The Malignant Brain Tumor (MBT) Domain Protein SFMBT1 Is an Integral Histone Reader Subunit of the LSD1 Demethylase Complex for Chromatin Association and Epithelial-to-mesenchymal Transition*

Received for publication, May 1, 2013, and in revised form, August 7, 2013 Published, JBC Papers in Press, August 8, 2013, DOI 10.1074/jbc.M113.482349

Ming Tang†,‡, Huaxuan Shen†,††, Yue Jin‡, Tong Lin§,§, Qingsong Cai†, Melissa A. Pinard§,§, Shyamasri Biswas‡,‡, Quyen Tran§§, Guangyao Li§§, Anitha K. Shenoy‡,§, Emily Tongdee‡,§, Shuabin Lin‡, Yumei Gu†, Brian K. Law†,§, Lei Zhou†, Robert Mckenna‡,‡, Lizi Wu‡,§, and Jianrong Lu†,‡‡

From the †Department of Biochemistry and Molecular Biology, ‡Department of Molecular Genetics and Microbiology, and §Department of Pharmacology and Therapeutics, Health Cancer Center, University of Florida College of Medicine, Gainesville, Florida 32610, ‡‡Genetics and Genomics Graduate Program, Genetics Institute, University of Florida, Gainesville, Florida 32610, ††State Key Laboratory of Ophthalmology, Zhongshan Ophthalmic Center, Sun Yat-sen University, Guangzhou, 510060 China, and **Agnes Scott College, Decatur, Georgia 30030

Background: The LSD1 histone demethylase regulates gene expression by demethylating chromatin.
Results: SFMBT1 associates with the LSD1 complex and is essential for its chromatin association, histone demethylation, and induction of EMT.
Conclusion: SFMBT1 is a novel, integral component of the LSD1 complex.
Significance: Understanding how the LSD1 complex is recruited to chromatin may offer new opportunities for therapeutic intervention.

Chromatin readers decipher the functional readouts of histone modifications by recruiting specific effector complexes for subsequent epigenetic reprogramming. The LSD1 (also known as KDM1A) histone demethylase complex modifies chromatin and represses transcription in part by catalyzing demethylation of dimethylated histone H3 lysine 4 (H3K4me2), a mark for active transcription. However, none of its currently known subunits recognizes methylated histones. The Sna1 family transcription factors are central drivers of epithelial-to-mesenchymal transition (EMT) by which epithelial cells acquire enhanced invasiveness. Sna1-mediated transcriptional repression of epithelial genes depends on its recruitment of the LSD1 complex and ensuing demethylation of H3K4me2 at its target genes. Through biochemical purification, we identified the MBT domain-containing protein SFMBT1 as a novel component of the LSD1 complex associated with Sna1. Unlike other mammalian MBT domain proteins characterized to date that selectively recognize mono- and dimethylated lysines, SFMBT1 binds di- and trimethyl histone H3 lysine 4 (H3K4me2), both of which are enriched at active promoters. We show that SFMBT1 is essential for Sna1-dependent recruitment of LSD1 to chromatin, demethylation of H3K4me2, transcriptional repression of epithelial markers, and induction of EMT by TGFβ. Carcinogenic metal nickel is a widespread environmental and occupational pollutant. Nickel alters gene expression and induces EMT. We demonstrate the nickel-initiated effects are dependent on LSD1-SFMBT1-mediated chromatin modification. Furthermore, in human cancer, expression of SFMBT1 is associated with mesenchymal markers and unfavorable prognosis. These results highlight a critical role of SFMBT1 in epigenetic regulation, EMT, and cancer.

Covalent posttranslational modifications of histones are crucial determinants of chromatin structure and gene transcription (1–4). Histone posttranslational modifications may serve as regulatory marks to be recognized by specific “reader” proteins (5, 6), which are typically part of large multisubunit chromatin-modifying or -remodeling complexes. Recruitment of such complexes can in turn further modify the epigenetic landscape and impact gene expression. Histone lysine methylation is associated with gene activation or repression, depending on the site and degree of the modification. In particular, methylation of lysine 4 of histone H3 (H3K4)§ correlates with gene activation. Monomethylation of H3K4 (H3K4me1) is a marker for enhancers, whereas di- and trimethylation of H3K4 (H3K4me2/3) are highly enriched at active promoters. Dynamic H3K4 methylation is critical for transcriptional regulation.

Lysine-specific demethylase 1 (LSD1, also known as KDM1A, AOF2), the first identified histone demethylase, is a flavin adenine dinucleotide (FAD)-dependent amine oxidase that specifically removes methyl groups from H3K4me1/2,

* This work was supported, in whole or in part, by National Institutes of Health Grant R01CA137021 (NCI; to J. L.). This work was also supported by the Florida Bankhead-Coley Cancer Research Program (08BN-12-23092 and 08BT01).
† Both authors contributed equally to this work.
‡ To whom correspondence may be addressed. E-mail: lzwu@ufl.edu.
§ To whom correspondence may be addressed. E-mail: jrlu@ufl.edu.
¶ The abbreviations used are: H3K4, lysine 4 of histone H3; LSD1, lysine-specific demethylase 1; EMT, epithelial-to-mesenchymal transition; SFMBT1, Scm-like with four MBT domains 1; MBT, malignant brain tumor; Bis-tris, 2-[bis(2-hydroxyethyl)amino]-2-(hydroxyethyl)propane-1,3-diol; KRT8, keratin 8; 4HT, 4-hydroxy-tamoxifen; ER, endoplasmic reticulum; IP, immunoprecipitation.
SFMBT1 Is Essential for LSD1 Binding to Chromatin and EMT

