Identification and Characterization of the Human Set1B Histone H3-Lys\(^4\) Methyltransferase Complex*

Jeong-Heon Lee\(^1\), Courtney M. Tate\(^2\), Jin-Sam You\(^3\), and David G. Skalnik\(^1\)

From the \(^4\)Wells Center for Pediatric Research, Section of Pediatric Hematology/Oncology, Departments of Pediatrics and Biochemistry and Molecular Biology, Indiana University School of Medicine, Indianapolis, Indiana 46202 and the \(^5\)Indiana Centers for Applied Protein Sciences, Indianapolis, Indiana 46202

We previously identified a mammalian Set1A complex analogous to the yeast Set1/COMPASS histone H3-Lys4 methyltransferase complex (Lee, J.-H., and Skalnik, D. G. (2005) J. Biol. Chem. 280, 41725–41731). Data base analysis indicates that human Set1A protein shares 39% identity with an uncharacterized SET domain protein, KIAA1076, hereafter denoted Set1B. Immunoprecipitation and mass spectrometry reveal that Set1B associates with a ~450 kDa complex that contains all five non-catalytic components of the Set1A complex, including CFP1, Rbbp5, Ash2, Wdr5, and Wdr82. These data reveal two human protein complexes that differ only in the identity of the catalytic histone methyltransferase. In vitro assays demonstrate that the Set1B complex is a histone methyltransferase that produces trimethylated histone H3 at Lys\(^4\). Both Set1A and Set1B are widely expressed. Inducible expression of the carboxyl terminus of either Set1A or Set1B decreases steady-state levels of both endogenous Set1A and Set1B protein, but does not alter the expression of the non-catalytic components of the Set1 complexes. A 123-amino acid fragment upstream of the Set1A SET domain is necessary for interaction with CFP1, Ash2, Rbbp5, and Wdr5. This protein domain is also required to mediate feedback inhibition of Set1A and Set1B expression, which is a consequence of reduced Set1A and Set1B stability when not associated with the methyltransferase complex. Confocal microscopy reveals that Set1A and Set1B each localize to a largely non-overlapping set of euchromatic nuclear speckles, suggesting that Set1A and Set1B each bind to a unique set of target genes and thus make non-redundant contributions to the epigenetic control of chromatin structure and gene expression.

Covalent modification of histone proteins, including methylation, acetylation, ubiquitination, and phosphorylation, confers critical epigenetic information that controls chromatin structure and regulation of gene expression (1–4). A large number of histone methyltransferases contain SET domains and catalyze the addition of methyl groups to lysine residues. Prominent sites of mammalian histone methylation include histone H3-Lys\(^4\), which is associated with euchromatin, and methylation of histone H3-Lys\(^9\), which is associated with heterochromatin. The yeast Set1 protein is the sole yeast histone H3-Lys\(^4\) methyltransferase and associates with a multimeric complex denoted COMPASS (5–7). Set1 is recruited by the RNA polymerase II machinery to actively expressed genes, and subsequent trimethylation of histone H3-Lys\(^4\) at these sites provides a localized mark of recent transcriptional activity (5–9).

Although yeast express only a single histone H3-Lys\(^4\) methyltransferase, mammalian cells contain approximately 12 histone methyltransferases that exhibit this specificity (10). The Set1-like family of histone H3-Lys\(^4\) methyltransferases includes Set1A, MLL, MLL2, MLL3, and MLL4 (10). Similar to the yeast Set1 protein, human MLL is localized to the 5’-end of actively expressed genes (11).

A major unanswered question concerns the significance of the complexity of mammalian histone H3-Lys\(^4\) methyltransferases. Although generally widely expressed, these mammalian methyltransferases provide non-redundant functions, as loss of a single member of the family can lead to disease or death. For example, chromosomal translocations involving the gene encoding the MLL histone H3-Lys\(^4\) methyltransferase are frequently found in leukemia (12–17); genetic disruption of the MLL or MLL2 genes leads to embryonic lethality (10, 18); and depletion of the SMYD3 histone H3-Lys\(^4\) methyltransferase by small interfering RNA treatment leads to suppression of cell growth (19). It is likely that non-redundant function of each histone H3-Lys\(^4\) methyltransferase is a result of distinct target gene specificity, but the nature of these gene targets and the mechanisms utilized to achieve unique subnuclear targeting of each methyltransferase are largely unknown.

We previously cloned CXXC finger protein 1 (CFP1)\(^3\) (20), the mammalian homologue of yeast Spp1, a component of the Set1/COMPASS histone H3-Lys\(^4\) methyltransferase complex. Biochemical purification of the human CFP1 complex revealed it to be analogous to the yeast COMPASS complex. The CFP1 structure and regulation of gene expression (1–4). A large number of histone methyltransferases contain SET domains and catalyze the addition of methyl groups to lysine residues. Prominent sites of mammalian histone methylation include histone H3-Lys\(^4\), which is associated with euchromatin, and methylation of histone H3-Lys\(^9\), which is associated with heterochromatin. The yeast Set1 protein is the sole yeast histone H3-Lys\(^4\) methyltransferase and associates with a multimeric complex denoted COMPASS (5–7). Set1 is recruited by the RNA polymerase II machinery to actively expressed genes, and subsequent trimethylation of histone H3-Lys\(^4\) at these sites provides a localized mark of recent transcriptional activity (5–9).

