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 affects 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 MLLL 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 coordinate 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 holocomplex [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, sSET1, 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. [1,6,16]. 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, protein EZH2, which are fully active only in the context of the complexes. Furthermore, 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.

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.
methylation is misregulated. In vivo, be used to help define the H3 K4 methylation machinery from direct enzymatic screen often have less specificity and potentially regardless of their substrate specificity. Therefore, inhibitors obtained domains are highly homologous among all lysine methyltransferases, domain [28,29]. One potential limit for such an approach is that the SET approached as viable therapeutic targets in cancer. The previous efforts sequential progression of cancer. It is conceivable that the HMTs are evidence suggests that epigenetic pathways play an integral role in the focused on kinases, including major successes with inhibitors developed 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