thereby repressing transcription (7). As precise gene regulation is instrumental in signal transduction, cell fate determination and differentiation, LSD1 has emerged as a pivotal player in various normal developmental processes, and its deregulation has been involved in diseases such as cancer. The importance of LSD1-dependent transcriptional regulation is reflected by its association with a wide variety of transcription factors. For instance, LSD1 occupies and controls H3K4 methylation at the promoters of a subset of OCT4- and NANOG-repressed developmental genes in human embryonic stem (ES) cells to maintain stem cell pluripotency (8). During ES cell differentiation, LSD1-mediated demethylation of H3K4me1 decommissions enhancers of active genes and facilitates stem cell differentiation (9). LSD1 interacts with orphan nuclear receptor TLX to regulate neural stem cell proliferation (10). In pituitary organogenesis, LSD1 is recruited by the transcription factor ZEB1 to repress its targets (11). During hematopoiesis, LSD1 associates with key hematopoietic regulators GFI1/B, TAL1, BLIMP1, BCL11A, and nuclear receptors TR2 and TR4 for differentiation of multiple hematopoietic lineages (12–16). In addition, LSD1 binds to the p53 tumor suppressor protein (17, 18). LSD1 also forms complexes with RBPJ to mediate Notch signaling in development and cancer (19, 20).

We and others recently showed that LSD1 functions as a corepressor physically associating with the zinc finger transcription factors Snai1 and Snai2 (21–23), which are crucial inducers of epithelial-to-mesenchymal transition (EMT) (24). EMT is a reprogramming event by which epithelial cells decrease expression of epithelial genes, lose cell adhesion, and acquire enhanced motility and invasiveness (25). It is a vital developmental program but is often reactivated in cancer, adversely contributing to malignant progression (25). Our previous study showed that LSD1 is essential for Snai1-mediated repression of epithelial markers (22). Consistently, LSD1 is overexpressed in advanced cancers, and its elevated expression correlates with tumor progression and unfavorable clinical outcome (26–30).

Although the critical role of LSD1 in a broad spectrum of physiological processes and its involvement in pathogenesis has been well documented, much remains to be understood about its functional regulation. LSD1 was mostly found in a core repressor complex with RCOR1 (CoREST) as well as histone deacetylase 1 and 2 (HDAC1/2) (7, 31–33). Recruitment of chromatin-modifying complexes to particular genomic sites generally requires multivalent interactions, involving combinatorial recognition of DNA sequence by associated sequence-specific transcription factors and local chromatin environment by the reader subunits. Although the LSD1 complex interacts with various transcription factors to regulate distinct target genes, it remains largely elusive how this complex recognizes epigenetic settings, in particular histone posttranslational modifications that may act as docking signals for its recruitment. Except for the catalytic domain of LSD1, no subunits in the complex have been known to recognize modified histone tails.

In the present study we purified the Snai1-LSD1 protein complex and identified a novel associated factor, SFMBT1 (Scm-like with four MBT domains 1) (34). SFMBT1 belongs to the malignant brain tumor (MBT) domain-containing protein family that is critical in chromatin regulation (35). Most MBT proteins characterized so far selectively bind mono- and dimethylated lysine residues on the amino-terminal histone tails. Our subsequent biochemical analysis showed that SFMBT1 associated with the LSD1 complex through physical interaction with the core subunit RCOR1. Unlike other MBT members that bind lower methylated lysines, SFMBT1 recognizes di- and tri-methylated H3K4 in vitro. Moreover, SFMBT1 is indispensable for Snai1-induced recruitment of LSD1 to chromatin, demethylation of H3K4me2, and transcriptional repression. Consistent with these findings, SFMBT1 is essential for induction of EMT by stimuli such as TGFβ and heavy metal nickel.

**EXPERIMENTAL PROCEDURES**

**Affinity Purification of Protein Complexes Associated with Snai1**—HEK293 cells were transfected with pcDNA3-Snai1-FLAG and selected by G418. Clones were picked and examined for Snai1 expression by immunoblotting with anti-FLAG antibodies (Sigma). Cells stably expressing Snai1-FLAG were subjected to affinity purification as previously described (36).Briefly, nuclear extracts (300 mg) were prepared and immunoprecipitated with anti-FLAG antibody-conjugated agarose (Sigma). After extensive washing, protein complexes were eluted using a 50 μg/ml FLAG peptide (Sigma) at room temperature for 30 min. The eluted protein samples were separated on the 4–12% NuPAGE Bis-tris gels (Invitrogen) and analyzed by silver staining and LC-MS/MS mass spectrometry at Taplin Biological Mass Spectrometry Facility of Harvard Medical School.

**RT-Quantitative PCR**—Cells were harvested in TRIzol (Invitrogen) followed by total RNA extraction. Reverse transcription of RNA was conducted using Moloney murine leukemia virus (MuLV) reverse transcriptase with random primers. The expression levels of selected genes were determined by real-time PCR with the SYBR Green PCR kit (Applied Biosystems). Data were normalized against β-actin.

**ChIP Analysis**—ChIP assay was conducted as previously reported (22). Briefly, cells were cross-linked with 1% formaldehyde for 10 min. The reaction was stopped by the addition of glycine. Cross-linked cells were washed in 1× PBS and collected. Cell pellets were washed with washing buffer (0.25% Triton X-100, 10 mM EDTA, 0.5 mM EGTA, 10 mM Tris, pH 8.0) resuspended in sonication buffer (1 mM EDTA, 0.5 mM EGTA, and 10 mM Tris pH 8.0), mixed with glass beads, and subjected to sonication. The sonicated samples were diluted in ChIP buffer (0.01% SDS 1.0% Triton X-100, 1.0 mM EDTA, 20 mM Tris, pH 8.1, 150 mM NaCl) and incubated with antibody against Snai1 (Santa Cruz), Snai2 (Cell Signaling), SFMBT1 (Bethyl), LSD1 (Upstate Biotechnology), H3K4me2 (Millipore), and H3K4me3 (Cell Signaling), and protein A slurry (Invitrogen). The immunoprecipitates were subjected to a series of washing steps to remove nonspecific binding material. After reversal of cross-linking, DNA samples were purified, and the enrichment of specific genomic regions was determined by real-time quantitative PCR. Final results represent the percentage of input chromatin.