Although yeast express only a single histone H3-Lys\(^4\) methyltransferase, mammalian cells contain approximately 12 histone methyltransferases that exhibit this specificity (10). The Set1-like family of histone H3-Lys\(^4\) methyltransferases includes Set1A, MLL, MLL2, MLL3, and MLL4 (10). Similar to the yeast Set1 protein, human MLL is localized to the 5’-end of actively expressed genes (11).

A major unanswered question concerns the significance of the complexity of mammalian histone H3-Lys\(^4\) methyltransferases. Although generally widely expressed, these mammalian methyltransferases provide non-redundant functions, as loss of a single member of the family can lead to disease or death. For example, chromosomal translocations involving the gene encoding the MLL histone H3-Lys\(^4\) methyltransferase are frequently found in leukemia (12–17); genetic disruption of the MLL or MLL2 genes leads to embryonic lethality (10, 18); and depletion of the SMYD3 histone H3-Lys\(^4\) methyltransferase by small interfering RNA treatment leads to suppression of cell growth (19). It is likely that non-redundant function of each histone H3-Lys\(^4\) methyltransferase is a result of distinct target gene specificity, but the nature of these gene targets and the mechanisms utilized to achieve unique subnuclear targeting of each methyltransferase are largely unknown.

We previously cloned CXXC finger protein 1 (CFP1)\(^3\) (20), the mammalian homologue of yeast Spp1, a component of the Set1/COMPASS histone H3-Lys\(^4\) methyltransferase complex. Biochemical purification of the human CFP1 complex revealed it to be analogous to the yeast COMPASS complex. The CFP1
Characterization of the Set1B Histone Methyltransferase Complex

complex contains human homologues of the COMPASS complex, including Set1A, Wdr5, Ash2, Rbbp5, and Wdr82 (previously denoted hSwd2) (21). The human Set1A-CFP1 complex exhibits histone H3-Lys4 methyltransferase activity in vitro.

Data base analysis reveals a predicted human protein that exhibits high homology to the human Set1A protein. The purpose of the studies reported here was to examine whether this novel protein, denoted Set1B, is a functional histone methyltransferase, whether it associates with a complex similar to COMPASS, and to assess possible functional interactions between these two Set1-like enzymes. The data presented demonstrate that Set1B associates with a complex indistinguishable from the Set1A complex (except for the identity of the catalytic methyltransferase), and that these histone methyltransferase complexes exhibit both overlapping and non-redundant properties.

EXPERIMENTAL PROCEDURES

Cell Lines—Human embryonic kidney cells (HEK293) were cultured and transfected as previously described (22), and transfected cells were selected in 200 μg/ml hygromycin B. Established stably transfected cell lines were maintained in Dulbecco’s modified Eagle’s medium supplemented with 10% bovine calf serum and 50 μg/ml hygromycin B. A T-REx HEK293 cell line (Invitrogen) that constitutively expresses the tetracycline repressor was maintained in Dulbecco’s modified Eagle’s medium supplemented with 10% bovine calf serum and 5 μg/ml blasticidin. To establish inducible cell lines, T-REx HEK293 cells were transfected, selected in 200 μg/ml hygromycin B and 5 μg/ml blasticidin, and maintained in media containing 50 μg/ml hygromycin B and 5 μg/ml blasticidin. Mouse embryonic stem (ES) cells were cultured and differentiated as previously described (23).

Plasmid Construction—Expression vectors carrying human Set1A, CFP1, Ash2, Rbbp5, Wdr5, and Wdr82 cDNAs were prepared using the pcDNA3.1/Hygro vector that carries an amino-terminal FLAG epitope, as previously described (21). Various deletion constructs of human Set1A were subcloned using the PCR and restriction enzyme digestion into the pcDNA5/TO vector (Invitrogen), which encodes an amino-terminal FLAG epitope. A partial cDNA of human Set1B (KIAA1076, accession number AB028999) was obtained from the Kazusa DNA Research Institute. The cDNA encoding amino acid (aa) residues 1120–1923 was subcloned into pcDNA5/TO vector, which encodes an amino-terminal FLAG epitope. The nucleotide sequence of all plasmid constructs was confirmed by DNA sequencing. The nucleotide sequences of oligonucleotide PCR primers are available upon request.

