ASH1L Links Histone H3 Lysine 36 di-methylation to MLL Leukemia

Li Zhu¹, Qin Li², Stephen H.K. Wong¹, Min Huang³, Brianna J. Klein⁴, Jinfeng Shen⁵, Larissa Ikenouye⁴, Masayuki Onishi², Dominik Schneidawind⁶, Corina Buechele¹, Loren Hansen², Jesús Duque-Afonso¹, Fangfang Zhu⁷, Gloria Martin⁸, Or Gozani⁸, Ravindra Majeti⁵, Tatiana G. Kutateladze⁴, and Michael L. Cleary¹*

¹ Department of Pathology, Stanford University School of Medicine, Stanford, CA
² Department of Genetics, Stanford University School of Medicine, Stanford, CA
³ Department of Medicine, Division of Oncology, Stanford University School of Medicine, Stanford, CA
⁴ Department of Pharmacology; University of Colorado School of Medicine; Aurora, CO
⁵ Department of Medicine, Division of Hematology, Cancer Institute, and Institute for Stem Cell Biology and Regenerative Medicine, Stanford University School of Medicine, Stanford, CA
⁶ Department of Medicine, Division of Blood and Marrow Transplantation, Stanford University School of Medicine, Stanford, CA
⁷ Departments of Pathology and Developmental Biology, and Institute for Stem Cell Biology and Regenerative Medicine, Stanford University School of Medicine, Stanford, CA
⁸ Department of Biology, Stanford University, Stanford, CA
⁹ These authors contributed equally.

Running title: Requirement for ASH1L and H3K36me2 in MLL leukemia

Keywords: Histone methylation; leukemia; ASH1L; MLL; LEDGF

Financial support: This work was supported by NIH grants GM106416 (T.G.K.), GM079641 (O.G.) and CA116606 (M.L.C.). J. D-A. was supported by the German Research Foundation (Deutsche Forschungsgemeinschaft, ref. DU 1287/2-1). C.B. and D.S. were supported by the Dr. Mildred Scheel Stiftung. We also acknowledge support from the Children’s Health Initiative of the Packard Foundation, the California Institute for Regenerative Medicine, and the Child Health Research Institute.
*Correspondence:
Michael L. Cleary
Department of Pathology
Stanford University School of Medicine
300 Pasteur Drive
Stanford, CA 94305
mcleary@stanford.edu
Ph: 650-723-5471

Conflict of interest: The authors have no conflicts of interest to declare.
ABSTRACT

Numerous studies in multiple systems support that histone H3 lysine 36 di-methylation (H3K36me2) is associated with transcriptional activation, however the underlying mechanisms are not well defined. Here we show that the H3K36me2 chromatin mark written by the ASH1L histone methyltransferase is preferentially bound in vivo by LEDGF, an MLL-associated protein that co-localizes with MLL, ASH1L and H3K36me2 on chromatin genome wide. Furthermore, ASH1L facilitates recruitment of LEDGF and wild type MLL proteins to chromatin at key leukemia target genes, and is a crucial regulator of MLL-dependent transcription and leukemic transformation. Conversely KDM2A, an H3K36me2 demethylase and Polycomb-group silencing protein, antagonizes MLL-associated leukemogenesis. Our studies are the first to provide a basic mechanistic insight into epigenetic interactions wherein placement, interpretation and removal of H3K36me2 contribute to the regulation of gene expression and MLL leukemia, and suggest ASH1L as a novel target for therapeutic intervention.
SIGNIFICANCE

Epigenetic regulators play vital roles in cancer pathogenesis and represent a new frontier in therapeutic targeting. Our studies provide basic mechanistic insight into the role of histone H3 lysine 36 di-methylation in transcription activation and MLL leukemia pathogenesis, and implicate ASH1L histone methyltransferase as a promising target for novel molecular therapy.
INTRODUCTION

Modulation of chromatin structure through covalent histone modifications serves a major role in normal and pathologic gene expression. A significant subset of modifications involves lysine methylation, which is a dynamic and reversible process regulated by two classes of counteracting enzymes: histone lysine methyltransferases and demethylases (1). Methyl groups on histone lysine residues serve as docking sites for reader proteins that bind specific modifications and translate the histone code into various transcriptional outputs (2). Among the diverse family of histone methyltransferases, TRX (trithorax) and ASH1 (absent, small, or homeotic discs 1) are distinguished as trithorax group (trxG) proteins that positively regulate gene transcription and counteract Polycomb group (PcG) mediated silencing as demonstrated by seminal genetic studies of Drosophila development (3). Furthermore, their mammalian homologs MLL (mixed lineage leukemia) and ASH1L (ASH1-like) are expressed in hematopoietic stem and progenitor cells (HSPCs), regulate HOX gene expression, and support the proliferation and/or survival of hematopoietic stem cells (4-7). ASH1L genetically cooperates with MLL to maintain hematopoiesis, as their compound deficiency results in more severe hematopoietic failure than knockout of either alone (7). At a biochemical level, MLL and ASH1L methylate lysine 4 and lysine 36, respectively, of histone H3 (8, 9), however the underlying molecular mechanisms for their synthetic genetic interactions are unknown.

Mis-regulation of histone methylation caused by perturbations of epigenetic writers, readers or erasers results in aberrant gene expression that can lead to the induction or progression of human cancers. Notably, the MLL gene is a frequent target of chromosomal translocations in acute leukemias (10). In leukemogenic translocations, MLL loses its methyltransferase activity and acquires novel epigenetic functionalities due to fusions with various partner proteins (10-13).
Furthermore, *MLL*-oncogene mediated leukemia retains a dependence on several MLL binding partners, including LEDGF (lens epithelium-derived growth factor) (14), a chromatin-associated protein implicated in cancer, autoimmunity, and HIV pathogenesis (15). LEDGF serves an essential role in MLL-dependent transcription and leukemic transformation (14) mediated in part through a conserved PWWP domain that is also present in a variety of chromatin-associated proteins and recently reported to bind methylated histone H3K36 (16-20) with a focus on H3K36me3, suggesting that LEDGF is a histone code reader. H3K36 methylation is associated with transcriptional activation and elongation, underscored by a progressive shift from H3K36 di-methylation at promoters to tri-methylation at the 3’ ends of actively transcribed genes (21). Although associated with actively expressed genes, H3K36me3 has been reported to recruit histone deacetylase Rpd3S and transcription co-repressor ZMYND11 in yeast and human, respectively, to maintain a repressive chromatin environment during transcriptional elongation (22, 23). However, the underlying mechanism for how H3K36 di-methylation may directly contribute to transcription activation is largely unknown.