**Histone Peptide Pulldown Assay**—Four micrograms of biotinylated histone H3 or H4 peptides (ANASPEC) were incubated...
with cell lysate from HEK293 cells transected with FLAG-SFMBT1 in 400 µl of binding buffer (50 mM Tris, pH 7.5, 100 mM NaCl, 0.1% Nonidet P-40) at 4 °C for overnight. The next day 12 µl of streptavidin beads (Promega) were added into the binding buffer and incubated at 4 °C for 1 h with rotation. After three washes with the binding buffer, the beads were boiled in protein loading buffer and subjected to SDS-PAGE and Western blotting with anti-FLAG antibodies.

**Generation of Snai1-ER Cells and Lentiviral shRNA Knockdown**—Cells were infected with lentivirus expressing both GFP and the Snai1-ER fusion protein. GFP-positive cells were sorted out by flow cytometry. These cells were subsequently transduced with lentiviral vector pLKO or shRNA targeting SFMBT1 (OpenBiosystems). After selection with puromycin (2 µg/ml) for 1 week, these cells were subjected to 4-hydroxytamoxifen (4HT) (100 nM) treatment for 2 days and followed by RT-quantitative PCR and ChIP assays.

**Treatments with TGFβ and Nickel and Transwell Invasion Assay**—A549 cells were transduced with lentiviral vector pLKO or shRNA targeting SFMBT1. After selection with puromycin (1 µg/ml) for 1 week, cells were cultured in 6-well plates with or without TGFβ (5 ng/ml) or nickel chloride (1 mM) for 2 days. Morphological changes were observed, and phase contrast images were taken under a microscope. These cells were then subjected to RT-quantitative PCR and Transwell assays. The membranes of Transwell chambers were coated with 20 µg/ml fibronectin at 37 °C overnight. Cells (4 × 10^4) were resuspended with serum-free medium and added into the top of the insert. Inserts were placed in full medium with serum. After 24 h, cells that migrated to the bottom of the membrane were fixed and stained with 0.1% crystal violet. Pictures were taken under a microscope, and the numbers of cells were counted.

**Gene Expression and Disease Association Analysis**—To investigate whether SFMBT1 may play a role in breast cancer patient survival, we first chose the GSE25055 dataset (from the NCBI GEO website) as the training dataset and performed a univariate Cox proportional hazards regression analysis to find genes whose expression was potentially associated with patient survival. This gene list was then overlapped with the EMT gene list (Sabioscience) and further limited by clustering analysis to six genes that are coexpressed with SFMBT1 in breast cancers (CDH2, VIM, and CALD1, which were up-regulated after EMT, and CLDN5, KRT19, and CDH1, which were down-regulated after EMT). We next constructed a multi-gene score using the following algorithm. Normalized probeset log2 values were subtracted with median values from all samples for one particular probeset. The multi-gene score is the averaged sum of the three EMT-up-regulated genes and SFMBT1 minus that of SFMBT1 and three EMT-down-regulated genes. Based on the gene expression heatmap, a cutoff was set to stratify patients into three equal number groups having low, medium, and high seven-gene scores. This effectively separated patients into poor and good prognosis groups.

**Statistical Analyses of Experimental Data**—Independent Student’s t test was used to analyze data from various experiments.

**RESULTS**

The MBT Domain-containing Protein SFMBT1 Is Essential for LSD1 Binding to Chromatin and EMT

**SFMBT1 Associates with Snai1 through Binding to the LSD1 Complex**—The LSD1 complex is known to directly bind to Snai1. The newly discovered SFMBT1-Snai1 interaction raised...
the possibility that SFMBT1 may interact with the LSD1 complex. When HEK293 cells were transfected with FLAG-SFMBT1, immunoprecipitation of exogenous SFMBT1 with anti-FLAG-agarose beads was sufficient to enrich endogenous LSD1 (Fig. 2A). Furthermore, in untransfected cells, immunoprecipitation of endogenous SFMBT1 with SFMBT1-specific antibodies also pulled down endogenous LSD1 (Fig. 2B). These results demonstrate the association between SFMBT1 and LSD1 and the existence of an endogenous SFMBT1-LSD1 complex. Consistently, in our large scale purification of SFMBT1-associated proteins, the LSD1 complex was readily identified (40).

Because of the apparently mutual associations between Snai1, SFMBT1, and the LSD1 complex, we suspected that SFMBT1 might associate with Snai1 through the LSD1 complex. To test this idea, we first depleted LSD1 in HEK293 cells with two independent lentiviral short hairpin RNAs (shRNAs) (22) (Fig. 2C). Control and LSD1-depleted cells were then transiently transfected with Snai1 and FLAG-SFMBT1 followed by co-IP assays. In control cells, immunoprecipitation of SFMBT1 with anti-FLAG antibodies efficiently pulled down both LSD1 and Snai1 (Fig. 2C, second lane from the left); however, in cells depleted of LSD1, the association between SFMBT1 and Snai1 was substantially reduced (Fig. 2C, third and fourth lanes from the left), suggesting that SFMBT1 interacts with Snai1 through the LSD1 complex.

