, 2009 284(36): p. 24242–56. [PubMed: 19556245]
18. Dou Y, et al., Physical Association and Coordinate Function of the H3 K4 Methyltransferase MLL1 and the H4 K16 Acetyltransferase MOF. Cell, 2005 121(6): p. 873–885. [PubMed: 15960975]
19. Cosgrove MS and Patel A, Mixed lineage leukemia: a structure-function perspective of the MLL1 protein. The FEBS journal, 2010 277(8): p. 1832–1842. [PubMed: 20236310]
20. Yang YW, et al., Essential role of lncRNA binding for WDR5 maintenance of active chromatin and embryonic stem cell pluripotency. Elife, 2014 3: p. e02046. [PubMed: 24521543]
21. Avdic V, et al., Structural and Biochemical Insights into MLL1 Core Complex Assembly. Structure, 2011 19: p. 101–108. [PubMed: 21220120]
22. Odho Z, Southall SM, and Wilson JR, Characterization of a Novel WDR5-binding Site That Recruits RbBP5 through a Conserved Motif to Enhance Methylation of Histone H3 Lysine 4 by Mixed Lineage Leukemia Protein-1. Journal of Biological Chemistry, 2010 285(43): p. 32967–32976. [PubMed: 20716525]
23. Patel A, et al., A Conserved Arginine-containing Motif Crucial for the Assembly and Enzymatic Activity of the Mixed Lineage Leukemia Protein-1 Core Complex. Journal of Biological Chemistry, 2008 283(47): p. 32162–32175. [PubMed: 18829457]
24. Hsu PL, et al., Crystal Structure of the COMPASS H3K4 Methyltransferase Catalytic Module. Cell, 2018 174: p. 1106–1116.e9. [PubMed: 30100181]
25. Li Y, et al., Structural basis for activity regulation of MLL family methyltransferases. Nature, 2016 530: p. 447. [PubMed: 26886794]
26. Cao F, et al., An Ash2L/RbBP5 heterodimer stimulates the MLL1 methyltransferase activity through coordinated substrate interactions with the MLL1 SET domain. PLoS One, 2010 5(11): p. e14102. [PubMed: 21124902]
27. Cao F, et al., Targeting MLL1 H3K4 Methyltransferase Activity in Mixed-Lineage Leukemia. Molecular Cell, 2014 53(2): p. 247–261. [PubMed: 24389101]
28. Ansari KI and Mandal SS, Mixed lineage leukemia: roles in gene expression, hormone signaling and mRNA processing. FEBS J, 2010 277(8): p. 1790–804. [PubMed: 20236313]
29. Krivtsov AV and Armstrong SA, MLL translocations, histone modifications and leukaemia stem-cell development. Nat Rev Cancer, 2007 7(11): p. 823–33. [PubMed: 17957188]
30. Wang P, et al., Global analysis of H3K4 methylation defines MLL family member targets and points to a role for MLL1-mediated H3K4 methylation in the regulation of transcriptional initiation by RNA polymerase II. Mol Cell Biol, 2009 29(22): p. 6074–85. [PubMed: 19703992]
31. Zhang Y, et al., Evolving Catalytic Properties of the MLL Family SET Domain. Structure, 2015 23(10): p. 1921–1933. [PubMed: 26320581]
32. Grebien F, et al., Pharmacological targeting of the Wdr5-MLL interaction in C/EBPalpha N-terminal leukemia. Nat Chem Biol, 2015 11(8): p. 571–578. [PubMed: 26167872]