Here we show that ASH1L maintains the H3K36me2 mark, which is preferentially read by LEDGF to recruit/stabilize MLL on promoter proximal chromatin to activate crucial target gene expression in leukemia cells. All of these molecular events, and leukemogenesis, are specifically antagonized by the histone demethylase KDM2A. Furthermore, H3K36me3 is not required for LEDGF and MLL occupancy, rather its reduction actually enhances recruitment of LEDGF and MLL to chromatin. Our studies reveal ASH1L, LEDGF and KDM2A as crucial writers, readers and erasers, respectively, of H3K36 di-methylation underlying MLL leukemia pathogenesis, illuminate the molecular mechanism by which H3K36 di-methylation directly
contributes to transcription activation, and identify ASH1L as a novel target for molecular therapy of acute leukemia.
RESULTS

LEDGF Preferentially Binds Histone H3K36me2 in the Context of Nucleosomes in vivo

The histone binding properties of LEDGF were initially assessed using NMR spectroscopy, which showed chemical-shift changes in the PWWP domain following titration of H3K36me3 and H3K36me2 peptides (Figures 1A and S1A). However, high dissociation constants (e.g. 4.4 mM for the H3K36me3 complex) (Figure S1B) suggested a weak interaction consistent with our inability to detect interactions in peptide pull-down assays, which are less sensitive than NMR (Figure S1C, D). Conversely, chromatin pull-down assays showed that recombinant LEDGF bound robustly to native mono-nucleosomes purified from Hela cells (Figure 1B, S1E). Nucleosome interaction was abolished by mutations (W21A or F44A) of conserved aromatic residues in the hydrophobic cavity of the PWWP domain (Figure 1B). In pull-down assays using nucleosomes purified from yeast strains containing different histone lysine mutations (Table S1), nucleosomes containing a K36A mutation of histone H3 were not bound by GST-LEDGF, indicating that H3K36 is required for the interaction of LEDGF with nucleosomes (Figure 1C, S1E).

A protein-protein chromatin immunoprecipitation assay using THP-1 leukemia cells showed that wild type LEDGF associated strongly with nucleosomes containing H3K36me2 in the context of physiologic chromatin, but only weakly with H3K36me1 or H3K36me3, and not at all with other methylated histones, suggesting that H3K36me2 is the highly preferred in vivo target of LEDGF (Figure 1D and S1F). The H3K36me2 interaction was mostly abolished by the W21A mutation that disrupted nucleosome binding in vitro (Figure 1D). Thus, LEDGF preferentially binds dimethylated histone H3K36 in the context of nucleosomes in vivo.
**LEDGF, MLL and H3K36me2 Co-occupy MLL Targets Genome-wide.**

Chromatin immunoprecipitation coupled with high-throughput sequencing (ChIP-seq) was performed to determine the genome-wide occupancy of LEDGF, and the extent of its overlap with H3K36me2 and MLL in the MV4-11 human leukemia cell line with *MLL-AF4* translocation. Approximately 89% of LEDGF-bound genes genome-wide overlapped with MLL and H3K36me2 co-occupied genes. High LEDGF and H3K36me2 occupancies were observed in genes enriched with MLL (Figure 2A). The average distribution of all LEDGF-bound peaks recapitulated the genomic distribution of MLL and H3K36me2 (Figure 2B) peaking around the transcription start site (TSS), in contrast to H3K36me3 distribution, which peaked at the transcription termination site (TTS). High MLL occupancies were observed in genes enriched with LEDGF, whereas MLL occupancies were much lower in genes lacking LEDGF enrichment (Figure 2C). Comparison of ChIP-seq peaks on selected MLL target genes revealed nearly complete overlap in regional occupancy for MLL, LEDGF and H3K36me2 around transcription start sites (Figure 2D), which was further confirmed by ChIP-qPCR analyses (Figure S2A). These results indicate that LEDGF co-localizes with MLL and H3K36me2 around the transcription start sites of MLL target genes.

**MLL and LEDGF Association with MLL Target Genes is Mutually Dependent.**

The observed co-occupancies raised the possibility that LEDGF functions as a reader of the histone code to recruit and/or stabilize MLL at H3K36me2-enriched promoter proximal chromatin. To test this, chromatin immunoprecipitation (ChIP) assays were performed in MLL-AF9 transformed hematopoietic stem and progenitor cells (HSPCs) transduced with LEDGF or control shRNAs. LEDGF was down-regulated to 5–10% of control levels (Figure 2E) without
affecting WT MLL and MLL-AF9 expression (Figure S2B). Using an antibody that specifically recognizes WT MLL, but not fusion MLL, ChIP assays demonstrated significantly reduced binding of WT MLL at proximal promoter regions of MLL target genes, and throughout the Hoxa9 locus, upon LEDGF knockdown (Figures 2F and S2C) concordant with substantially reduced transcript levels (Figure 2E). Unexpectedly, occupancy of the MLL fusion protein (MLL-AF9) was consistently increased at the respective target loci in ChIP assays using either an anti-AF9 antibody or an anti-Flag antibody to detect MLL-AF9 or Flag-tagged MLL-AF9, respectively (Figures 2F and S2C, D). Furthermore, a missense mutant of MLL-AF9 (F129A) that cannot interact with LEDGF (14) retained an ability to associate with MLL target genes (Figure S2E). Decreased occupancies of WT MLL and increased occupancies of MLL fusion proteins were also observed for MLL-AF10 and MLL-ENL in LEDGF knockdown cells (Figure S2D). These results indicate that LEDGF is required for retention of WT MLL, but not MLL fusion proteins, at target gene loci in MLL-transformed HSPCs. Consistent with these results, MLL oncogene mediated leukemogenesis is critically dependent on the WT MLL allele (24).