Purification of the Set1B Complex and Identification of Its Components—A T-REx HEK293 cell line that inducibly expresses FLAG-Set1B (aa 1120–1923) was treated with 1 μg/ml of doxycycline for 4 days to induce protein expression. Nuclear extracts were prepared from cells (5 ml cell pellet volume) carrying the FLAG-Set1B or parental expression vector and used for FLAG immunoprecipitation as described (21). Bound proteins were eluted twice with 0.5 ml of binding buffer (10 mM PIPES, pH 7.0, 300 mM NaCl, 300 mM sucrose, 3 mM MgCl₂, 1 mM EGTA, supplemented with the protease inhibitors leupeptin, aprotinin, pepstatin (1 μg/ml each), and 1 mM phenylmethylsulfonyl fluoride, and 0.5% Triton X-100) containing 250 μg/ml FLAG peptide. The eluants were combined and loaded onto 10–50% sucrose gradients and subjected to centrifugation at 38,000 × g for 18 h using an SW41 rotor (Beckman). Five-hundred-milliliter fractions were collected. Molecular weight markers were applied to a parallel gradient, and their migration was analyzed by Coomassie Blue staining. Fractions enriched for Set1B were identified by Western blotting analysis and were pooled (fractions 11–14). The pooled sample was then subjected to a second round of FLAG immunoprecipitation. Bound proteins were eluted twice with 0.25 ml of binding buffer containing 250 μg/ml FLAG peptide, and the eluant was concentrated using an Amicon Ultra-4 concentrator (100 kDa cutoff). The sample was denatured and separated by 4–12% SDS-PAGE. The gel was stained with Coomassie Brilliant Blue or subjected to silver staining. Protein bands were excised and subjected to in-gel trypsin digestion using standard protocols. Peptides were extracted with 0.1% trifluoroacetic acid for 30 min at 37 °C and injected onto a 75 μm × 5-cm C-18 reverse-phase column (Waters, Picofrit column), and were eluted with a 5–45% acetonitrile gradient developed over 30 min at a flow rate of 250 nl/min using an Agilent 1100 series nanopump. The column was interfaced with a LTQ ion trap mass spectrometer (Thermo), and data were collected in the triple-play mode. MS/MS spectra were searched against the IPI human protein data base using SEQUEST.

Generation of Polyclonal Antisera—Anti-Set1A, anti-Set1B, anti-CFP1, anti-Wdr5, and anti-Wdr82 antisera were generated in rabbits (Proteintech Group) by using as antigens a recombinant GST-Set1A fragment (aa 258–458), recombinant GST-Set1B fragment (aa 1444–1596), recombinant GST-CFP1 fragment (aa 213–367), recombinant GST-Wdr5 fragment (aa 1–334), and recombinant GST-Wdr82 fragment (aa 1–173). Each human antigen was expressed in Escherichia coli and affinity purified. Goat antisera recognizing Set1A was also similarly generated against a recombinant GST-Set1A fragment (aa 795–993) (Proteintech Group).

Immunoprecipitation and Western Blotting Analysis—Nuclear extracts were incubated with anti-FLAG M2-agarose beads (Sigma) for 3 h and extensively washed. Bound proteins were eluted with SDS sample buffer. Anti-FLAG (mouse monoclonal M2) antibody was obtained from Sigma, anti-Ash2 (rabbit polyclonal) and anti-Rbpb5 (rabbit polyclonal) antisera were obtained from Bethyl Laboratories, anti-Brg1 (rabbit polyclonal) antiserum was obtained from Santa Cruz Biotechnology, anti-H3K4me2 (rabbit polyclonal) and anti-H3K9me2 (rabbit polyclonal) antisera were obtained from Upstate Biotechnology, and anti-H3K4me3 (rabbit polyclonal) and anti-histone H3 (rabbit polyclonal) antisera were obtained from Abcam.

Pulse-Chase Analysis of Protein Stability—T-REx HEK293 cell lines carrying empty vector or FLAG-Set1A fragment (aa 1082–1707) were treated with doxycycline for 4 days. Cells were rinsed twice with phosphate-buffered saline (PBS) and T-REx medium lacking methionine and cysteine, and were incubated in medium lacking methionine and cysteine for 15 min. Cells were then metabolically labeled for 45 min in

JOURNAL OF BIOLOGICAL CHEMISTRY

VOLUME 282 • NUMBER 18 • MAY 4, 2007

13420
Characterization of the Set1B Histone Methyltransferase Complex

Histone Methyltransferase Assay—Histone methyltransferase assays were performed as described previously (21, 24). Either 10 μg of purified core histones (Upstate Biotechnology) or 2 μg of recombinant human histone H3 purified from E. coli (Upstate Biotechnology) were used for methyltransferase reactions. Reaction products were analyzed by SDS-PAGE followed by Coomassie Blue staining and fluorography. To assess methyltransferase specificity, reaction products were transferred to polyvinylidene difluoride membrane and analyzed by Western blotting using modification-specific antisera.

