CpG-binding Protein (CXXC Finger Protein 1) Is a Component of the Mammalian Set1 Histone H3-Lys4 Methyltransferase Complex, the Analogue of the Yeast Set1/COMPASS Complex*

Jeong-Heon Lee and David G. Skalnik

From the Herman B 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

CpG-binding protein (CXXC finger protein 1 (CFP1)) binds to DNA containing unmethylated CpG motifs and is required for mammalian embryogenesis, normal cytosine methylation, and cellular differentiation. Studies were performed to identify proteins that interact with CFP1 to gain insight into the molecular function of this protein. Immunoprecipitation and mass spectrometry reveal that human CFP1 associates with a ~450-kDa complex that contains the mammalian homologues of six of the seven components of the Set1/COMPASS complex, the sole histone H3-Lys4 methyltransferase in yeast. In vitro assays demonstrate that the human Set1/CFP1 complex is a histone methyltransferase that produces mono-, di-, and trimethylated histone H3 at Lys4. Confocal microscopy reveals that CFP1 and Set1 co-localize to nuclear speckles associated with euchromatin. A Set1 complex of reduced mass persists in murine embryonic stem cells lacking CFP1. These cells carry elevated levels of methylated histone H3-Lys4 and reduced levels of methylated histone H3-Lys9. Together with the previous finding of reduced levels of cytosine methylation, these data indicate that cells lacking CFP1 contain reduced levels of heterochromatin. Furthermore, ES cells lacking CFP1 exhibit a 4-fold excess of histone H3-Lys4 methylation following induction of differentiation, indicating that CFP1 restricts the activity of the Set1 histone methyltransferase complex. These results reveal a mammalian counterpart to the yeast Set1/COMPASS complex. The presence of CFP1 in this complex implicates this protein as a critical epigenetic regulator of histone modification in addition to cytosine methylation and reveals one mechanism by which this protein intersects with the epigenetic machinery.

CpG-binding protein exhibits a unique DNA binding specificity for unmethylated CpG motifs and acts as a transcriptional activator (1). This factor, encoded by the CXXC1 gene, has recently been designated CXXC finger protein 1, and will hereafter be referred to as CFP1.2 Originally identified in mammals, homologues of CFP1 have been detected in Drosophila, Caenorhabditis elegans, and both Saccharomyces cerevisiae and Schizosaccharomyces pombe (1, 2). CFP1 contains a cysteine-rich CXXC DNA-binding domain (1, 3), which is present in several other proteins, including DNA methyltransferase 1 (Dnmt1) (4), the major maintenance DNA methyltransferase; human trithorax (HRX) (also known as ALL-1 or MLL), a histone H3-Lys4 methyltransferase encoded by a gene frequently involved in chromosomal translocations in leukemia (5–10); methyl-binding domain protein 1 (MBD1), which is characteristic of chromatin-associated proteins and/or regulators of gene expression (11, 12); and mediate protein/protein interactions (17–19). CFP1 is a component of the nuclear matrix and localizes to nuclear speckles associated with euchromatin (20).

Targeted disruption of the CXXC1 gene results in peri-implantation embryonic lethality in mice (21), a developmental stage associated with global remodeling of chromatin structure and cytosine methylation patterns (22–24). CXXC1−/− embryonic stem (ES) cells exhibit a 60–80% decrease in global cytosine methylation, including hypomethylation of repetitive genomic elements, single copy genes, and imprinted genes (25). Maintenance DNA methyltransferase activity is reduced 60% in CXXC1−/− ES cells and expression of Dnmt1 protein is also reduced, although de novo DNA methyltransferase activity is normal (25). The cytosine methylation deficiencies observed in CXXC1−/− ES cells, although dramatic, cannot account for the severity of the observed phenotype. For example, ES cells lacking Dnmt1 exhibit a 90% reduction in Dnmt activity and cytosine methylation, yet exhibit normal growth prior to in vitro differentiation (26). In contrast, undifferentiated CXXC1−/− ES cells exhibit a lengthened doubling time as a consequence of an increased rate of apoptosis (25). In addition, mouse embryos lacking Dnmt1 die later in gestation (8.5–9.5 days post-coitus) (26) compared with CXXC1−/− embryos (4.5–6.5 days post-coitus) (21). Finally, decreased Dnmt1 activity and Dnmt1 protein levels in CXXC1−/− ES cells are unlikely to fully explain the observed deficiency in genomic cytosine methylation, as Dnmt1−− ES cells express reduced Dnmt1 protein but retain normal levels of cytosine methylation (27, 28).