Despite reduced occupancy of WT MLL at target gene loci following LEDGF knockdown, H3K4me3 levels were not altered (Figure S2F). This is consistent with previous studies showing that knockout of MLL in HSPCs has no effect on H3K4 methylation at target genes, and the histone methyltransferase activity of MLL is dispensable for leukemogenesis (25). Rather, MLL regulates target gene expression by recruitment of acetyltransferase MOF, which forms a stable complex with WT MLL but not MLL fusion proteins and acetylates chromatin at histone H4 lysine 16 to recruit the BRD4/pTEFb complex and facilitate transcriptional elongation (25-27). Notably, histone H4K16ac levels were reduced at Hoxa9 and Meis1 loci in
LEDGF knockdown cells, and the chromatin occupancies of BRD4, P-TEFb and elongating POL II (serine 2 phosphorylated) were substantially decreased (Figure 2F).

The foregoing results raised questions regarding how LEDGF may affect MLL fusion protein functions that are essential for mis-regulation of MLL target genes and MLL-induced transformation. Several translocation partners of MLL, including AF9, coexist in higher-order protein complexes (e.g. AEP or SEC), which contain known transcription elongation factors such as AF4 and P-TEFb (28, 29). MLL oncoproteins fused with AEP components constitutively form MLL/AEP hybrid complexes to cause sustained target gene expression, which leads to transformation of HSPCs. To investigate whether LEDGF plays a role in the formation of MLL/AEP complexes on chromatin, ChIP assays were performed for AEP components AF4 and CDK9 at the Hoxa9 and Meis1 loci in MLL-AF9 transformed HSPCs (Figure 2F). AF4 and CDK9 occupancies were significantly reduced upon LEDGF knockdown, suggesting that recruitment of the components of MLL fusion-AEP complexes at target genes is dependent on LEDGF, although LEDGF is not required for retention of MLL fusion proteins on chromatin (Figure 2F).

The direct interaction and genome-wide co-occupancy of MLL and LEDGF raised the possibility that LEDGF chromatin binding is MLL-dependent. To test this, occupancy of LEDGF was investigated by ChIP assays in MLL-AF9 transformed HSPCs transduced with MLL or control shRNAs. Knockdown of MLL (Figure 2G) resulted in reduced MLL and LEDGF binding at proximal promoter regions of MLL target genes (Figure 2H), indicating that MLL promotes and/or stabilizes LEDGF chromatin binding. Thus, LEDGF and MLL target association is mutually dependent, which may account for their promoter proximal co-occupancies.
ASH1L Promotes Recruitment of LEDGF and MLL to Chromatin.

Given the oncogenic requirement for LEDGF (14) and its ability to selectively bind H3K36me2 in the context of nucleosomes, we assessed the potential role for H3K36 di-methylation in MLL-transformed cells. Several histone methyltransferases mediate H3K36 di-methylation (21) most notably ASH1L, which genetically cooperates with MLL to maintain hematopoiesis (7).

ChIP-seq of MV4-11 leukemia cells showed that ASH1L displayed overlapping distribution profiles with MLL, LEDGF and H3K36me2 around transcription start sites at MLL target genes (Figure 3A, B, and S3A-B), which was further confirmed by ChIP-qPCR analyses of selected MLL target genes (Figure S3C). Gene set enrichment analysis (GSEA) showed that the subset of genes co-occupied by MLL, LEDGF and ASH1L were highly enriched in genes with elevated mRNA expression in human MLLr ALL versus non-MLLr ALL (Figure 3C). Thus, ASH1L co-localizes with H3K36me2, MLL and LEDGF around the transcription start sites of MLL target genes whose expression highly distinguishes MLLr human leukemias.

Knockdown of ASH1L in MLL-transformed HSPCs (Figure 3D and S3D) resulted in reduced H3K36me2 but not H3K36me3 levels at MLL target gene promoters (Figure 3E and S3E), which correlated with reduced LEDGF binding (Figure 3F) and \textit{Hoxa9} transcript levels (Figure 3D). ASH1L knockdown had no effect on global levels of H3K36me2 and H3K36me3 (Figure S3F). Occupancy of WT MLL, but not fusion MLL, was reduced at key target gene promoters and throughout the \textit{Hoxa9} locus (Figure 3F and S3G), and ASH1L was required for AEP complex recruitment, maintenance of H4K16ac levels, and chromatin occupancies of BRD4 and elongating POL II (Figure 3F). These results phenocopied those observed for LEDGF knockdown, consistent with its specific binding to nucleosomes marked by H3K36me2.
Thus, ASH1L maintains histone H3 lysine 36 dimethylation and promotes LEDGF and MLL recruitment at key target genes.

Activating mutations in NSD2, an H3K36 dimethyltransferase, were found in several pediatric ALL subtypes (30, 31), raising a question whether NSD2 plays any role in LEDGF and MLL recruitment. Efficient knockdown of NSD2 in MLL-AF10 transformed HSPCs resulted in modest reductions of LEDGF and MLL occupancy at the *Hoxa9* gene promoter, but not *Meis1* gene promoter (Figure S3H), which correlated with modest reduction of *Hoxa9*, but not *Meis1* transcript levels (Figure S3I). This contrasted with ASH1L knockdown, which caused major, consistent reduction of LEDGF and MLL recruitment at MLL target genes.

**ASH1L is an Essential Regulator of MLL-mediated Leukemogenesis.**

The requirement for ASH1L and H3K36 di-methylation in MLL-mediated oncogenesis was initially assessed in MLL-transformed HSPCs. Cells transduced with shRNAs targeting ASH1L (Figure 4A) showed morphologic and phenotypic differentiation, and underwent apoptosis (Figure 4B-D and S4A, B). ASH1L knockdown substantially compromised the in vitro growth and clonogenic activity of HSPCs transformed by *MLL* oncogenes, but did not affect cells transformed by a non-*MLL* oncogene (Figure 4E). In transplantation assays, MLL-AF10 transformed HSPCs knocked down for ASH1L showed a substantial impairment in their ability to induce leukemia as evidenced by a marked increase in disease latency and reduced penetrance (Figure 4F).