Analysis of Set1A and Set1B Expression—Total RNA was isolated from undifferentiated ES cells, differentiated embryoid bodies (5 and 10 days following the removal of leukemia inhibitory factor), and HEK293, HEL, K562, Jurkat, and HeLa cells using TriReagent solution (Molecular Research Center) per the manufacturer’s recommended protocol. Total RNA (5 μg) was reverse transcribed using Moloney murine leukemia virus reverse transcriptase (Invitrogen) and random hexamers (Roche) at 42 °C for 60 min. Single-stranded cDNA (0.1 μg) was amplified in a 50-μl reaction mixture that included 0.2 μM of each deoxynucleoside triphosphate, 50 pmol of sense and antisense primers, and 1 unit of Tag DNA polymerase (Roche) in buffer supplied by the manufacturer. Samples were heat denatured at 94 °C for 2 min, followed by 30 cycles at 94 °C for 1 min, 55 °C for 1 min, 72 °C for 1 min, and finally 10 min at 72 °C. Amplified DNA was subjected to restriction enzyme digestion and analyzed by agarose gel electrophoresis along with the undigested amplified DNA. PCR primers were selected that amplify both Set1A and Set1B cDNA. Human PCR products were subjected to restriction enzyme digestion with PmlI, which does not cleave the Set1A PCR product but generates 189- and 261-bp fragments of the Set1B PCR product. Murine PCR products were subjected to restriction enzyme digestion with AflIII, which does not cleave the Set1A PCR product but generates 1563–1707), 60% identical and 78% similar throughout the mammalian proteins exhibit significant similarity to the yeast Set1 protein, including Set1A (25), KIAA1076, MLL (26), MLL2 (27), MLL3 (28), and MLL4 (28). Data base analysis indicates that yeast Set1 is most closely related with human Set1A and the uncharacterized KIAA1076 protein (10). Hereafter, we will refer to KIAA1076 as Set1B.

RESULTS

KIAA1076 Is Highly Homologous to Set1A—We previously reported that the human Set1A complex is analogous to the yeast Set1-COMPASS complex (21), which is the sole histone H3-Lys4 methyltransferase in yeast. Several mammalian proteins exhibit significant similarity to the yeast Set1 protein, including Set1A (25), KIAA1076, MLL (26), MLL2 (27), MLL3 (28), and MLL4 (28). Data base analysis indicates that yeast Set1 is most closely related with human Set1A and the uncharacterized KIAA1076 protein (10). Hereafter, we will refer to KIAA1076 as Set1B.

Human Set1A and Set1B proteins share 35 and 37% identity to yeast Set1, respectively, and human Set1A and Set1B proteins exhibit 39% identity and 56% similarity (Fig. 1). Human Set1A and Set1B proteins are 85% identical and 97% similar throughout the catalytic SET and post-SET domains (Set1A residues 1563–1707), 60% identical and 78% similar throughout the region upstream of the SET domain (N-SET domain, Set1A residues 1414–1562), 46% identical and 66% similar in a central...
Characterization of the Set1B Histone Methyltransferase Complex

region (Set1A residues 786–916), and 61% identical and 83% similar throughout an NH2-terminal region that includes the RNA recognition motif domain (Set1A residues 85–173) (7). Set1A additionally contains an HCF-1 binding motif that interacts with HCF-1 \textit{in vivo} (25), but Set1B lacks this HCF-1 binding motif domain. The extensive homology between Set1A and Set1B, particularly throughout the SET domain, suggests that Set1B functions as a histone methyltransferase.

**Identification of the Set1B Complex**—Sucrose gradient equilibrium centrifugation was performed to analyze the size of the

---

**FIGURE 1. Human Set1A is highly homologous to KIAA1076.** Amino acid alignments between Set1A (NP_055527) and KIAA1076 (Set1B, XP_037523) were analyzed by ClustalW (1.82) (34). Asterisks denote sequence identity and dots indicate similarity between the two proteins. RNA recognition motif, HCF-1 binding motif, N-SET, and post-SET domains are shaded, and the SET domain is boxed.
Characterization of the Set1B Histone Methyltransferase Complex

FIGURE 2. Identification of the Set1B complex. A, sucrose gradient centrifugation analysis of the Set1 complexes. Nuclear extracts from HEK293 cells stably transfected with FLAG-CFP1, a component of the Set1A complex (21), were immunoprecipitated with anti-FLAG antibody. Bound proteins were eluted with 250 μg/ml FLAG peptide and the eluant was analyzed by 10–50% sucrose gradient equilibrium centrifugation as described (21). An equal volume of each fraction was analyzed by Western blotting using the indicated antisera. B, flow chart of the procedure used to purify and characterize the Set1B complex. C, identification of proteins that interact with the carboxyl-terminal region of Set1B. T-REx HEK293 cell lines carrying the native Set1B complex, and to compare it with the Set1A complex. Nuclear extracts isolated from cells expressing FLAG-CFP1, a component of the Set1A complex (21), were prepared and subjected to FLAG immunoprecipitation. Bound proteins were eluted with FLAG peptide and separated on a 10–50% sucrose gradient centrifugation, and gradient fractions were analyzed by Western blotting. Set1B and Set1A proteins were both co-immunoprecipitated with FLAG-CFP1 and co-migrate at ~450 kDa (Fig. 2A).