33. Zhang H, et al., MLL1 Inhibition Reprograms Epiblast Stem Cells to Naive Pluripotency. Cell Stem Cell, 2016 18(4): p. 481–94. [PubMed: 26996599]
34. Aho ER, et al., Displacement of WDR5 from Chromatin by a WIN Site Inhibitor with Picomolar Affinity. Cell Rep, 2019 26(11): p. 2916–2928 e13. [PubMed: 30865883]
35. Patel A, et al., A novel non-set domain multi-subunit methyltransferase required for sequential nucleosomal histone H3 methylation by the mixed lineage leukemia protein-1 (MLL1) core complex. Journal of Biological Chemistry, 2010.
36. Kim T-M, et al., The mutational burdens and evolutionary ages of early gastric cancers are comparable to those of advanced gastric cancers. The Journal of Pathology, 2014 234: p. 365–374. [PubMed: 25042771]
37. Biankin AV, et al., Pancreatic cancer genomes reveal aberrations in axon guidance pathway genes. Nature, 2012 491: p. 399–405. [PubMed: 23103869]
38. Zheng S, Wyrick JJ, and Reese JC, Novel trans-tail regulation of H2B ubiquitylation and H3K4 methylation by the N terminus of histone H2A. Mol Cell Biol, 2010 30(14): p. 3635–45. [PubMed: 20498280]
39. Sun ZW and Allis CD, Ubiquitination of histone H2B regulates H3 methylation and gene silencing in yeast. Nature, 2002 418(6893): p. 104–8. [PubMed: 12077605]
40. Dover J, et al., Methylation of histone H3 by COMPASS requires ubiquitination of histone H2B by Rad6. J Biol Chem, 2002 277(32): p. 28368–71. [PubMed: 12070136]
41. Kim J, et al., The n-SET Domain of Set1 Regulates H2B Ubiquitylation-Dependent H3K4 Methylation. Molecular Cell, 2013 49(6): p. 1121–1133. [PubMed: 23453808]
42. McGinty RK, et al., Chemically ubiquitylated histone H2B stimulates hDot1L-mediated intranucleosomal methylation. Nature, 2008 453(7196): p. 812–816. [PubMed: 18449190]
43. Dehé P-M, et al., Protein Interactions within the Set1 Complex and Their Roles in the Regulation of Histone 3 Lysine 4 Methylation. Journal of Biological Chemistry, 2006 281: p. 35404–35412. [PubMed: 16921172]
44. Vethantham V, et al., Dynamic Loss of H2B Ubiquitylation without Corresponding Changes in H3K4 Trimethylation during Myogenic Differentiation. Molecular and Cellular Biology, 2012 32(6): p. 1044–1055. [PubMed: 22252316]
45. Wang Z, Cui B, and Gorovsky MA, Histone H2B Ubiquitylation Is Not Required for Histone H3 Methylation at Lysine 4 in Tetrahymena. Journal of Biological Chemistry, 2009 284(50): p. 34870–34879. [PubMed: 19822522]
46. Wu L, et al., ASH2L regulates ubiquitylation signaling to MLL: trans-regulation of H3 K4 methylation in higher eukaryotes. Molecular cell, 2013 49(6): p. 1108–1120. [PubMed: 23453805]
47. Hyun K, et al., Writing, erasing and reading histone lysine methylations. Exp Mol Med, 2017 49(4): p. e324. [PubMed: 28450737]
48. Ha M, et al., Coordinated histone modifications are associated with gene expression variation within and between species. Genome Res, 2011 21(4): p. 590–8. [PubMed: 21324879]
49. Barski A, et al., High-resolution profiling of histone methylations in the human genome. Cell, 2007 129(4): p. 823–37. [PubMed: 17512414]
50. Ruthenburg AJ, Allis CD, and Wysocka J, Methylation of lysine 4 on histone H3: intricacy of writing and reading a single epigenetic mark. Mol Cell, 2007 25(1): p. 15–30. [PubMed: 17218268]
51. Rose NR and Klose RJ, Understanding the relationship between DNA methylation and histone lysine methylation. Biochim Biophys Acta, 2014 1839(12): p. 1362–72. [PubMed: 24560929]
52. Deaton AM and Bird A, CpG islands and the regulation of transcription. Genes Dev, 2011 25(10): p. 1010–22. [PubMed: 21576262]
53. Birke M, The MT domain of the proto-oncoprotein MLL binds to CpG-containing DNA and discriminates against methylation. Nucleic Acids Research, 2002 30: p. 958–965. [PubMed: 11842107]
54. Erfurth FE, et al., MLL protects CpG clusters from methylation within the Hoxa9 gene, maintaining transcript expression. Proc Natl Acad Sci U S A, 2008 105(21): p. 7517–22. [PubMed: 18483194]