The role of ASH1L in MLL-mediated leukemogenesis was further investigated in human leukemia cells. Knockdown of ASH1L resulted in reduced *HOXA9* expression in MLL-rearranged (MLLr) leukemia cell lines (Figure 5A), which correlated with reduced growth in
methylcellulose and liquid cultures, whereas non-MLLr leukemia cells were unaffected (Figure 5B, C). Xenotransplantation studies showed that knockdown of ASH1L markedly impaired bone marrow engraftment efficiency of MV4-11 cells in NSG mice by over 90% compared to control (sh-Luc) cells at 4 weeks (Figure 5D). Similarly, xenotransplantation studies of primary MLLr AMLs showed that knockdown of ASH1L substantially reduced engraftment efficiency (~80%) compared to control at 10 weeks post-transplant (Figure 5E). Thus, ASH1L is required for MLL oncogene-induced leukemogenesis.

**Reduced H3K36me3 Increases Recruitment of LEDGF and MLL to Chromatin.**

In protein–protein ChIP assays, LEDGF associated preferentially with H3K36me2, which is essential for MLL-mediated oncogenesis, and weakly with H3K36me3 (Figure 1D). The potential contribution of H3K36 tri-methylation to MLL-mediated transformation was assessed by knocking down SETD2, the only known histone H3K36me3 methyltransferase (21). SETD2 knockdown in MLL-AF9-transformed HSPCs (Figure 6A) reduced global levels of H3K36me3 (Figure 6B). Compared to ASH1L knockdown, SETD2 knockdown had minor effects on Hoxa9 gene expression (Figure 6A), and clonogenic activity of MLL-transformed HSPCs (Figure 6C). SETD2 plays an important role in maintaining genome stability (32), which may account for partially reduced leukemia cell growth.

ChIP assays showed that SETD2 knockdown resulted in reduction of H3K36me3 levels at MLL target genes (Figure 6D). In contrast, H3K36me2 was increased, which correlated with increased LEDGF and MLL occupancy, further supporting the key role of H3K36me2 in MLL/LEDGF recruitment. Fusion MLL was maintained at key target gene promoters in SETD2
Zhu et al.

knockdown cells (Figure 6D). Thus, SETD2 and H3K36 tri-methylation do not directly regulate MLL oncogene-induced leukemogenesis.

**KDM2A Antagonizes MLL Oncogene Induced Leukemogenesis.**

The role of reduced H3K36me2 through forced demethylation was also assessed in MLL-mediated transformation. Several families of histone demethylases demethylate H3K36me2 in mammalian cells (1). Among them, the KDM2 family is of particular interest since genetic evidence indicates antagonism between ASH1 and dKDM2 in Drosophila (33). In humans, two paralogs (KDM2A and KDM2B) preferentially demethylate H3K36me2, and repress transcription (1, 34).

Over-expression of KDM2A, but not KDM2B, in MLL-AF10-transformed HSPCs resulted in reduced Hoxa9 expression (Figure 7A and S5A). KDM2A over-expression also resulted in reduced H3K36me2 but not H3K36me3 levels at MLL target genes, and reduced occupancy of LEDGF and WT MLL, but not fusion MLL (Figure 7B). ASH1L binding to MLL target genes was not significantly affected by forced KDM2A expression (Figure S5B). KDM2A over-expression had no effect on global levels of H3K36me2 and H3K36me3 (Figure S5C). Thus, KDM2A demethylates H3K36me2 and promotes the chromatin dissociation of MLL and LEDGF concomitant with reduced expression of MLL target genes.

Over-expression of KDM2A, but not KDM2B, induced monocytic differentiation of MLL-AF10 transformed cells (Figures S5D), and markedly reduced CFC activity of MLL-transformed cells, whereas cells transformed by E2A-HLF, which acts through non-HOX pathways, were unaffected (Figure 7C, D). Forced expression of KDM2A also antagonized the leukemogenic properties of MLL-AF10-transformed HSPCs causing increased latency and
reduced penetrance in transplant assays (Figure 7E). These results demonstrate that KDM2A specifically opposes MLL-induced leukemia, further supporting the role for H3K36me2 in MLL-mediated oncogenesis.

**DISCUSSION**

The epigenetic pathways and mechanisms involved in MLL leukemia pathogenesis are multiple and complex, and the effectiveness of early stage targeted molecular therapies not yet established. The recent discovery of BET inhibitor resistance further underscores the need for discovery of new drug targets and development of combined therapies (35, 36). Our studies are the first to define the epigenetic interactions involving placement, interpretation, and removal of H3K36me2 underlying MLL leukemia pathogenesis. Furthermore, our studies are particularly notable for implicating ASH1L as a promising target for novel molecular therapy in addition to providing important mechanistic insight into the role of H3K36 di-methylation in transcription.

**ASH1L, LEDGF and MLL cooperatively regulate target gene expression and leukemogenesis.**

Methyl lysine residues ‘written’ by histone methyltransferases are recognized and bound by ‘reader’ proteins that recruit various components of the transcription complex to translate the histone code into specific gene expression states (2). In yeast, H3K36me3 recruits histone deacetylase Rpd3S to reset the coding region of actively transcribed genes to a relatively repressive state, thus suppressing intragenic transcription initiation (23). In human, H3K36me3 recruits transcription co-repressor ZMYND11 to highly expressed genes to fine-tune transcription by modulating RNA polymerase II at the elongation stage (22). It remains
unknown how H3K36me2 may directly contribute to activation. Our studies answer this longstanding question and support a model in which ASH1L writes the H3K36me2 mark, which is read by LEDGF to recruit/stabilize MLL on chromatin to maintain crucial target gene expression in leukemia cells. These findings delineate the functional inter-relationships of two lysine histone methyltransferases involved in MLL leukemia pathogenesis and illuminate their genetic interactions as evolutionary conserved trithorax group proteins. Together they counteract Polycomb group mediated silencing, to regulate cell fate through maintaining homeotic gene expression (3). This provides a molecular context for the conserved interactions of trxG proteins in development, hematopoiesis and leukemogenesis.

