# MLL/SET1 Complex: From Yeast to Human

**Fang Cao***

*Department of Pathology, University of Michigan, USA*

The past decade witnessed the significant advancement of our knowledge in chromatin biology. Histone post-translational modifications are linked to different biological pathways. Among various histone modifications, H3 K4 methylation has been proposed as a critical component in regulating gene expression, epigenetic states and cellular identities. It antagonizes the functions of another histone methylation mark H3 K27 methylation by PcG group proteins, to set the chromatin activation state by the developmentally regulated genes. The founding member of the H3 K4 methyltransferase is MLL1 (Mixed Lineage Leukemia protein), which is essential for embryonic development and hematopoiesis. Relatively, its mis-regulation is associated with a variety of diseases including leukemia, multiple myeloma and brain tumors. In the past ten years, there has been an exponential increase in our knowledge regarding the MLL family of H3 K4 methyltransferase, and the biological role of this histone modification from yeast to humans.

## Role of MLL Mediated H3 K4 Methylation in Hox Gene Activation and Leukemogenesis

MLL is essential for definitive hematopoiesis by regulating transcription activation of Hox genes (e.g. Hoxa9 and Meis1), which encode transcription/regulatory factors promoting hematopoietic stem cell expansion. MLL effects mono-, di-, and tri-methylation through its evolutionarily conserved SET domain [1]. Both MLL and H3 K4 methylation are localized broadly across promoter, 5' transcribed and coding regions of the critical target genes, and facilitate the recruitment of RNA Pol-II and other chromatin remodeling activities involved in transcription activation [1,2]. Deregulation of MLL is associated with acute lymphoid and myeloid leukemia. In most cases, balanced chromosome translocations occur on one MLL allele and result in leukemogenic MLL fusion proteins (e.g. MLL-AF9, MLL-ENL), lacking the C-terminal SET domain [2]. However, MLL fusion proteins cooperate with the remaining copy of wild type MLL in leukemogenesis [3,4]. It was shown recently that in leukemia cells transformed by MLL-AF9, both wild-type MLL and oncogenic MLL-AF9 fusion proteins were recruited to Hox gene loci. Furthermore, wild-type MLL is required to promote Hox gene expression through persistent H3 K4 methylation, and to maintain MLL-AF9-transformed leukemia cells. In addition to MLL translocation, MLL amplification and tandem duplication are also reported in AML and in patients with myelodysplastic syndrome (MDS) [5].

MLL1 encodes 3969 amino acids. MLL is proteolytically cleaved into two fragments: 320KD MLLN and 180KD MLLC immediately after translation [6], which are then incorporated into the same complex. MLL has many domains implicated in chromatin functions: starting from N-terminus, a DNA methyltransferase (DNMT) homology domain, four PHD domains (Plant Homeo Domain), one of which was shown to bind H3 K4me3, a bromo-domain, a trans-activation domain (TA) [7] and a C-terminal SET domain. These domains play important roles in MLL function. In three reported MLL knockout mouse models, progressive deletion of these functional domains leads to progressively more severe phenotypes [8-10]. MLL resides in a 1.6 mDa complex, with a dozen of polypeptides as the tightly associated components [1].

The MLL complex functions coordinately with histone acetyltransferase MOF and activates transcription both on a recombinant chromatin template *in vitro,* and on its targets (Hoxa9 and Meis1) *in vivo* [1]. Using in vitro biochemical reconstitution approach, MLLC and three highly conserved components (i.e. RbBP5, Ash2L, and WDR5) are sufficient to recapitulate most of the methyltransferase activity of the MLL holo-complex [11]. Later studies showed that WDR5 makes direct contact with MLL through a conserved arginine (R3765) residue in the MLL pre-SET domain [12,13], and with RbBP5 to maintain the integrity of the whole complex. The functions of other core components in the MLL complex remain largely unknown.

## MLL/SET1 Family HMTs

To date, more than 10 HMTs have been reported for methylating H3 K4, a scenario that is dramatically different from yeast, where only one enzyme, ySET1, is present [14,15]. Among the mammalian H3 K4 HMTs, six belong to MLL family HMTs: SET1a and SET1b, the mammalian orthologues of yeast SET1 (ySET1), and four MLLs (MLL1-4), which share limited homology with ySET1 beyond the SET domain. [1,16-19]. Multiplicity of MLL family HMTs in higher eukaryotes reflects functional specialization, as indicated by distinct phenotypes, when they were knocked out in mice [9,20,21]. It was shown that like MLL1, other MLL/SET1 family members are generally required for key developmental programs, and their deregulation often leads to malignant transformation. A common feature of MLL/SET1 family HMTs is the conserved core configuration, including RbBP5, Ash2L, and WDR5 [11,17,21-23], and it likely contains the essential information underlying the shared enzymatic activities and substrate specificities. Thus, any knowledge gained from the mechanistic studies for MLL can be extrapolated to the whole MLL/SET1 family HMTs in mammals [11] (Table 1).