55. Hughes CM, et al., Menin associates with a trithorax family histone methyltransferase complex and with the hoxc8 locus. Mol Cell, 2004 13(4): p. 587–97. [PubMed: 14992727]
56. Yokoyama A and Cleary ML, Menin critically links MLL proteins with LEDGF on cancer-associated target genes. Cancer Cell, 2008 14(1): p. 36–46. [PubMed: 18598942]
57. Milne TA, et al., Menin and MLL cooperatively regulate expression of cyclin-dependent kinase inhibitors. Proc Natl Acad Sci U S A, 2005 102(3): p. 749–54. [PubMed: 15640349]
58. Thiel AT, et al., MLL-AF9-induced leukemogenesis requires coexpression of the wild-type Mll allele. Cancer Cell, 2010 17(2): p. 148–59. [PubMed: 20159607]
59. Chen YX, et al., The tumor suppressor menin regulates hematopoiesis and myeloid transformation by influencing Hox gene expression. Proc Natl Acad Sci U S A, 2006 103(4): p. 1018–23. [PubMed: 16415155]
60. Zhu L, et al., ASH1L Links Histone H3 Lysine 36 Dimethylation to MLL Leukemia. Cancer Discov, 2016 6(7): p. 770–83. [PubMed: 27154821]
61. Jones M, et al., Ash1l controls quiescence and self-renewal potential in hematopoietic stem cells. J Clin Invest, 2015 125(5): p. 2007–20. [PubMed: 25866973]
62. van Nuland R, et al., Quantitative Dissection and Stoichiometry Determination of the Human SET1/MLL Histone Methyltransferase Complexes. Molecular and Cellular Biology, 2013 33: p. 2067–2077. [PubMed: 23508102]
63. Yokoyama A, et al., The menin tumor suppressor protein is an essential oncogenic cofactor for MLL-associated leukemogenesis. Cell, 2005 123(2): p. 207–18. [PubMed: 16239140]
64. Caslini C, et al., Interaction of MLL amino terminal sequences with menin is required for transformation. Cancer Res, 2007 67(15): p. 7275–83. [PubMed: 17671196]
65. El Ashkar S, et al., LEDGF/p75 is dispensable for hematopoiesis but essential for MLL-rearranged leukemogenesis. Blood, 2018 131(1): p. 95–107. [PubMed: 29084774]
66. Borkin D, et al., Pharmacologic inhibition of the Menin-MLL interaction blocks progression of MLL leukemia in vivo. Cancer Cell, 2015 27(4): p. 589–602. [PubMed: 25817203]
67. Ayton PM, Chen EH, and Cleary ML, Binding to nonmethylated CpG DNA is essential for target recognition, transactivation, and myeloid transformation by an MLL oncoprotein. Mol Cell Biol, 2004 24(23): p. 10470–8. [PubMed: 15542854]
68. Cierpicki T, et al., Structure of the MLL CXXC domain-DNA complex and its functional role in MLL-AF9 leukemia. Nat Struct Mol Biol, 2010 17(1): p. 62–8. [PubMed: 20010842]
69. Milne TA, et al., Multiple interactions recruit MLL1 and MLL1 fusion proteins to the HOXA9 locus in leukemogenesis. Mol Cell, 2010 38(6): p. 853–63. [PubMed: 20541448]
70. Speck NA and Vakoc CR, PAF is in the cabal of MLL1-interacting proteins that promote leukemia. Cancer Cell, 2010 17(6): p. 531–2. [PubMed: 20541697]
71. Muntean AG, et al., The PAF complex synergizes with MLL fusion proteins at HOXloci to promote leukemogenesis. Cancer Cell, 2010 17(6): p. 609–21. [PubMed: 20541477]
72. Bach C, et al., Alterations of the CxxC domain preclude oncogenic activation of mixed-lineage leukemia 2. Oncogene, 2009 28(6): p. 815–23. [PubMed: 19060922]
73. Clouaire T, et al., Cfp1 integrates both CpG content and gene activity for accurate H3K4me3 deposition in embryonic stem cells. Genes & development, 2012 26: p. 1714–28. [PubMed: 22855832]
74. Brown DA, et al., The SET1 Complex Selects Actively Transcribed Target Genes via Multivalent Interaction with CpG Island Chromatin. Cell Rep, 2017 20(10): p. 2313–2327. [PubMed: 28877467]
75. Xu C, et al., The structural basis for selective binding of non-methylated CpG islands by the CFP1 CXXC domain. Nat Commun, 2011 2: p. 227. [PubMed: 21407193]
76. Thomson JP, et al., CpG islands influence chromatin structure via the CpG-binding protein Cfp1. Nature, 2010 464(7291): p. 1082–6. [PubMed: 20393567]
77. Clouaire T, Webb S, and Bird A, Cfp1 is required for gene expression-dependent H3K4 trimethylation and H3K9 acetylation in embryonic stem cells. Genome Biol, 2014 15(9): p. 451. [PubMed: 25201068]