**H3K36me2, not H3K36me3, serves a key role in MLL leukemia.**

The role of LEDGF as a histone code reader that mediates chromatin association of MLL is consistent with previous studies showing that its contribution in MLL leukemia requires its conserved PWWP domain (14, 17). The latter has recently been reported to bind H3K36me2/3 (16-19) with a focus on H3K36me3, however our data strongly support H3K36me2 as the biologically relevant target. The differences of our results with previous studies (18, 37) are likely accounted for by prior use of ChIP conditions not efficiently precipitating the MLL-associated form of LEDGF (18) and more limited genomic assessment using customized arrays covering only *Hox* loci and several other developmental genes (18, 37). Although both the tri- and di-H3K36 methylation histone marks have been implicated in transcription activation (21), LEDGF preferentially binds H3K36me2 in the context of physiologic chromatin in human leukemia cells. In support of this, LEDGF and MLL co-localize with H3K36me2 peaking at the transcription start site, in contrast to H3K36me3 distribution, which peaks at the transcription
termination site. Furthermore, knockdown of SETD2, the only known H3K36 trimethyltransferase, reduced H3K36me3 levels genome-wide and particularly at MLL target genes, whereas H3K36me2 was respectively increased, which correlated with increased LEDGF and MLL occupancy. Consistent with the lack of a role for SETD2 and H3K36me3 in MLL-dependent transcription, SETD2 inactivating mutations occur with a significantly higher frequency in patients with MLL-rearranged leukemia and contribute to leukemogenesis through a global reduction of H3K36me3 (38, 39). Our results indicate that a molecular consequence of prevalent SETD2 loss-of-function mutations in MLL leukemias is increased occupancy of H3K36me2, LEDGF and MLL at MLL target genes. Thus, H3K36me3 is not essential for MLL occupancy and leukemogenesis, rather its relative reduction actually enhances molecular features that sustain MLL leukemia.

**The different roles of LEDGF in leukemogenic activity of MLL wild type versus fusion proteins.**

An unexpected observation is that ASH1L and LEDGF are not required for the chromatin association of MLL fusion proteins at crucial leukemia target genes. In fact, depletion of either ASH1L or LEDGF resulted in increased occupancy of MLL fusion proteins suggesting a likely competition for chromatin sites vacated by WT MLL. Both WT MLL and MLL-fusion proteins are required for sustained expression of genes that are essential for leukemogenesis (24), however it is not fully understood how they are recruited to chromatin. The differential dependence on LEDGF for occupancy at target gene loci in MLL-transformed cells indicates that WT and fusion MLL proteins have different recruitment mechanisms. In support of this, the recruitment of WT MLL requires both its CXXC/RD1 domain and PHD3 finger, which mediate
direct interactions with the PAF1 elongation complex and H3K4me3, respectively. However, the CXXC/RD1 domain alone, which binds CpG-rich DNA as well as PAF1, is sufficient for the recruitment of MLL fusion proteins (40). Thus, WT MLL has more complex requirements for chromatin association including a selective dependence on LEDGF and H3K36me2, which appear to be bypassed by MLL fusion proteins. Although LEDGF is not required for the chromatin occupancy of MLL fusion proteins, their leukemogenic activity is critically dependent on LEDGF with which they specifically interact (14). This raised a question regarding whether LEDGF has a recruitment-independent role in modulating MLL-fusion protein functions in their roles as components of aberrant AEP/SEC complexes, which contain transcription elongation factors including MLL fusion partners essential for leukemia. Our data show that the chromatin association of AEP/SEC components AF4 and CDK9 is significantly reduced upon LEDGF knockdown, suggesting that the recruitment of components of the fusion protein complex at target genes is dependent on LEDGF, although LEDGF is not required for MLL fusion protein retention on chromatin.

ASH1L is a novel target for therapeutic intervention in acute leukemia.

The dependence on ASH1L establishes it as a candidate target for molecular therapy of MLLr acute leukemias, which are generally associated with a poor prognosis (10). Our results show that ASH1L is particularly enriched at a subset of genes (e.g. HOXA9, MEIS1, and CDK6) that are differentially expressed in MLLr leukemias and crucial for leukemia pathogenesis. Their constitutive expression is mediated by the combined actions of MLL WT and fusion proteins (24), and targeting either factor effectively antagonizes MLL leukemia. Although small molecule inhibitors are not yet available, genetic studies suggest that ASH1L inhibition may not
be unmanageably toxic. Homozygous ASH1L mutation was reported to result in decreased LT-HSC numbers, however increased self-renewal of progenitors compensated for HSC loss and sustained relatively normal mature hematopoietic cell output (7). Partial reduction in ASH1L activity shows greater cytotoxicity for MLL leukemia cells defining it as a selective target for therapeutic intervention of leukemia. Future studies are warranted when inhibitors are developed to further assess the efficacy of targeting ASH1L as a therapeutic strategy in MLLr leukemia and possibly other cancer types dependent on elevated HOX gene expression.

**KDM2A counteracts ASH1L in MLL oncogene induced leukemogenesis.**

Maintenance of HOX gene expression and MLL oncogene-induced leukemogenesis are opposed by the histone code ‘eraser’ KDM2A, a demethylase that counteracts the actions of ASH1L. This parallels results in Drosophila, where dKDM2 is a component of the dRING-associated factor complex, a Polycomb group silencing complex, and cooperates with Polycomb to counteract homeotic gene activation by trxG histone methyltransferases TRX and ASH1 (33). In humans, KDM2 has two homologues (KDM2A and KDM2B) that demethylate H3K36me2 and repress transcription (41, 42). KDM2A interacts with SUZ12, a component of Polycomb repressive complex 2 (43). Overexpression of KDM2A reduced MLL-dependent transcription and leukemic transformation. KDM2A demethylates H3K36me2 at MLL target genes, and promotes the chromatin dissociation of MLL and LEDGF, elucidating a molecular pathway for how KDM2A counteracts trxG proteins to repress transcription.