## Unique Structure for the MLL SET Domain

Most histone lysine methylation is catalyzed by the evolutionarily conserved SET domain [24-26]. In most cases, the SET domain is fully active in catalyzing methyl-transfer reactions. Biochemical and structural studies revealed that SET domains usually adopt structures featuring a narrow hydrophobic channel, that links the substrate lysine and cofactor S-adenosyl-L-methionine (SAM). Notable exceptions to the rule are the MLL/SET1 family HMTs and the polycomb group protein EZH2, which are fully active only in the context of the complexes. Co-crystal structure of the MLL SET domain in complex with cofactor

**\*Corresponding author:** Fang Cao, Department of Pathology, University of Michigan, USA, E-mail: fangc@med.umich.edu

| Years Set1 Complex | MLL complex | MLL2 complex | MLL3/4 complex | hSET1 complex |
|---|---|---|---|---|
| Set1 | MLL1 | MLL2 | Mll3 | hSet1 |
| Bre2 Spp1 } → | Ash2L | Ash2L | Ash2L | Ash2L CFP1 |
| Swd1 → Swd2 | RbBP5 | RbBP5 | RbBP5 | RbBP5 hSwd2 |
| Swd3 Sdc1 | WDR5 | WDR5 hDPY-30 | WDR5 hDPY-30 | WDR5 |
| | HCF1/2 | HCF2 | UTX | HCF1 |
| | menin | menin | PTIP | Sin3A |
| | MOF | Rpb2 | ASC2 PA1 | MEP50 |

**Table 1:** MLL Family H3 K4 Methyltransferase Complexes in Mammali.an cells.

product AdoHcy and an H3 peptide clearly points out why MLLSET is deficient in H3 K4 methylation by itself [27]. In contrast to other SET domains, the active center of MLLSET was fully exposed to solvent, and optimal methyltransferase activity requires additional interactions to stabilize H3-SAM interaction [27]. Thus, MLL core components could enhance MLL methyltransferase activity, by either directly participating in the formation of lysine access channel, or inducing conformational change in MLLSET. Future structure studies for the MLL core complex are needed for providing the exact molecular mechanism.

Over the past several decades, targeted therapies for cancer have focused on kinases, including major successes with inhibitors developed against ABL, RAF, ALK and EGFR family kinases. Accumulating evidence suggests that epigenetic pathways play an integral role in the sequential progression of cancer. It is conceivable that the HMTs are approached as viable therapeutic targets in cancer. The previous efforts for developing histone methyltransferase inhibitor have mostly used high throughput screening for compounds that inhibit the catalytic SET domain [28,29]. One potential limit for such an approach is that the SET domains are highly homologous among all lysine methyltransferases, regardless of their substrate specificity. Therefore, inhibitors obtained from direct enzymatic screen often have less specificity and potentially more toxicity for cells. Small molecular inhibitors targeting MLL could be used to help define the H3 K4 methylation machinery *in vitro* and *in vivo*, and could be used as possible drugs in situations where H3 K4 methylation is misregulated.

## References

1. Dou Y, Milne TA, Tackett AJ, Smith ER, Fukuda A, et al. (2005) Physical association and coordinate function of the H3 K4 methyltransferase MLL1 and the H4 K16 acetyltransferase MOF. Cell 121: 873-885.

2. Milne TA, Briggs SD, Brock HW, Martin ME, Gibbs D, et al. (2002) MLL targets SET domain methyltransferase activity to Hox gene promoters. Mol Cell 10: 1107-1117.

3. Thiel AT, Blessington P, Zou T, Feather D, Wu X, et al. (2010) MLL-AF9-induced leukemogenesis requires coexpression of the wild-type Mll allele. Cancer Cell 17: 148-159.

4. Milne TA, Martin ME, Brock HW, Slany RK, Hess JL (2005) Leukemogenic MLL fusion proteins bind across a broad region of the Hox a9 locus, promoting transcription and multiple histone modifications. Cancer Res 65: 11367-11374.

5. Dou Y, Hess JL (2008) Mechanisms of transcriptional regulation by MLL and its disruption in acute leukemia. Int J Hematol 87: 10-18.

6. Hsieh JJ, Ernst P, Erdjument-Bromage H, Tempst P, Korsmeyer SJ (2003) Proteolytic cleavage of MLL generates a complex of N- and C-terminal fragments that confers protein stability and subnuclear localization. Mol Cell Biol 23: 186-194.

7. Prasad R, Yano T, Sorio C, Nakamura T, Rallapalli R, et al. (1995) Domains with transcriptional regulatory activity within the ALL1 and AF4 proteins involved in acute leukemia. Proc Natl Acad Sci U S A 92: 12160-12164.