The LSD1 complex consists of many subunits, and it was unclear to which subunit SFMBT1 might bind. We decided to use the GST pulldown assay to address this question. SFMBT1 contains four MBT repeats at the amino terminus, an SPM motif at the carboxyl terminus and a highly conserved domain (DUF3588) with unknown function in the middle (Fig. 3A) (35). Each of these domains was fused to GST, and the resulting fusion proteins were purified from bacteria (Fig. 3A). When these GST fusion proteins were incubated with extracts prepared from HEK293 cells transfected with LSD1, none of them

---

**FIGURE 1.** The MBT domain-containing protein SFMBT1 is associated with Snai1. A, identification of Snai1-associated proteins. Left, silver staining of Snai1-associated proteins. Nuclear extracts prepared from cells stably expressing Snai1-FLAG or empty vector (mock) were immunoprecipitated with anti-FLAG-agarose beads. The protein complex was eluted with FLAG peptides after extensive wash and separated by 4–12% gradient SDS-PAGE. Right, a list of potential Snai1-associated proteins was identified by mass spectrometric analysis. B, a phylogenetic tree of fly and human MBT domain-containing proteins derived from multiple sequence alignment. d(3)MBT, dSCM, and dSFMBT are three MBT proteins in Drosophila. WB, Western blot. C, SFMBT1 interacts with Snai1. HEK293 cells were transiently transfected with Snai1 together with or without FLAG-SFMBT1, lysed, and followed by co-IP with anti-FLAG antibodies and Western blotting with anti-Snai1 and Anti-FLAG antibodies. Snai1 was specifically co-precipitated with SFMBT1. D, SFMBT1 interacts with Snai2. Snai2 and FLAG-SFMBT1 were co-expressed in HEK293 cells and subjected to co-IP analysis similar to B.
exhibited detectable binding of LSD1 (Fig. 3A), suggesting that SFMBT1 does not directly interact with LSD1. We thus further examined other subunits of the LSD1 complex. When incubated with lysate from cells transfected with RCOR1, a dimerization partner of LSD1 and a core subunit of the complex (31–33), the middle part of SFMBT1 showed strong association with RCOR1 (Fig. 3A), suggesting that SFMBT1 associates with the LSD1 complex through binding to RCOR1.

The interaction between SFMBT1 and RCOR1 was further validated by a reciprocal pulldown assay. RCOR1 contains several functional domains (32) (Fig. 3B). In this experiment, different domains of RCOR1 were fused to GST (Fig. 3B) and incubated with cellular extract containing FLAG-SFMBT1. The linker region of RCOR1, which forms an elongated structure with the “tower domain” of LSD1 (33), interacted with SFMBT1 (Fig. 3B). As a control, L3MBTL1, another MBT domain protein that does not have the unique middle domain of SFMBT1, did not bind to RCOR1 (Fig. 3B). Collectively, these results suggest that SFMBT1 indirectly associates with Snai1 through binding to the LSD1 complex core subunit RCOR1, resulting in assembly of a larger complex minimally comprising Snai1-LSD1-RCOR1-SFMBT1.

FIGURE 3. SFMBT1 interacts with the LSD1 complex subunit RCOR1. A, SFMBT1 interacts with RCOR1. Top, structural diagram of SFMBT1. Indicated regions of SFMBT1 (N, amino acids (aa) 1–479; M, aa 494–699; C, aa 721–866) were fused to GST. The fusion proteins were purified from bacteria and incubated with extracts prepared from HEK293 cells transfected with RCOR1 or LSD1. After extensive wash, the protein complex was subjected to Western blotting (WB) with indicated antibodies. B, the linker region of RCOR1 is responsible for SFMBT1 binding. Top, domains of RCOR1. Different fragments of RCOR1 containing various domains (E1–S1, amino acids (aa) 1–293; S1, aa 181–293; L1–S2, aa 283–482; S2, aa 381–482) were fused to GST and subjected to pulldown assays with cell lysates expressing FLAG-SFMBT1 or FLAG-L3MBTL1 followed by Western blotting with anti-FLAG antibodies.

SFMBT1 Is Recruited to Snai1 Target Genes—The LSD1 complex is recruited to epithelial gene promoters by Snai1 (22). Due to its association with the LSD1 complex, SFMBT1 was expected to be recruited to these genes by Snai1 as well. To verify this idea, we employed an inducible form of Snai1, namely Snai1-ER, in which Snai1 is fused to a mutant form of the estrogen receptor ERα ligand binding domain that recognizes the synthetic ligand 4HT (41). The Snai1-ER fusion protein is retained in the cytoplasm and thus unable to repress transcription in the nucleus. Upon the addition of 4HT, the fusion protein translocates into the nucleus and binds and represses Snai1 target genes. This inducible model has been validated in various cultured cell lines and in vivo (41). When MCF7 human breast cancer epithelial cells were transduced with lentivirus expressing Snai1-ER, ChIP analysis with anti-Snai1 antibodies revealed that the binding of the Snai1 fusion protein to the E-cadherin promoter was minimal in cells treated with vehicle (DMSO) but was strongly induced after cells were treated with 4HT (Fig. 4A).

We next investigated whether SFMBT1 could be recruited to epithelial genes in a Snai1-dependent manner. MCF7 cells expressing Snai1-ER were treated with DMSO (vehicle) or 4HT followed by ChIP assay with IgG or anti-Snai1 antibody. Binding of Snai1-ER to the E-cadherin promoter (Ecad-pro) and a 5-kb upstream region was examined. Binding of Snai1-ER to the E-cadherin promoter (Ecad-pro) and a 5-kb upstream region was examined. Binding of SFMBT1 to epithelial gene promoters is induced by Snai1. MF7 Snai1-ER cells were treated with DMSO or 4HT followed by ChIP assay with IgG (control) or antibodies for SFMBT1. Binding at promoters of E-cadherin and KRT8 was determined. Experiments were conducted in triplicate, and error bars indicate S.D.