The action of KDM2A in suppressing MLL leukemia by opposing ASH1L activity may reflect an analogous role in normal hematopoiesis. KDM2A transcripts are low in HSPCs and increase with myeloid differentiation, which is the inverse of expression profiles for MLL,
LEDGF and ASH1L (Microarray Database of Gene Expression Commons). KDM2A is recruited to non-methylated CpG islands through its CXXC domain, where it removes histone H3K36me2 (41, 44). Our studies show that ASH1L, MLL and LEDGF co-occupy CpG islands at MLL target genes in leukemia cells. Taken together, these findings support a model in which KDM2A initiates differentiation of normal and leukemic myeloid progenitors by binding to non-methylated CpG islands resulting in removal of histone H3K36me2, which in turn promotes the dissociation of MLL and LEDGF, and subsequent silencing of MLL target genes.
METHODS

Cell Culture

Phoenix and 293T cells were cultured in Dulbecco’s modified Eagle’s medium (DMEM) supplemented with 10% FBS and penicillin-streptomycin-glutamine. Human leukemia cell lines were cultured in RPMI-1640 medium supplemented with 10% FBS and penicillin-streptomycin-glutamine. Cell lines were obtained from the ATCC or DSMZ in 2013 and not further authenticated. Cell lines were tested periodically for Mycoplasma by PCR (MD Bioproducts), and results verified by IDEXX RADIL Laboratories.

Chromatin Immunoprecipitation and ChIP-qPCR

ChIP assays using mouse myeloid progenitors transformed by MLL-fusion genes were performed using 0.5-1×10^7 cells per IP under conditions described previously (45). ChIP-qPCR signals were calculated as percent of input. Primer sequences and antibodies used in ChIP assays are provided in Tables S2, S3 and S4.

ChIP-seq

Whole genome ChIP using MV4-11 cells was performed under conditions described previously (45) using 2-4×10^7 cells per IP with specific antibodies (Table S4). ChIP-seq libraries were prepared according to the Illumina protocol and sequenced using Illumina HiSeq. Sequencing reads were mapped to the human reference genome (hg19) using Bowtie with parameters: -v 2 -k 2 -m 1 --best --strata. Clonal reads were removed before enriched peak/region calling with MACS version2.1.0. Enriched regions (fold-change ≥ 5) were filtered with false discovery rate ≤ 0.01 to be considered as statistically significant ChIP peaks. Downstream analyses including
genome-wide localization density and average profile plots were performed by using ngs.plot and R/Bioconductor.

**Leukemogenesis Assays**

MLL-AF10 transformed mouse HSPCs were transduced with control or appropriate shRNAs against LEDGF or ASH1L. Following puromycin selection, cells ($1 \times 10^6$) were transplanted intravenously into lethally irradiated C57BL/6 mice (900 rads) with $2 \times 10^5$ syngeneic bone marrow cells. Similar transplant conditions were employed for leukemogenesis assays using MLL-AF10 transformed mouse HSPCs transduced with KDM2A expression vector. All experiments on mice in this study were performed with the approval of and in accordance with Stanford University’s Administrative Panel on Laboratory Animal Care.

**Xenotransplant Assays**

Xenotransplant assays were performed as previously described (46). MV4-11 human leukemia cells were transduced with control or ASH1L shRNA virus by spinoculation. After 48 hrs, four million cells in 0.2 ml PBS containing 0.5% FBS were transplanted into sublethally irradiated (200 rad) NOD.Cg-$Prkdc^{scid} \text{Il2rg}^{+/-}\text{J}/\text{SzJ}$ (NSG) mice (8-10 week females). After four weeks, the bone marrow was collected and analyzed for the presence of transduced (GFP+) human leukemia cells by flow cytometry.

Human AML samples were obtained from patients at the Stanford University Medical Center with informed consent and institutional review board approval in accordance with the Declaration of Helsinki. Primary AML specimens (MLLr) were thawed and cultured overnight in myelocult medium containing human cytokines (100 ng/ml SCF, 100 ng/ml Flt-3, 50 ng/ml...
TPO, 20 ng/ml IL-3, 20 ng/ml IL-6, and 1 μM SR1). In triplicate, three million cells were
transduced with ASH1L shRNA virus using retronectin (Clontech) according to the
manufacturer’s instructions. After 48 hrs, two million cells in 0.2 ml PBS containing 0.5% FBS
were transplanted into sublethally irradiated (200 rad) NSG mice. After ten weeks, the bone
marrow was collected and analyzed for the presence of transduced (GFP+) human leukemia cells
by flow cytometry.

Expression and Purification of $^{15}$N-labeled LEDGF PWWP

The wild type LEDGF PWWP (residues 1-120) was expressed in E. coli BL21(DE3) pLysS
grown in $^{15}$NH$_4$Cl-supplemented M9-minimal medium. Bacteria were harvested by
centrifugation after IPTG induction and lysed via sonication. The GST-fusion protein was
purified on a glutathione Sepharose 4B column, cleaved with PreScission protease and
concentrated in Millipore concentrators using 25 mM Tris pH 6.8 buffer containing 150 mM
NaCl and 8% $^2$H$_2$O.

NMR Spectroscopy

NMR experiments were performed at 298 K on a Varian INOVA 600 MHz spectrometer using
uniformly $^{15}$N-labeled LEDGF PWWP. The histone binding was characterized by monitoring
chemical shift changes in $^1$H,$^{15}$N H$\alpha$QC spectra of 0.1-0.2 mM PWWP as histone peptides were
titrated in. The spectra were processed with NMRPipe and analyzed with nmrDraw. The
dissociation constant (Kd) was determined by a nonlinear least-squares analysis using a
Kaleidagraph program and the equation: $\Delta \delta = \Delta \delta_{\text{max}}(([L]+[P]+K_D)-\sqrt{\sqrt{([L]+[P]+K_D)-(4*[P]*[L]))}/(2*[P]))$, where [L] is concentration of the peptide, [P] is concentration of PWWP,
Δδ is observed chemical shift change, and Δδmax is the difference in chemical shifts of the free and the ligand-bound protein.

**Nucleosome Purification**

Yeast nucleosomes were purified as described previously (47). At the last step, nuclei were resuspended in 400 µl NP buffer (20mM Tris-HCL pH 8.0, 150 mM NaCl, 5 mM KCl, 1 mM EDTA, 5 mM CaCl2, 10% glycerol, 1 mM PMSF, protease Inhibitors). Micrococcal nuclease was added to a final concentration of 10 u/mL, and incubated for 5 min at 37°C to obtain nucleosomes. The reaction was stopped with 10 µl of 0.5M EDTA. After spinning with max speed for 20 min at 4°C, the supernatant that contained nucleosomes was aliquoted and kept at -80°C. Yeast strains are listed in Table S1.