The existence of CFP1 homologues in lower eukaryotes that lack CpG methylation, such as yeast and C. elegans, provides additional circumstantial evidence for a function of CFP1 that is independent of cytosine methylation. Interestingly, sequence alignment reveals that CFP1 homologues in organisms that lack CpG methylation lack the CXXC DNA-binding domain (data not shown). The yeast CFP1 homologue, Spp1, is a component of the multimeric Set1 histone H3-Lys4
CFP1 Is a Component of the Set1 Complex

methyltransferase complex (also known as COMPASS) (2, 29, 30). This is the sole histone H3-Lys4 methyltransferase in yeast and is also required for telomeric and rDNA silencing (31, 32). Set1 is recruited by the RNA polymerase II elongation machinery to actively expressed genes (33, 34). The subsequent deposition of H3K4me3 at these genomic sites provides a localized mark of recent transcriptional activity. Spy1 is dispensable for histone H3-Lys4 methyltransferase activity in S. cerevisiae but is necessary for methyltransferase activity in S. pombe (35).

Given the presence of yeast Spy1 within a histone-modifying complex and the severity of the CXXC1"/" phenotype in mice, we hypothesized that mammalian CFP1 may play a role in the control of histone modification and chromatin structure in addition to facilitating cytosine methylation. The studies reported here demonstrate that CFP1 is a component of a mammalian Set1 histone H3-Lys4 methyltransferase that is analogous to the yeast Set1/COMPASS complex. Furthermore, CXXC1"/" ES cells exhibit perturbed global patterns of histone modification, demonstrating that CFP1 is a critical epigenetic regulator of both cytosine methylation and histone methylation.

EXPERIMENTAL PROCEDURES

Cell Lines—Human embryonic kidney cells (HEK-293) were cultured and transfected as described (20), and transfected cells were selected in 200 μg/ml hygromycin B. Mouse ES cells were cultured and differentiated in vitro as described (22). ES cell lines either homozygous or heterozygous for a disrupted CXXC1 allele were generated from blastocysts derived from matings between CXXC1"/" mice (21), as described (25). CXXC1"/" ES cells were transfected with a CFP1 cDNA expression vector to produce rescued cells (CXXC1"/"/CFP1) or were transfected with empty expression vector as a control (CXXC1"/"/vector), as described previously (25).

Plasmid Constructions—FLAG-tagged human CFP1 cDNA was subcloned into pcDNA3.1/Hygro (Invitrogen). cDNA of human Set1 (KIAA0339) was obtained from the Kazusa DNA Research Institute. This cDNA clone carried a single nucleotide deletion, which was corrected by reverse transcription PCR amplification of this region and substitution into the cDNA clone. Ash2 (BC015936), Rbbp5 (BC001635) cDNA were amplified from expressed sequence-tagged clones obtained from the American Type Culture Collection. hSwd2 cDNA was amplified by reverse transcription PCR from HEK-293 RNA. The cDNA clones were subcloned into the pcDNA3.1/Hygro vector that contains an amino-terminal FLAG epitope. The cDNA for human CFP1 was subcloned into the pEGFP-C2 vector. The nucleotide sequence of all constructs was confirmed. Set1 was similarly cloned into the pEGFP-C2 vector. The nucleotide sequence of all constructs was confirmed.

Plasmid Constructions—FLAG-tagged human CFP1 cDNA was subcloned into pcDNA3.1/Hygro (Invitrogen). cDNA of human Set1 (KIAA0339) was obtained from the Kazusa DNA Research Institute. This cDNA clone carried a single nucleotide deletion, which was corrected by reverse transcription PCR amplification of this region and substitution into the cDNA clone. Ash2 (BC015936), Rbbp5 (BC037284), and Wdr5 (BC001635) cDNA were amplified from expressed sequence-tagged clones obtained from the American Type Culture Collection. hSwd2 cDNA was amplified by reverse transcription-PCR from HEK-293 RNA. The cDNA clones were subcloned into the pcDNA3.1/Hygro vector that contains an amino-terminal FLAG epitope. The cDNA of human CFP1 (1) was subcloned into the pEGFP-C2 and pEGFP-FLAG vectors (20). The cDNA for human Set1 was similarly cloned into the pEGFP-C2 vector. The nucleotide sequence of all constructs was confirmed.