FIGURE 4. SFMBT1 is recruited to Snai1 target genes. A, validation of Snai1-ER, an inducible form of Snai1. MCF7 cells expressing Snai1-ER were treated with DMSO (vehicle) or 4HT followed by ChIP assay with IgG or anti-Snai1 antibody. Binding of Snai1-ER to the E-cadherin promoter (Ecad-pro) and a 5-kb upstream region was examined. B, binding of SFMBT1 to epithelial gene promoters is induced by Snai1. MCF7 Snai1-ER cells were treated with DMSO or 4HT followed by ChIP assay with IgG (control) or antibodies for SFMBT1. Binding at promoters of E-cadherin and KRT8 was determined. Experiments were conducted in triplicate, and error bars indicate S.D.
SFMBT1 Recognizes Tri- and Dimethylated H3K4—Recruitment of SFMBT1 to the Snai1 target genes prompted us to investigate its role in Snai1-mediated epigenetic regulation. MBT domains characterized to date are readers of post-translational modifications of histones, more specifically, mono- and dimethylated lysine residues (35). SFMBT1 contains four MBT repeats that may potentially bind methyl lysines. Interestingly, none of the previously known subunits of the LSD1 complex recognize methylated histone tails. Therefore, we postulate that SFMBT1 may be a candidate histone methyl-lysine reader of the LSD1 complex.

Before Snai1-initiated repression in epithelial cells, Snai1 target genes (including many epithelial genes) are actively expressed, possessing histone marks associated with active transcription. It is hence conceivable that the putative histone reader may recognize certain active marks and facilitate the recruitment/activation of the LSD1 complex. We thus focused on active lysine methylation marks, most notably the H3K4 methylations at various degrees. In addition, monomethylation of H3K9, H3K27, and H4K20 is linked to gene activation (42). Therefore, we evaluated the binding capacity of SFMBT1 with candidate histone species that harbor various forms of lysine methylation in vitro. Biotinylated H3 and H4 peptides were incubated with cellular extract expressing FLAG-SFMBT1. Histone peptides and possible bound SFMBT1 were purified with streptavidin beads. Among the examined peptides, both tri- and dimethylated H3K4 exhibited detectable binding to SFMBT1 (Fig. 5A). H3K4me3 appeared to have a higher affinity toward SFMBT1 than H3K4me2. Unmodified or mono-methylated H3 or H4 tails did not show evidence of association with SFMBT1 (Fig. 5A). This result suggests that SFMBT1 may recognize specific histone methyl marks, in particular H3K4me2/3, and its binding specificity may differ from that of previously characterized other human MBT proteins that demonstrate selective binding only to mono- and dimethylated histone peptides (35).

SFMBT1 Is Required for Snai1-induced Recruitment of LSD1, Demethylation of Dimethyl-H3K4, and Transcriptional Repression—Based on the SFMBT1 ability to bind histone marks and its potential as a histone reader, we next tested whether SFMBT1 contributed to Snai1-dependent chromatin recruitment of LSD1. We used shRNAs (in the pLKO lentiviral vector) to efficiently deplete SFMBT1 in Snai1-ER MCF7 cells (Fig. 5B). Consistent with our previous finding that Snai1 recruited LSD1 to specific genomic sites (22), 4HT treatment stimulated the association of LSD1 to the promoters of epithelial genes E-cadherin and KRT8 in control cells transduced with the empty pLKO vector, but this effect was virtually abolished in SFMBT1-depleted cells (Fig. 5C), demonstrating that SFMBT1 is essential for chromatin association of LSD1 induced by Snai1.

LSD1 specifically demethylates mono- and dimethylated H3K4. We further determined the role of SFMBT1 in the demethylation process using the inducible Snai1-ER cells. Similar to our previous observation (22), induction of Snai1 by 4HT potently reduced the level of H3K4me2 at the epithelial gene promoters in control cells (Fig. 5D). In cells depleted of SFMBT1, however, demethylation of H3K4me2 was greatly impaired (Fig. 5D). LSD1 cannot demethylate trimethylated H3K4. Indeed, this mark was not responsive to 4HT treatment regardless of the SFMBT1 depletion status (Fig. 6A). These results suggest that SFMBT1 is indispensable for Snai1-dependent recruitment of LSD1 to chromatin and subsequent H3K4me2 demethylation.

Removal of H3K4me2 by LSD1 is required for Snai1-mediated transcriptional repression (22); we hence expected that SFMBT1 was also essential for Snai1 repressive activity. In control pLKO-transduced MCF7 Snai1-ER cells, treatment with 4HT robustly inhibited endogenous E-cadherin RNA expression, but this inhibition was significantly compromised when SFMBT1 was depleted (Fig. 6B).

We noticed that depletion of SFMBT1 did not completely abolish Snai1-mediated repression of E-cadherin and wondered whether this might be related to the expression levels of Snai1. We transduced another transformed mammary epithelial cell line, MCF10DCIS, with lentiviral Snai1-ER-IRES-GFP. GFP-positive cells were sorted and, based on signal intensity, grouped as GFP-low (i.e. expressing lower levels of Snai1-ER) and GFP-high (i.e. expressing higher levels of Snai1-ER). Each group of cells was further infected with lentiviral vector pLKO or shRNA targeting SFMBT1. In both groups treatment of cells with 4HT decreased RNA levels of endogenous E-cadherin, and the repression was more robust in cells expressing higher levels of GFP (i.e. Snai1-ER) (Fig. 6C). As expected, depletion of
SFMBT1 Is Essential for LSD1 Binding to Chromatin and EMT

SFMBT1 weakened the 4HT-induced repression on E-cadherin expression in both GFP-low and -high groups (Fig. 6C). But in GFP-high cells, de-repression of E-cadherin by SFMBT1 depletion is less robust than that in GFP-low cells (Fig. 6C). These observations suggest that SFMBT1 is required for Snai1-mediated transcriptional repression and may also imply that high levels of Snai1 may possess SFMBT1-independent repressive activity.