Because the carboxyl-terminal SET domain of the MLL protein interacts with three protein components of the Set1A-CFP1 complex (Ash2, Rbbp5, and Wdr5) (26, 29), we reasoned that the highly conserved carboxyl terminus of Set1B interacts with components of its cognate complex. The carboxyl-terminal region of Set1B (aa 1120–1923) was conditionally expressed in T-REx HEK293 cells. To purify Set1B-associated proteins, cells were induced with doxycycline for 4 days and nuclear extracts were prepared (Fig. 2B). Nuclear extracts were subjected to FLAG immunoprecipitation, bound proteins were eluted by FLAG peptide, and eluants were separated on 10–50% sucrose gradients. Gradient fractions enriched for the FLAG-Set1B carboxyl-terminal fragment were pooled and subjected to a second round of FLAG immunoprecipitation and then visualized by SDS-PAGE. Several protein bands were specifically associated with the FLAG-tagged Set1B fragment, and were not apparent in samples that were isolated from T-REx HEK293 cells transfected with empty expression vector (Fig. 2C). Set1B-interacting protein bands were excised and analyzed by mass spectrometry. These were determined to be components of the Set1A complex, including CFP1 (accession number NP_055408), Ash2 (accession number NP_004665.1), Rbbp5 (accession number NP_005048.1), and Wdr5 (accession number NP_060058.1).

To further evaluate the authenticity of the putative Set1B complex, nuclear extracts isolated from cells expressing the carboxyl terminus of FLAG-Set1B (aa 1120–1923) were subjected to FLAG immunoprecipitation followed by Western blot analysis. CFP1, Ash2, Rbbp5, and Wdr5 were detected by specific antisera, but were not detected in FLAG immunoprecipitates prepared from cells carrying the empty expression vector (Fig. 3A). Similar analysis was performed following immunoprecipitation of endogenous Set1A or Set1B proteins (Fig. 3B). Western analysis of immunoprecipitated material reveals the presence of all members of the previously described Set1 complex (21), including Wdr82. The failure to detect Wdr82 in co-immunoprecipitation studies using the carboxyl half of Set1B (Fig. 2C) indicates that Wdr82 interacts with the amino half of the Set1B protein. Similar to the inability of FLAG-tagged Set1B fragment to immunoprecipitate Set1A (Fig. 3A), immunoprecipitation of endogenous Set1B fails to pull down Set1A, and immunoprecipitation of endogenous Set1A fails to pull down Set1B (Fig. 3B). Thus, these methyltransferases are found in distinct complexes.

Reciprocal co-immunoprecipitation analysis was performed to further characterize the Set1B complex. Stably transfected HEK293 cell lines were established that constitutively express FLAG-Set1A, FLAG-CFP1, FLAG-Ash2, FLAG-Rbbp5, FLAG-Wdr5, or FLAG-Wdr82. Nuclear extracts were prepared from each cell line and subjected to FLAG immunoprecipitation. The association of FLAG-tagged protein with endogenous Set1A or Set1B proteins was detected by Western blot analysis using anti-Set1A and anti-Set1B antisera. As expected, Set1B is associated with CFP1, Ash2, Rbbp5, Wdr5, and Wdr82 (Fig. 3C). As previously reported (21), Set1A associates with CFP1, Ash2, Rbbp5, Wdr5, and Wdr82, but does not associate with Set1B protein. These results indicate that Set1B forms a complex analogous to the yeast Set1/COMPASS that is similar to but distinct from the human Set1A complex. Summation of the predicted size of the six complex components (~480 kDa) agrees well with the observed size of the Set1B complex (Fig. 2), suggesting a 1:1 stoichiometry for the subunits, similar to the Set1A complex (21).

In Vitro Histone Methyltransferase Activity of the Set1B Complex—The yeast Set1/COMPASS complex and related Set1-like mammalian complexes function as histone H3-Lys4 methyltransferases (6, 7, 21, 25–28, 30, 31). The enzymatic activity of the purified human Set1B complex was examined using both purified core histones and recombinant human histone H3 substrates. As expected, the Set1B complex exhibits histone H3 methyltransferase activity, whereas samples recov-
Characterization of the Set1B Histone Methyltransferase Complex

FIGURE 3. Confirmation of the Set1B complex. A, confirmation of proteins that were identified in the Set1B complex by mass spectrometry. T-REx HEK293 cell lines carrying the empty expression vector or expressing FLAG-Set1B (aa 1120–1923) were induced with 1 μg/ml of doxycycline for 4 days. Nuclear extracts were prepared, subjected to FLAG immunoprecipitation, and immunoprecipitates were analyzed by Western blotting using the indicated antisera. B, nuclear extracts isolated from HEK293 cells were subjected to immunoprecipitation using antisera directed against Set1A or Set1B, and associated proteins were detected by Western analysis using the indicated antisera and the TrueBlot detection system. C, nuclear extracts from HEK293 cells that stably express the indicated full-length FLAG-tagged Set1 complex components were immunoprecipitated with anti-FLAG antibody. The immunoprecipitates were analyzed by Western blotting using the indicated antisera. Asterisks in the FLAG Western blot indicate the expected size of each FLAG-tagged protein.