Purification of a CFP1 Complex—HEK-293 cells expressing FLAG-CFP1 were incubated in hypotonic buffer for 10 min on ice and homogenized, and nuclei were harvested by centrifugation. Ten milliliters of extraction 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) was added and incubated for 10 min on ice. The pellets were extensively Dounce homogenized on ice. The solubilized nuclear fraction was separated by centrifugation and used for immunoprecipitation. FLAG-IgG-agarose slurry (Sigma) was added, incubated for 3 h, and washed five times with 12 ml of extraction buffer. Bound proteins were eluted twice with 0.5 ml of extraction buffer containing 250 μg/ml FLAG peptide. Combined eluants were loaded onto 10–50% sucrose gradients and centrifuged at 38,000 rpm for 18 h using a SW41 rotor (Beckman). Five-hundred-milliliter fractions were collected. Molecular markers were applied to a parallel gradient, and their migration was analyzed by Coomassie Blue staining. Fractions enriched for CFP1 (#9-12) were pooled and subjected to a second round of FLAG immunoprecipitation. Bound proteins were similarly eluted and concentrated using an Amicon Ultra-4 concentrator (100 kDa cut-off). Samples were separated by 4–12% SDS-PAGE and stained with Coomassie Brilliant Blue or silver stain, and protein bands were excised for in-gel trypsin digestion. Peptides were extracted and desalted on C18-ZipTips (Millipore) according to the manufacturer’s protocol and resuspended in 10 μl of 30% methanol, 0.5% acetic acid. The peptides were analyzed by MALDI-TOF mass spectrometry (Indiana University Protein Analysis and Research Center), and protein identity was determined using the ProFound search engine (129.85.19.192/profound_bin/WebProFound.exe).

Histone Methyltransferase Assay—Histone methyltransferase assays were performed on the immunoprecipitated Set1/CFP1 complex, as described previously (36). Either 10 μg of core histones purified from chicken erythrocytes (Upstate Biotechnology), 2 μg of recombinant histone H3 or recombinant histone H4 purified from Escherichia coli (Upstate Biotechnology), or 10 μg of core histones purified by acid extraction (22) from HEK-293 cells were used for methyltransferase reactions. Reaction products were analyzed by SDS-PAGE followed by Coomassie Blue staining and fluorography. Alternatively, reaction products were transferred to polyvinylidene difluoride membrane and analyzed by Western blotting using modification-specific antisera.

Co-immunoprecipitation and Western Blotting Analysis—Nuclear extract preparations and FLAG immunoprecipitations were performed as described above. For CFP1, Ash2, and Rbbp5 immunoprecipitations, nuclear extracts were precleared and incubated with antibodies for 1 h and incubated with protein-A-agarose beads (Roche Applied Science) for 3 h. Protein-A-agarose beads were washed four times with extraction buffer containing 300 mM NaCl. Proteins were eluted with SDS sample buffer and analyzed by Western blotting.

Polyclonal anti-CFP1 antibody has been described previously (1). Anti-FLAG antibody was obtained from Sigma; anti-H3K4me2 and anti-H3K9me2 antibodies were obtained from Upstate Biotechnology; anti-BRG1 antibody was obtained from Santa Cruz Biotechnology; anti-H3K4me3, anti-H3K9me3, anti-H3K4me1, and anti-histone H3 antibodies were obtained from Abcam; anti-Ash2 and anti-Rbbp5 antibodies were obtained from Bethyl Laboratories; and anti-Set1 amino-terminal and carboxyl-terminal antibodies were obtained from Abgent. Anti-Set1 antibody was a kind gift of Winship Herr (Cold Spring Harbor Laboratory). Anti-MEP50 antibody was a kind gift of Gideon Dreyfuss (University of Pennsylvania).