Based on these findings, we propose the following working model (Fig. 6D): SFMBT1 generally exists as a part of the LSD1 corepressor complex (at least consisting of LSD1-RCOR1-SFMBT1). This complex is present in epithelial cells but is not particularly enriched at epithelial genes. When Snai1-like factors are induced, Snai1 recruits this complex to its genomic targets through physical interaction with LSD1. Binding of SFMBT1 to histone marks facilitates this recruitment and/or further stabilizes the association of the complex with chromatin. In the absence of SFMBT1, Snai1 alone appears to be insufficient to tether the complex to chromatin.

SFMBT1 Is Essential for Induction of EMT by TGFβ—The aforementioned experimental settings mostly involve ectopic expression of Snai1. We decided to investigate the role of SFMBT1 in more physiological EMT models. TGFβ is a potent driver of EMT, tumor progression, and metastasis (43). Treatment with TGFβ is often sufficient to induce EMT in various epithelial cells. Depending on the context, TGFβ up-regulates endogenous Snai1 and/or Snai2, which are essential for TGFβ-induced EMT phenotype (44–46). Snai1 and Snai2 share the nearly identical SNAG repressive domain, which is sufficient for binding LSD1 (21, 22). The Snai2-LSD1 interaction was confirmed experimentally (23).

TGFβ treatment of A549 human lung cancer epithelial cells activated expression of Snai2 (Fig. 7A). To define the role of SFMBT1 in TGFβ-induced EMT, we depleted SFMBT1 in A549 cells (Fig. 7B). Upon TGFβ treatment, control A549 cells displayed decreased expression of epithelial genes, such as E-cadherin and KRT8 (Fig. 7, C and D) and increased expression of the mesenchymal marker N-cadherin (Fig. 7D). By contrast, in SFMBT1-depleted cells down-regulation of epithelial markers by TGFβ was significantly impaired, whereas induction of N-cadherin was not affected (Fig. 7, C and D).

During EMT, epithelial cells lose cell-cell adhesion and acquire a mesenchymal, fibroblast-like morphology. A549 cells display a typical epithelial phenotype; cells are cohesive and clustered with extensive cell-cell contact (Fig. 7E). After treatment with TGFβ, these cells became elongated and dissociated from each other (Fig. 7E), which is characteristic of EMT and consistent with the changes in gene expression. In contrast, SFMBT1-depleted cells largely resisted the “scattering” effect of TGFβ (Fig. 7E).

By promoting morphological changes, EMT confers increased invasiveness on cells. To evaluate whether SFMBT1 regulates cell invasion, we performed Transwell invasion assays. Transwell cell-permeable membranes were coated with fibronectin to mimic the extracellular matrix that cells encounter during migration and invasion. A549 cells had limited capacity to transverse the membrane. However, after TGFβ treatment, control cells exhibited a dramatically enhanced invasive property as more cells seen invaded through the membrane (Fig. 7F). Depletion of SFMBT1 essentially abolished TGFβ-induced invasion (Fig. 7F). Taken together, these results suggest that SFMBT1 is required for TGFβ-induced EMT and cell invasion.

FIGURE 6. SFMBT1 is essential for Snai1-mediated transcriptional repression. A, H3K4me3 levels at the E-cadherin promoter are not affected by Snai1. Control (pLKO) or SFMBT1-depleted MCF7 Snai1-ER cells were treated with DMSO or 4HT for 2 days followed by ChIP analysis with antibodies recognizing H3K4me3. Ecad-pro, cadherin promoter. B, SFMBT1 is required for repression of endogenous E-cadherin by Snai1. MCF7 cells stably expressing Snai1-ER were infected with lentiviral vector pLKO or shRNA targeting SFMBT1 followed by treatment with DMSO or 4HT. Cells were subjected to RT-PCR to determine the RNA levels of E-cadherin. C, SFMBT1 is required for Snai1-dependent repression of E-cadherin in MCF10DCIS cells. MCF10DCIS cells were transduced with lentiviral Snai1-ER-IRES-GFP, and GFP + cells were divided into high and low expression groups. These cells were further infected with lentiviral vector pLKO or shRNA targeting SFMBT1 followed by 4HT treatment. E-cadherin expression was determined by RT-PCR. Results were derived from triplicate experiments, and error bars represent S.D. D, a working model for SFMBT1 in recruiting the LSD1 complex. Left, in epithelial cells without Snai1/2, the LSD1-RCOR1-SFMBT1 complex is present, but it does not stably associate with epithelial gene promoters. Middle, when Snai1/2 are induced, multivalent binding events tether the LSD1 complex to chromatin at epithelial genes. Right, in the absence of SFMBT1, LSD1 fails to steadily bind to chromatin (the Snai1-LSD1 interaction alone appears to be insufficient to anchor the complex on chromatin).
**SFMT1 Is Essential for Nickel-induced EMT**—Many environmental pollutants (including some heavy metals) are carcinogenic. Exposure to nickel and its compounds is fairly extensive in both environmental and occupational settings. The metal carcinogenicity is in large part attributed to their epigenetic effect (47), although the underlying mechanism remains poorly understood. Carcinogenic nickel compounds induce EMT and promote malignant progression during carcinogenesis (48). Exposure of A549 cells to nickel chloride (NiCl2) strongly up-regulated Snai2 and concomitantly down-regulated epithelial markers including E-cadherin (Fig. 8A).

Induction of Snai2 prompted us to determine whether nickel-induced EMT might involve LSD1-SFMT1-dependent epigenetic regulation. ChIP analysis showed that after nickel treatment, the binding of SFMT1 and LSD1 to epithelial gene promoters was potently increased (Figs. 8B and 9A, respectively), whereas H3K4me2 levels at these promoters significantly decreased (Fig. 9B). Importantly, these effects were negated by depletion of SFMT1 (Fig. 9, A and B), demonstrating that SFMT1 is required for nickel-induced recruitment of LSD1 to epithelial genes and subsequent histone demethylation. Consistent with the epigenetic changes, expression of E-cadherin persisted in SFMT1-depleted cells under nickel treatment (Fig. 9C). Although control cells acquired a mesenchymal cell morphology after exposure to nickel, SFMT1-depleted cells remained epithelial, indicating that SFMT1 is required for nickel-induced EMT.