FIGURE 4. The Set1B complex is a histone H3-Lys4 methyltransferase. A, FLAG immunoprecipitation was performed on nuclear extracts isolated from HEK293 cells expressing FLAG-Set1B (aa 1120–1923) or carrying the empty expression vector. Immunoprecipitated material was incubated with core histones isolated from chicken erythrocytes or recombinant histone H3 in the presence of S-[methyl-3H]adenosylmethionine. Reaction products were resolved by 12% SDS-PAGE and examined by Coomassie Blue staining (lower panel) or fluorography (upper panel). Arrows indicate the position of histone H3. The diagonal lines visible in the Coomassie staining of histone H3 are cracks that occurred during gel drying. B, recombinant histone H3 was methylated using FLAG-immunoprecipitated material as in A, and reaction products were analyzed by Western blotting using the indicated modification-specific antisera. Core histones purified from HEK293 cells were used as a positive control.

Steady-state Levels of Set1A and Set1B Are Regulated through a Conserved Carboxyl-terminal Domain—To gain insight into the functional roles of the Set1A and Set1B methyltransferase complexes, we investigated whether Set1A and Set1B are differentially expressed in various murine tissues, human cell lines, and during murine ES cell differentiation. Expression of both methyltransferases was detected by RT-PCR and Western blot analysis in all examined cell lines (Fig. 5), and did not change upon induction of ES cell differentiation (Fig. 5C). ES cell differentiation was confirmed by demonstration of Oct4 suppression (23). Quantitative RT-PCR revealed ubiquitous expression of both Set1A and Set1B in all murine tissues examined (Fig. 5D). There were no significant differences in the pattern of expression between Set1A and Set1B.

Given their sequence similarity and overlapping expression patterns, experiments were conducted to determine whether there is regulatory cross-talk between the Set1A and Set1B methyltransferase complexes. Doxycycline-inducible cell lines were generated that carry an empty expression vector, or vectors expressing FLAG-Set1A (aa 1288–1707) or FLAG-Set1B (aa 1120–1923). Expression levels of each methyltransferase complex component were compared between untreated cells...
and following 4 days of doxycycline treatment (Fig. 6A). Doxycycline treatment of the cell line carrying the empty expression vector does not alter the expression level of any component of the methyltransferase complexes (Fig. 6A). However, expression of the carboxyl-terminal region of either Set1A or Set1B decreases the steady-state levels of both endogenous Set1A and Set1B, whereas other components of the complexes are unaffected.

Because Set1A and Set1B associate with a common group of proteins, the simplest explanation for the observed coordinate down-regulation upon overexpression of the carboxyl terminus of either Set1A or Set1B is that the Set1 fragment competes for binding with the full methyltransferase complex, thus resulting in endogenous Set1A and Set1B proteins that are not associated with the complex. The reduction in steady-state levels of the endogenous factors suggests that these proteins are unstable when not bound to the Set1 complex. This model was examined by co-immunoprecipitation and pulse-chase studies. Co-immunoprecipitation analysis reveals that very little of the endogenous Set1A and Set1B proteins remain associated with CFP1 following induction of the carboxyl fragment of Set1A (Fig. 6B). Instead, CFP1 efficiently associates with the Set1A carboxyl fragment (aa 1288–1707). In addition, pulse-chase analysis reveals that the half-life of both endogenous Set1A and Set1B is dramatically shortened following expression of the carboxyl fragment of Set1A. Both endogenous Set1A and Set1B exhibit a long half-life (far in excess of the 7 h time course) in cells carrying the empty vector, but become much less stable (half-lives of ~2 h) following induction of the Set1A carboxyl fragment (Fig. 6C).

Various truncated versions of the Set1A carboxyl domain were similarly analyzed to define the Set1A protein domain necessary for decreased expression of endogenous Set1A and Set1B (Fig. 7A). These studies reveal that the 1415–1538-amino acid region of Set1A is essential for regulation of the steady-state levels of Set1A and Set1B (Fig. 7B). This region is also referred to as the N-SET domain (7), and is highly conserved between Set1A and Set1B.
Characterization of the Set1B Histone Methyltransferase Complex

Importantly, the majority of Set1A and Set1B signals do not overlap, suggesting that these two similar methyltransferase complexes have different in vivo gene targets.