8. Terranova R, Agherbi H, Boned A, Meresse S, Djabali M (2006) Histone and DNA methylation defects at Hox genes in mice expressing a SET domain-truncated form of Mll. Proc Natl Acad Sci U S A 103: 6629-6634.

9. Yu BD, Hess JL, Horning SE, Brown GA, Korsmeyer SJ (1995) Altered Hox expression and segmental identity in Mll-mutant mice. Nature 378: 505-508.

10. Yagi H, Deguchi K, Aono A, Tani Y, Kishimoto T, et al. ( 1998) Growth disturbance in fetal liver hematopoiesis of Mll-mutant mice. Blood 92: 108-117.

11. Dou Y, Milne TA, Ruthenburg AJ, Lee S, Lee JW, et al. (2006) Regulation of MLL1 H3K4 methyltransferase activity by its core components. Nat Struct Mol Biol 13: 713-719.

12. Patel A, Dharmarajan V, Cosgrove MS (2008) Structure of WDR5 bound to mixed lineage Leukemia protein-1 peptide. J Biol Chem 283: 32158-32161.

13. Song JJ, Kingston RE (2008) WDR5 interacts with mixed lineage Leukemia (MLL) protein via the histone H3-binding pocket. J Biol Chem 283: 35258-35264.

14. Shilatifard A (2006) Chromatin modifications by methylation and ubiquitination: implications in the regulation of gene expression. Annu Rev Biochem 75: 243-269.

15. Martin C, Zhang Y (2005) The diverse functions of histone lysine methylation. Nat Rev Mol Cell Biol 6: 838-849.

16. Nakamura T, Mori T, Tada S, Krajewski W, Rozovskaia T, et al. (2002) ALL-1 is a histone methyltransferase that assembles a supercomplex of proteins involved in transcriptional regulation. Mol Cell 10: 1119-1128.

17. Hughes CM, Rozenblatt-Rosen O, Milne TA, Copeland TD, Levine SS, et al. ( 2004) Menin associates with a trithorax family histone methyltransferase complex and with the hoxc8 locus. Mol Cell 13: 587-597.

18. Goo YH, Sohn YC, Kim DH, Kim SW, Kang MJ, et al. (2003) Activating signal cointegrator 2 belongs to a novel steady-state complex that contains a subset of trithorax group proteins. Mol Cell Biol 23: 140-149.

19. Wysocka J, Myers MP, Laherty CD, Eisenman RN, Herr W (2003) Human Sin3 deacetylase and trithorax-related Set1/Ash2 histone H3-K4 methyltransferase are tethered together selectively by the cell-proliferation factor HCF-1. Genes Dev 17: 896-911.

20. Glaser S, Schaft J, Lubitz S, Vintersten K, van der Hoeven F, et al. ( 2006) Multiple epigenetic maintenance factors implicated by the loss of Mll2 in mouse development. Development 133: 1423-1432.

21. Lee S, Lee DK, Dou Y, Lee J, Lee B, et al. ( 2006) Coactivator as a target gene specificity determinant for histone H3 lysine 4 methyltransferases. Proc Natl Acad Sci U S A 103: 15392-15397.

22. Lee JH, Skalnik DG (2005) CpG-binding protein (CXXC finger protein 1) is a component of the mammalian Set1 histone H3-Lys4 methyltransferase complex, the analogue of the yeast Set1/COMPASS complex. J Biol Chem 280: 41725-41731.

23. Yokoyama A, Wang Z, Wysocka J, Sanyal M, Aufiero DJ, et al. ( 2004) Leukemia proto-oncoprotein MLL forms a SET1-like histone methyltransferase complex with menin to regulate Hox gene expression. Mol Cell Biol 24: 5639-5649.

24. Dillon SC, Zhang X, Trievel RC, Cheng X (2005) The SET-domain protein superfamily: protein lysine methyltransferases. Genome Biol 6: 227.

25. Yeates TO (2002) Structures of SET domain proteins: protein lysine methyltransferases make their mark. Cell 111: 5-7.

26. Qian C, Zhou MM (2006) SET domain protein lysine methyltransferases: Structure, specificity and catalysis. Cell Mol Life Sci 63: 2755-2763.

27. Southall SM, Wong PS, Odho Z, Roe SM, Wilson JR (2009) Structural basis for the requirement of additional factors for MLL1 SET domain activity and recognition of epigenetic marks. Mol Cell 33: 181-191.

28. Chang Y, Ganesh T, Horton JR, Spannhoff A, Liu J, et al. ( 2010) Adding a lysine mimic in the design of potent inhibitors of histone lysine methyltransferases. J Mol Biol 400: 1-7.

29. Kubicek S, O'Sullivan RJ, August EM, Hickey ER, Zhang Q, et al. ( 2007) Reversal of H3K9me2 by a small-molecule inhibitor for the G9a histone methyltransferase. Mol Cell 25: 473-481.