**SFMT1 Is Essential for Nickel-induced EMT**—Many environmental pollutants (including some heavy metals) are carcinogenic. Exposure to nickel and its compounds is fairly extensive in both environmental and occupational settings. The metal carcinogenicity is in large part attributed to their epigenetic effect (47), although the underlying mechanism remains poorly understood. Carcinogenic nickel compounds induce EMT and promote malignant progression during carcinogenesis (48). Exposure of A549 cells to nickel chloride (NiCl2) strongly up-regulated Snai2 and concomitantly down-regulated epithelial markers including E-cadherin (Fig. 8A).

Induction of Snai2 prompted us to determine whether nickel-induced EMT might involve LSD1-SFMT1-dependent epigenetic regulation. ChIP analysis showed that after nickel treatment, the binding of SFMT1 and LSD1 to epithelial gene promoters was potently increased (Figs. 8B and 9A, respectively), whereas H3K4me2 levels at these promoters significantly decreased (Fig. 9B). Importantly, these effects were negated by depletion of SFMT1 (Fig. 9, A and B), demonstrating that SFMT1 is required for nickel-induced recruitment of LSD1 to epithelial genes and subsequent histone demethylation. Consistent with the epigenetic changes, expression of E-cadherin persisted in SFMT1-depleted cells under nickel treatment (Fig. 9C). Although control cells acquired a mesenchymal cell morphology after exposure to nickel, SFMT1-depleted cells remained epithelial, indicating that SFMT1 is required for nickel-induced EMT.

**FIGURE 7.** **SFMT1 is required for TGFβ-induced repression of epithelial genes and EMT.** A, induction of Snai2 by TGFβ. A549 epithelial lung cancer cells were treated with TGFβ, and RNA levels of Snai1/2 were subsequently determined by RT-PCR. C and D, SFMT1 is required for down-regulation of epithelial markers by TGFβ. A549 cells were transduced with lentiviral vector pLKO or shRNA targeting SFMT1 followed by TGFβ treatment. Cells were harvested to measure E-cadherin and KRT8 RNA levels by RT-PCR (C) as well as E-cadherin and N-cadherin protein levels by Western blotting (D). Error bars indicate S.D. from triplicate experiments. E, SFMT1 is required for TGFβ-induced morphological changes in A549 cells. Phase-contrast images of control (pLKO) and SFMT1-depleted A549 cells treated with TGFβ for 2 days. F, depletion of SFMT1 blocked TGFβ-induced cell invasion. Control and SFMT1-depleted A549 cells were treated with TGFβ for 2 days and plated into Transwell (pre-coated with fibronectin). One day later cells invaded through the membranes were counted.

**FIGURE 8.** Nickel treatment up-regulates Snai2, down-regulates epithelial markers, and enhances binding of SFMT1 to epithelial gene promoters. A, induction of Snai2 and repression of epithelial genes E-cadherin (Ecad), KRT8, claudin 7 (CLDN7) by nickel treatment. A549 epithelial lung cancer cells were treated with nickel for 2 days, and RNA levels of the indicated genes were subsequently determined by RT-PCR. Error bars indicate S.D. from triplicate experiments. B, enhanced binding of SFMT1 to epithelial gene E-cadherin and KRT8 promoters by nickel treatment. A549 cells were treated with nickel for 2 days followed by ChIP assay with IgG (control) or antibodies for SFMT1.
completed cells maintained the epithelial phenotype (Fig. 9D). These observations suggest that SFMBT1 is necessary for induction of EMT by nickel.

**SFMBT1 Is Associated with Mesenchymal Markers and Poor Prognosis in Human Cancer**—EMT is believed to facilitate tumor cell dissemination and contribute to carcinoma progression and metastasis. EMT inducers Snai1 and Snai2 are associated with tumor metastasis and an adverse clinical outcome in a variety of human cancers (24). Given the role of SFMBT1 in transcriptional repression by Snai1 and induction of EMT, we sought to investigate its potential relevance in human cancer. We chose a publicly available gene expression data set GSE25055, which represents a relatively large cohort (comprising 310 pretreatment breast cancer biopsies) with information of distant-relapse-free survival (49). We analyzed the correlation of SFMBT1 with established epithelial and mesenchymal markers. Consistent with its role in EMT, SFMBT1 expression positively correlated with mesenchymal markers including N-cadherin and vimentin and negatively correlated with epithelial markers including E-cadherin in human breast cancer (Fig. 10A).

Using SFMBT1 and the associated epithelial/mesenchymal markers as a gene signature, the cancer samples were clustered accordingly (Fig. 10A). Patients grouped into the high expression tertile of the mesenchymal markers (including SFMBT1) (group II) had significantly shorter recurrence-free survival than the low expression group (group I) ($p = 0.001$) (Fig. 10B), suggesting this gene signature has prognostic value. The findings implicate SFMBT1 in EMT and tumor progression in human cancer.

**DISCUSSION**

The Snail family transcription factors are central regulators of EMT and are implicated in tumor progression and metastasis (24). To gain insights into the epigenetic mechanism underlying their transcriptional repression, we used a proteomic approach to purify protein complexes associated with Snai1. The LSD1 demethylase complex was found to be the predomi-
SFMBT1 Is Essential for LSD1 Binding to Chromatin and EMT

SFMBT1 is essential for LSD1 binding to chromatin and EMT. SFMBT1 forms a complex with LSD1. MBT domain is a structural motif consisting of ~100 amino acids, and its overall structure is highly conserved. MBT domains form a narrow and deep cavity of aromatic residue-lined pocket that specifically recognizes low methylated lysines (35). However, in SFMBT1, such methyl-lysine recognition residues are not all conserved, making it possible to form a larger pocket to accommodate trimethyllysine. Interestingly, the Caenorhabditis elegans MBT protein LIN-61 was shown to preferentially bind trimethyl H3K9 (52). Future structural studies are warranted to delineate the methyl lysine binding features of SFMBT1.