**DISCUSSION**

Mammalian cells contain a multitude of histone modifying enzymes, including numerous histone H3-Lys4 methyltransferases. The association between histone H3-Lys4 methylation and euchromatin and active gene expression is firmly established. However, the molecular mechanisms that control the targeting and activity of histone methyltransferase complexes are not well understood. In addition, the significance of multiple histone H3-Lys4 methyltransferases is unclear, although severe phenotypes that are produced upon loss of a single member of this family indicate that individual enzymes contribute unique functions. The studies reported here describe the identification and characterization of the novel human Set1B histone H3-Lys4 methyltransferase complex.

With the exception of the catalytic component, the subunit composition of the Set1B complex is identical to that of the previously described Set1A complex (21). These include proteins such as Rbbp5, Wdr5, and Ash2, which are common among all of the Set1-like family of histone H3-Lys4 methyltransferase complexes and comprise a structural platform to which each of the Set1-like proteins interact (29). In contrast, the Wdr82 and CFP1 proteins have only been detected in the Set1A and Set1B complexes (21). The function of most of the non-catalytic components within the mammalian Set1-like complexes is unclear, although Wdr5 has been found to recognize dimethylated histone H3-Lys4 (32, 33), and to mediate interaction between the catalytic methyltransferase subunit and the histone substrate (29). Furthermore, absence of Ash2, Rbbp5, or Wdr5 significantly inhibits the methyltransferase activity of the MLL complex (29). Analysis of murine ES cells that lack CFP1 reveals elevated levels of histone H3-Lys4 methylation, suggesting that this factor inhibits or restricts the activity of the Set1A and Set1B complexes (21). The behavior of these complexes in the absence of CFP1 requires further study.

Interestingly, overexpression of the carboxyl fragment of Set1A leads to down-regulation of both endogenous Set1A and Set1B proteins as a consequence of reduced stability of the Set1A and Set1B proteins when not associated with the methyltransferase complexes.
Characterization of the Set1B Histone Methyltransferase Complex

is consistent with the finding that the MLL enzyme continues to be targeted to appropriate genomic sites following individual depletion of the Ash2, Rbbp5, or Wdr5 components of the methyltransferase complex (29). In addition, proteins common to Set1-like complexes, such as Rbbp5 and Wdr5, fail to localize to MLL-target genes in the absence of MLL (29), indicating that other methyltransferase complexes that also contain Rbbp5 and Wdr5 fail to provide redundant function at these genomic sites. However, the distinct subnuclear distributions of the Set1A and Set1B proteins contrasts with the finding that the MLL member of the Set1-like family of histone H3-Lys4 methyltransferases is detected at 90% of RNA pol II-occupied genes (11). These results suggest that Set1A and Set1B each act to regulate a more restricted subset of target genes. Studies to catalogue the genomic binding sites of these factors and to determine the structural features of these proteins that confer enzyme-specific genomic targeting will provide important insights into the epigenetic control of chromatin structure and gene expression.

REFERENCES

1. Shilatifard, A. (2006) Annu. Rev. Biochem. 75, 243–269
2. Peterson, C. L., and Laniel, M. A.-A. (2004) Curr. Biol. 14, R546–R551
3. Roth, S. Y., Denu, J. M., and Allis, C. D. (2001) Annu. Rev. Biochem. 70, 81–120
4. Martin, C. C., and Zhang, Y. (2005) Nat. Rev. Mol. Cell Biol. 6, 838–849
5. Miller, T., Krogan, N. J., Dover, J., Ejdummet-Bromage, H., Tempst, P., Johnston, M., Greenblatt, J. F., and Shilatifard, A. (2001) Proc. Natl. Acad. Sci. U. S. A. 98, 12902–12907
6. Nagy, P. L., Griesenbeck, J., Kornberg, R. D., and Cleary, M. L. (2002) Proc. Natl. Acad. Sci. U. S. A. 99, 90–94
7. Roguev, A., Schaft, D., Shevchenko, A., Pin Pijnappel, W., Wilm, M., Aasland, R., and Stewart, A. F. (2001) EMBO J. 20, 7137–7148
8. Krogan, N. J., Dover, J., Wood, A., Schneider, J., Heidt, I., Boetang, M. A., Dean, K., Ryan, O. W., Golshani, A., Johnston, M., Greenblatt, J. F., and Shilatifard, A. (2003) Mol. Cell 11, 709–719
9. Glaeser, S., Schaft, J., Lubitsh, Z., Vintersten, K., van der Hoeven, F., Tufte, K., Raasland, A., Rastas, M., and Stewart, A. F. (2006) Development 133, 1423–1432
10. Quintiliani, M. G., Jenner, R. G., Chevalier, B., Nakamura, Y., and Young, R. A. (2005) Proc. Natl. Acad. Sci. U. S. A. 102, 8603–8608
11. Domer, P. H., Fakharzadeh, S. S., Chen, C. S., Jockel, J., Johansen, L., Silverman, G. A., Kersey, J. H., and Korsmeyer, S. J. (1993) Proc. Natl. Acad. Sci. U. S. A. 90, 7884–7888
12. Gu, Y., Nakamura, T., Alder, H., Prasad, R., Canaani, O., Cimino, G., Croce, C. M., and Canaani, E. (1992) Cell 71, 701–708
13. Gu, Y., Alder, H., Nelson, K. K., Chatterjee, D., Gu, Y., Nakamura, T., Canaani, E., Croce, C. M., Siracusa, L. D., and Buchberg, A. M. (1993) Proc. Natl. Acad. Sci. U. S. A. 90, 6350–6354
14. Prasad, R., Yano, T., Soro, C., Nakamura, T., Fellapali, R., Gu, Y., Leshkowitz, D., Croce, C. M., and Canaani, E. (1995) Proc. Natl. Acad. Sci. U. S. A. 92, 12160–12164
15. Tkachuk, D. C., Kohler, S., and Cleary, M. L. (1992) Cell 71, 691–700
16. Zalesznik-Le, N., Harden, A. M., and Rowley, J. D. (1994) Proc. Natl. Acad. Sci. U. S. A. 91, 10610–10614
17. Aytion, P., Sneddon, S. F., Palmer, D. B., Rosewell, I. R., Owen, M. J., Young, B., Presley, R., and Subramanian, V. (2001) Genesis 30, 201–212
18. Hamamoto, R., Furukawa, Y., Morita, M., Iimura, Y., Silva, F. P., Li, M., Yagyu, R., and Nakamura, Y. (2004) Natl. Cell. Biol. 6, 731–740
19. Voo, K. S., Carbone, D., L., Jacobsen, B. M., Flodin, A., and Skalnik, D. G. (2000) Mol. Cell. Biol. 20, 2108–2121
Characterization of the Set1B Histone Methyltransferase Complex