78. Yu C, et al., CFP1 Regulates Histone H3K4 Trimethylation and Developmental Potential in Mouse Oocytes. Cell Rep, 2017 20(5): p. 1161–1172. [PubMed: 28768200]

79. Chang PY, et al., Binding of the MLL PHD3 finger to histone H3K4me3 is required for MLL-dependent gene transcription. J Mol Biol, 2010 400(2): p. 137–44. [PubMed: 20452361]

80. Xia ZB, et al., MLL repression domain interacts with histone deacetylases, the polycomb group proteins HPC2 and BMI-1, and the corepressor C-terminal-binding protein. Proc Natl Acad Sci U S A, 2003 100(14): p. 8342–7. [PubMed: 12829790]

81. Wang Z, et al., Pro Isomerization in MLL1 PHD3-Bromo Cassette Connects H3K4me Readout to CyP33 and HDAC-Mediated Repression. Cell, 2010 141(7): p. 1183–1194. [PubMed: 20541251]

82. Park S, et al., The PHD3 domain of MLL acts as a CYP33-regulated switch between MLL-mediated activation and repression. Biochemistry, 2010 49(31): p. 6576–86. [PubMed: 20677832]

83. Fair K, et al., Protein interactions of the MLL PHD fingers modulate MLL target gene regulation in human cells. Mol Cell Biol, 2001 21(10): p. 3589–97. [PubMed: 11313484]

84. Muntean AG, et al., The PHD fingers of MLL block MLL fusion protein-mediated transformation. Blood, 2008 112(12): p. 4690–3. [PubMed: 18796627]

85. Zhang Y, et al., Selective binding of the PHD6 finger of MLL4 to histone H4K16ac links MLL4 and MOF. Nat Commun, 2019 10(1): p. 2314. [PubMed: 31127101]

86. Wang J, et al., A subset of mixed lineage leukemia proteins has plant homeodomain (PHD)-mediated E3 ligase activity. J Biol Chem, 2012 287(52): p. 43410–6. [PubMed: 23129768]

87. Ernst P, et al., MLL and CREB Bind Cooperatively to the Nuclear Coactivator CREB-Binding Protein. Molecular and Cellular Biology, 2001 21(7): p. 2249. [PubMed: 11259575]

88. De Guzman RN, et al., Structural Basis for Cooperative Transcription Factor Binding to the CBP Coactivator. Journal of Molecular Biology, 2006 355(5): p. 1005–1013. [PubMed: 16253272]

89. Swarnalatha M, Singh AK, and Kumar V, Promoter occupancy of MLL1 histone methyltransferase seems to specify the proliferative and apoptotic functions of E2F1 in a tumour microenvironment. J Cell Sci, 2013 126(Pt 20): p. 4636–46. [PubMed: 23868976]

90. Petruk S, et al., Trithorax and dCBP acting in a complex to maintain expression of a homeotic gene. Science, 2001 294(5545): p. 1331–4. [PubMed: 11701926]

91. Tang Z, et al., SET1 and p300 Act Synergistically, through Coupled Histone Modifications, in Transcriptional Activation by p53. Cell, 2013 154(2): p. 297–310. [PubMed: 23870121]

Highlights

- Regulation of the MLL/SET1 activity by a conserved mechanism
- The MLL/SET1 complexes exhibit dynamic conformations on the NCP
- H2BK120ub alleviates repression of ySET1 activity on the NCP
- Multivalent chromatin interactions by the MLL1 complex