Chromatin remodeling and modifying complexes are recruited to chromatin primarily through interactions with sequence-specific DNA binding transcription factors. Such complexes typically contain reader subunits that recognize specific posttranslational modifications of histone tails, which are docking signals independent of underlying DNA sequence. Successful loading of chromatin regulatory complexes at specific genomic loci requires multivalent bindings of transcription factors to specific DNA sequence as well as histone readers to local chromatin environment. Cooperation among these multiple weak interactions not only enhances overall binding strength but also achieves avid specificity. Based on the current study, a multivalent engagement model assembled by Snail-LSD1-RCOR1-SFMBT1 is expected to be critical for recruiting the complex to chromatin template during Snail-induced epigenetic reprogramming (Fig. 6D): before Snail’s action, epithelial genes are actively transcribed in epithelial cells and possess di- and trimethyl H3K4 at their promoters. The LSD1 complex is present in these cells but may only have limited affinity toward the epithelial genes (possibly due to recognition of these marks by SFMBT1).

When Snail is induced in epithelial cells, it recognizes binding sites at many actively transcribed epithelial gene promoters. Snail recruits the LSD1 complex to these genomic sites via physical interaction with LSD1 (22); RCOR1 dimerizes with LSD1 and tethers it to nucleosome by contacting DNA backbone (31–33); SFMBT1 associates with RCOR1 as well as binds to the pre-existing active marks H3K4me2/3 present at the promoters of target genes. All these interactions cooperatively contribute to the initial recruitment of the LSD1 repressor complex to Snail1 target genes and/or stabilization of the complex on chromatin and are required for LSD1-mediated H3K4 demethylation. LSD1 specifically demethylates H3K4me2 to an unmethylated form, generating a recognition surface for BHC80 (another subunit) that stabilizes the LSD1 complex on chromatin and may maintain an unmethylated H3K4 status and the repressed state (51). The H3K4me3 mark persists during the process (Ref. 22 and Fig. 6A) and may continue to serve as an anchor for SFMBT1 binding and help retain the complex on chromatin. These events lead to an activation-to-repression transition at the Snail1 target promoters and reset the chromatin state at these loci. The LSD1 complex can be recruited to a variety of target genes through interactions with many different transcription factors. It is likely that SFMBT1 participates in most, if not all, LSD1-mediated epigenetic regulation.

It was unexpected that SFMBT1 was able to bind tri-methyl H3K4. There are nine MBT members in human and most of them can be evolutionary linked to one of three Drosophila orthologs: dL(3)MBT (the founding member of the MBT family), dSCM, and dSFMBT (Fig. 1B). Despite the nomenclature, human SFMBT1 and related SFMBT2 form a rather distant subgroup without an obvious ortholog in fly (Fig. 1B). Different groups of MBT proteins exist in discrete large protein complexes and display divergent biological functions (35). Only SFMBT1 forms a complex with LSD1. MBT domain is a structural motif consisting of ~100 amino acids, and its overall structure is highly conserved. MBT domains form a narrow and deep cavity of aromatic residue-lined pocket that specifically recognizes low methylated lysines (35). However, in SFMBT1, such methyl-lysine recognition residues are not all conserved, making it possible to form a larger pocket to accommodate trimethyllysine. Interestingly, the Caenorhabditis elegans MBT protein LIN-61 was shown to preferentially bind trimethyl H3K9 (52). Future structural studies are warranted to delineate the methyl lysine binding features of SFMBT1.
Epigenetic therapies conventionally target the catalytic activities of histone regulatory complexes. Chromatin readers recognize histone posttranslational modifications and critically interpret their biological significance. The protein–protein interactions between the reader modules and histone modifications are amenable to small molecule inhibitors. Given their importance in epigenetic signaling and chemical tractability, chromatin readers have emerged as new therapeutic targets (53, 54). Small-molecule MBT antagonists that interfere with the MBT-histone interaction have been under development (54, 55). LSD1 is deregulated in cancer (26–30). The LSD1 complex contains two enzymatic subunits, LSD1 and HDAC1/2, that are currently explored as therapeutic targets (54, 56). SFMBT1 may represent another epigenetic vulnerability of this complex that can be exploited for pharmacological intervention of EMT-dependent cancer progression.

Acknowledgments—We thank M. Angela Nieto (Instituto de Neurociencias, Consejo Superior de Investigaciones Científicas-Universidad Miguel Hernández, Spain) for the Snail-ER construct and Jorg Bungert for critically reading the manuscript.

REFERENCES
SFMBT1 Is Essential for LSD1 Binding to Chromatin and EMT

53. Dawson, M. A., Kouzarides, T., and Huntly, B. J. (2012) Targeting epige-
The Malignant Brain Tumor (MBT) Domain Protein SFMBT1 Is an Integral Histone Reader Subunit of the LSD1 Demethylase Complex for Chromatin Association and Epithelial-to-mesenchymal Transition

Ming Tang, Huangxuan Shen, Yue Jin, Tong Lin, Qingsong Cai, Melissa A. Pinard, Shyamasri Biswas, Quyen Tran, Guangyao Li, Anitha K. Shenoy, Emily Tongdee, Shuibin Lin, Yumei Gu, Brian K. Law, Lei Zhou, Robert Mckenna, Lizi Wu and Jianrong Lu

J. Biol. Chem. 2013, 288:27680-27691.
doi: 10.1074/jbc.M113.482349 originally published online August 8, 2013

Access the most updated version of this article at doi: 10.1074/jbc.M113.482349

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 56 references, 17 of which can be accessed free at http://www.jbc.org/content/288/38/27680.full.html#ref-list-1