**In Vitro Binding Assays**

Histone peptide binding assays was performed as described (48). For mononucleosome binding assays, 1 µg of mononucleosome purified from Hela cells or *S. cerevisiae* and 5 µg of purified GST-LEDGF wild type or mutant proteins were incubated in 300 µl binding buffer (300 mM NaCl, 0.1% NP40 and 50 mM Tris pH 7.5, 10% glycerol) at 4°C for 4 hours, followed by an additional 1 hour with glutathione beads (Amersham). After extensive washing with binding buffer, beads were boiled with SDS buffer and subjected to western blot analysis using anti-H3 antibody (Abcam ab1791).

**Protein-protein ChIP**
Modified protein-protein ChIP assays using THP-1 human leukemia cells for detection of in situ LEDGF-histone interactions were performed as described (48).

**Microarray Data Processing and Gene Set Enrichment Analysis**

Raw microarray data of MLL-rearranged (20 cases) and MLL-germline ALL (112 cases) (49) was processed and normalized by Robust Multi-Array (RMA) method using R/Bioconductor package, affy. Normalized expression values were subjected to Gene Set Enrichment Analysis (GSEA) as described previously (50).

**Vector Construction**

pMSCV-neo constructs encoding MLL-AF9, MLL-AF10, MLL-ENL and E2A-HLF were described previously (14). Expression vectors for flag-KDM2A and flag-KDM2B were constructed by PCR using pCMV-SPORT6-KDM2A and pCMV-SPORT6-KDM2B (Open Biosystems) as templates and cloned into pMSCV-puro vector (Clontech).

**shRNA Lentivirus Production**

ASH1L (human) shRNA constructs were purchased from Open Biosystems (Table S5). For knockdown of LEDGF, ASH1L and SETD2 in mouse cells, shRNAs (Table S5) were cloned into the lentiviral pLKO.1 vector. Lentivirus was produced by co-transfection of viral vectors with pCMV–dR8.2 and pCMV-VSVG packaging constructs into 293T cells. Medium containing virus was collected 48 hrs post transfection and used for transductions.

**RNA Preparation and Real-time qPCR**
Total RNA was isolated and converted to cDNA as described previously (14), then subjected to RT-PCR on a CFX384 real-time PCR system (BIO-RAD) using TaqMan assays. Relative expression was normalized to beta-Actin or Gapdh. TaqMan probes used in real-time PCR are provided in Table S6.

Myeloid Transformation Assays

HSPC transformation assays were performed as previously described (14). For knock down experiments, MLL-fusion transformed cells were transduced with shRNA lentivirus by spinoculation, cultured in methylcellulose medium for 2 days, then selected in 2 μg/ml puromycin for 3 days, before harvesting for further analysis.

ACCESSION NUMBERS

The Gene Expression Omnibus accession number for the ChIP-seq data reported in this article is GSE73528.

SUPPLEMENTAL INFORMATION

Supplemental Information includes 5 figures and 6 tables.

ACKNOWLEDGEMENTS

We thank Cita Nicolas and Maria Ambus for technical support, Ziming Weng and Dr. Eliezer Calo for support on ChIP-sequencing, Dr. Xianxin Hua for MLL knockdown constructs, and Dr. ...
Matt Barber for suggestions on nucleosome binding assays. We thank members of the Cleary lab for discussions. We acknowledge the Stanford Hematology Division Tissue Bank and patients for providing clinical samples.
REFERENCES


FIGURE LEGENDS

Figure 1. Nucleosome-specific H3K36me2 recognition by LEDGF.

(A) The PWWP domain of LEDGF recognizes H3K36me2/3 peptide. Superimposed 1H,15N HSQC spectra of 0.1 mM PWWP collected in the absence (black) and presence (red) of the indicated histone peptides. The protein-peptide ratio is 1:10.

(B) The ability of WT LEDGF or PWWP domain mutants to bind nucleosomes purified from Hela cells was assessed by western blot analysis for histone H3 in a GST pull-down assay.

(C) Western blot analysis for histone H3 was performed on a GST-LEDGF pull down assay of nucleosomes isolated from various wildtype and mutant S. cerevisiae strains (indicated at top). Input nucleosomes are shown in lower panel. The detail information for yeast strains can be found in Table S1.

(D) THP-1 cells expressing Flag-tagged LEDGF (WT or W21A mutant) were subjected to protein-protein ChIP. Chromatin pulled-down with anti-Flag antibody was analyzed by western blot using antibodies indicated on the left.

Figure 2. LEDGF regulates MLL occupancy at target genes.

(A) A ChIP-seq density heat map representation of promoter region occupancies (3 kb flanking transcriptional start sites (TSS)) for the indicated proteins and histone marks. Gene promoters are rank-ordered by the level of MLL enrichment at the TSS.

(B) Average genome-wide occupancies are shown for LEDGF, MLL, H3K36me2 and H3K36me3 along the transcription unit. The gene body length is aligned by percentage from the TSS to TTS. Five kb upstream of TSS and 5 kb downstream of TTS are also included.
(C) Average genome-wide MLL occupancies are shown for genes enriched with LEDGF, or genes with no LEDGF enrichment.

(D) Genome browser representation of ChIP-seq peaks for the indicated proteins and histone marks displaying calculated ChIP-Seq enrichment values at the *MEIS1* and *CDK6* loci in MV4-11 leukemia cells. Relative positions of genes are shown at the top and CpG islands are shown at the bottom.

(E) Mouse MLL-AF9 leukemia cells were transduced with lentiviral vectors expressing control or LEDGF shRNAs. Relative transcript and protein levels were quantified by qRT-PCR and western blot analysis, respectively.

(F) MLL-AF9 transformed HSPCs transduced with either LEDGF or control shRNAs were subjected to ChIP assays using the indicated antibodies. Promoter regions amplified by qPCR are indicated below the respective panels.

(G) Mouse MLL-AF9 leukemia cells were transduced with lentiviral vectors expressing control or MLL shRNAs. Relative transcript levels were quantified by qRT-PCR.

(H) MLL-AF9 transformed HSPCs transduced with either MLL or control shRNAs were subjected to ChIP assays using the indicated antibodies. The regions amplified by qPCR are indicated below the respective panels.