21. Lee, J.-H., and Skalnik, D. G. (2005) J. Biol. Chem. 280, 41725–41731
22. Lee, J.-H., and Skalnik, D. G. (2002) J. Biol. Chem. 277, 42259–42267
23. Lee, J.-H., Hart, S. R. L., and Skalnik, D. G. (2004) Genesis 38, 32–38
24. Strahl, B. D., Ohba, R., Cook, R. G., and Allis, C. D. (1999) Proc. Natl. Acad. Sci. U. S. A. 96, 14967–14972
25. Wysocka, J., Myers, M. P., Laherty, C. D., Eisenmann, R. N., and Herr, W. (2003) Genes Dev. 17, 896–911
26. Yokoyama, A., Wang, Z., Wysocka, J., Sanyal, M., Aufiero, D. J., Kitabayashi, I., Herr, W., and Cleary, M. L. (2004) Mol. Cell. Biol. 24, 5639–5649
27. Hughes, C. M., Rozenblatt-Rosen, O., Milne, T. A., Copeland, T. D., Levine, S. S., Lee, J. C., Hayes, D. N., Shanmugam, K. S., Bhattacharjee, A., Biondi, C. A., Kay, G. F., Hayward, N. K., Hess, J. L., and Meyerson, M. (2004) Mol. Cell 13, 587–597
28. Goo, Y.-H., Sohn, Y. C., Kim, D.-H., Kim, S.-W., Kang, M.-J., Jung, D.-J., Kwak, E., Barlev, N. A., Berger, S. L., Chow, V. T., Roeder, R. G., Azorsa, D. O., Meltzer, P. S., Suh, P.-G., Song, E. J., Lee, K.-J., Lee, Y. C., and Lee, I. W. (2003) Mol. Cell. Biol. 23, 140–149
29. Dou, Y., Milne, T. A., Ruthenburg, A. J., Lee, S., Lee, J. W., Verdin, G. L., Allis, C. D., and Roeder, R. G. (2006) Nat. Struct. Mol. Biol. 13, 713–719
30. Briggs, S. D., Bryk, M., Strahl, B. D., Cheung, W. L., Davie, J. K., Dent, S. Y. R., Winston, F., and Allis, C. D. (2001) Genes Dev. 15, 3286–3295
31. Krogan, N. J., Dover, J., Khorrami, S., Greenblatt, J. F., Schneider, J., Johnston, M., and Shilatifard, A. (2002) J. Biol. Chem. 277, 10753–10755
32. Han, Z., Guo, L., Wang, H., Shen, Y., Deng, X. W., and Chai, J. (2006) Mol. Cell 22, 137–144
33. Wysocka, J., Swigut, T., Milne, T. A., Dou, Y., Zhang, X., Brurlingame, A. L., Roeder, R. G., Brivanlou, A. H., and Allis, C. D. (2005) Cell 121, 859–872
34. Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T. J., Higgins, D. G., and Thompson, J. D. (2003) Nucleic Acids Res. 31, 3497–3500