Figure 3. *ASH1L*, an H3K36me2 methyltransferase, controls LEDGF and MLL occupancy at target genes.

(A) Genome browser representation of ChIP-seq peaks for the indicated proteins and histone marks displaying calculated ChIP-Seq enrichment values at the *RUNX2* locus in human MV4-11
leukemia cells. Relative positions of RUNX2 are shown at the top and CpG islands are shown at the bottom.

(B) Average genome-wide occupancies are shown for ASH1L, LEDGF, MLL, H3K36me2 and H3K36me3 along the transcription unit in MV4-11 cells.

(C) GSEA plot shows enrichment profile of the genes co-occupied by MLL, LEDGF and ASH1L in MV4-11 cells in genes differentially expressed in MLL-rearranged (20 cases) versus MLL-germline ALL (112 cases) (49). Heat map shows expression in ALLs for the top 25 differentially expressed genes identified from GSEA. Normalized enrichment score (NES) = 0.508; p <0.01.

(D) Mouse MLL-AF9 leukemia cells were transduced with lentiviral vectors expressing the indicated shRNAs. Transcript and protein levels were quantified by qRT-PCR and western blot analysis after three days of selection with puromycin.

(E-F) MLL-AF9 transformed HSPCs transduced with ASH1L or control shRNAs were subjected to ChIP assays using the indicated antibodies. Promoter regions amplified by qPCR are indicated below the respective panels.

Figure 4. ASH1L is required for leukemic transformation by MLL oncoproteins.

(A) Mouse MLL-AF9 leukemia cells were transduced with lentiviral vectors expressing the indicated shRNAs. Transcript levels were quantified by qRT-PCR after three days of selection with puromycin and expressed relative to control shRNA transduced cells.

(B) Light microscopy of May-Grunwald/Giemsa-stained mouse MLL-AF9 leukemia cells after five days of selection in methylcellulose culture with puromycin.
(C-D) The effects of ASH1L knockdown on differentiation (C) and apoptosis (D) in mouse MLL-AF9 leukemia cells (by flow cytometry) are shown.

(E) Mouse HSPCs transformed by the indicated oncogenes (below) were transduced with lentiviral vectors expressing control or ASH1L shRNAs. Colony numbers were quantified after 5 days of selection in methylcellulose culture with puromycin. Results are expressed relative to control shRNA transduced cells.

(F) Survival curves are shown for cohorts of mice transplanted with mouse MLL-AF10 leukemia cells (1x10⁶) transduced with control or ASH1L shRNAs (n = 5 each cohort). Acute leukemia was confirmed by peripheral blood leukocyte count and necropsy. Log-rank test was used for statistical analysis (p = 0.0017).

Figure 5. ASH1L is required for transformation of MLL-rearranged human leukemia cells.

(A) MV4-11 leukemia cells were transduced with lentiviral vectors expressing control or ASH1L shRNAs. Transcript levels for the indicated genes were quantified by qRT-PCR after three days of selection with puromycin and expressed relative to control shRNA transduced cells.

(B, C) Human leukemia cell lines were transduced with lentiviral vectors expressing the indicated shRNAs. Colony numbers were quantified after seven days of selection in methylcellulose culture with puromycin, and expressed relative to control shRNA transduced cells (B). Cell numbers were quantified after indicated times of selection in liquid culture with puromycin (C).

(D, E) MV 4-11 leukemia cells (D) or MLLr primary human AML cells (E) were transduced with ASH1L or control shRNAs and transplanted into NSG mice. The percentage of transduced (RFP⁺) cells was assessed before transplant and in the bone marrow after 4 weeks (D) and 10
weeks (E) in vivo. Normalized RFP reflects the ratio of RFP+ human cells at a given time point to the level immediately prior to transplant.

**Figure 6. H3K36me3 is not essential for LEDGF and MLL target gene occupancy.**

(A) Mouse MLL-AF9 leukemia cells were transduced with lentiviral vectors expressing the indicated shRNAs. Transcript levels for genes indicated below were quantified by qRT-PCR after three days of selection with puromycin, and expressed relative to control shRNA transduced cells.

(B) Global levels of histone methylation were determined by western blot analysis of MLL-AF9 leukemia cells transduced with SETD2 or control shRNAs.

(C) Mouse HSPCs transformed by MLL-AF9 were transduced with lentiviral vectors expressing the indicated shRNAs. Colony numbers were quantified after five days of selection in methylcellulose culture with puromycin. Results are expressed relative to control shRNA transduced cells.

(D) MLL-AF9 transformed HSPCs transduced with SETD2 or control shRNAs were subjected to ChIP assays using the indicated antibodies. Promoter regions amplified by qPCR are indicated below the respective panels.

**Figure 7. KDM2A impairs leukemic transformation by MLL oncoproteins.**

(A) Mouse MLL-AF10 leukemia cells were transduced with KDM2A or KDM2B over-expression or control vectors. Transcript levels for genes indicated below the respective panels were quantified by qRT-PCR after three days of selection with puromycin and expressed relative to control vector transduced cells.
(B) MLL-AF10 transformed HSPCs transduced with KDM2A over-expression or control vectors were subjected to ChIP assays using the indicated antibodies. Promoter regions amplified by qPCR are indicated below the respective panels.

(C) MLL-AF10 transformed HSPCs were transduced with KDM2A or KDM2B over-expression or control vectors. Colony numbers were quantified after five days of selection in methylcellulose culture with puromycin. Results are expressed relative to control vector transduced cells.

(D) Mouse HSPCs transformed by the indicated oncogenes were transduced with KDM2A over-expression or control vectors. Colony numbers were quantified after five days of selection in methylcellulose culture with puromycin. Results are expressed relative to control vector transduced cells.

(E) Survival curves are shown for cohorts of mice transplanted with MLL-AF10 leukemia cells (1x10^6) transduced with KDM2A over-expression or control vectors. (n = 5 each cohort). Acute leukemia was confirmed by peripheral blood leukocyte count and necropsy. Log-rank test was used for statistical analysis (p = 0.0018).
ASH1L links histone H3 lysine 36 di-methylation to MLL leukemia

Li Zhu, Qin Li, Stephen H. K. Wong, et al.

Cancer Discov  Published OnlineFirst May 6, 2016.