Core histones were isolated from ES cells as described (22). ES cells (~10⁶ cells) were harvested and washed with ice-cold phosphate-buffered saline (PBS) buffer containing 10 mM sodium butyrate. Cells were resuspended with 10 ml of cell lysis buffer and lysed by 10 strokes of a Dounce homogenizer. The nuclei were pelleted and resuspended in 5 ml of wash buffer (250 mM sucrose, 10 mM Tris-HCl (pH 7.4), 10 mM sodium butyrate, 4 mM MgCl2, 0.1 mM phenylmethylsulfonyl fluoride, 0.1 mM benzamidine). The nuclear suspension was overlaid onto 5 ml of a sucrose cushion (30% sucrose in wash buffer) and subjected to centrifugation at 24,000 × g for 5 min at 4 °C. Nuclei were resuspended in a minimum volume of wash buffer. For acid extraction of histones, an equal volume of ice-cold 0.8 M HCl was added and stirred on ice for 2 h, and the acid-insoluble material was removed by centrifugation at 12,000 × g for 10 min at 4 °C. The histone-containing supernatant was saved. The pellet was re-extracted as above, and the supernatants were
combined. The histones were precipitated by adding 8 volumes of acetone and incubating at −20 °C overnight. The histones were recovered by centrifugation. The histone pellets were washed twice with freshly prepared acetone/100 mM HCl (10:1) solution and then washed three times with acetone. The histone pellets were dried at room temperature and reconstituted with water. Protein concentration was determined by the Bradford method (37). For Western blot analysis, an equivalent amount of core histone was solubilized in Laemmli sample buffer. Following electrophoresis on 12% SDS-PAGE, proteins were analyzed by Western blotting.

Immunofluorescence and Confocal Microscopy—Cells were seeded onto a cover glass at 2–5 × 10^4 cells/well in a 24-well dish and grown for 24 h, fixed with 4% (v/v) paraformaldehyde in PBS for 20 min, and washed with PBS. Cells were permeabilized with 0.2% Triton X-100 in PBS for 10 min, and a blocking solution (PBS containing 2.5% normal serum and 0.2% Tween 20) was added and incubated for 1 h. Anti-FLAG antibody (1:1000) was added and incubated for 2 h. Cells were washed three times with PBS containing 0.2% Tween 20 for 5 min. Secondary antibody labeled with Texas Red (2 μg/ml in blocking solution) (Santa Cruz Biotechnology) was added and incubated for 1 h. Cells were washed three times with PBS containing 0.2% Tween 20 for 5 min. Nuclear counterstaining was performed with 0.1 μg/ml 4,6-diamidino-2-phenylindole (DAPI) in PBS for 5 min followed by washing with PBS. Cells were mounted with 10 μl of Fluoromount G (Southern Biotechnology Associates) and scanned with a Zeiss LSM 510 laser scanning confocal microscope.

RESULTS

CFP1 Is a Component of the Mammalian Set1 Complex—HEK-293 cells were stably transfected with an expression vector containing the FLAG-CFP1 cDNA under the control of the cytomegalovirus promoter/enhancer (or the empty expression vector). These cells express FLAG-CFP1 at a level ~80% of endogenous CFP1 (data not shown). Nuclear extracts isolated from these cells were subjected to sucrose gradient equilibrium centrifugation (Fig. 1A). Both endogenous CFP1 and FLAG-CFP1 are present in a complex of ~450 kDa. The migration of CFP1 is not affected by the presence of 500 mM NaCl (data not shown). A protein band corresponding to CFP1 was detected in sucrose gradient equilibrium centrifugation. Sucrose gradient fractions enriched for CFP1 were pooled, subjected to SDS-PAGE following FLAG immunoprecipitation (Fig. 1, B and C). Six major protein bands were detected. Analysis by mass spectrometry reveals that many of these proteins represent human homologues of components of the yeast Set1/COMPASS histone H3-Lys4 methyltransferase complex, including Set1 (XP_049380.2), homologue to yeast Set1 (BLAST score of 214); Ash2 (NP_004665.1), homologue to yeast Bre2 (BLAST score of 62); Wdr5 (also known as Big-3, NP_060058.1), homologue to the yeast WD repeat protein Swd3 (BLAST score of 162); hSwd2 (XP_293514.1), homologue of the yeast protein Swd2 (BLAST score of 168); and Rbbp5 (also known as RBQ-3, NP_005048.1), the homologous of the yeast WD40 repeat protein Swd1 (BLAST score of 150). Last, the WD40 domain protein MEP50 (38) was also detected by mass spectrometry, although this protein does not correspond to a Set1/COMPASS homologue.

Co-immunoprecipitation analysis was performed to further assess the authenticity of the putative Set1/CFP1 complex. Set1, Ash2, MEP50, and Rbbp5 were each detected by immunoblot following immunoprecipitation of FLAG-CFP1 (Fig. 1D). However, MEP50 was also detected following immunoprecipitation of the vector control extract, indicating that the presence of this factor in the original FLAG immunoprecipitation was an artifact. Antibodies directed against hSwd2 and Wdr5 were not available for these studies.

Reciprocal co-immunoprecipitation studies were also performed. Because immunoprecipitating antibodies directed against Set1, Wdr5, and hSwd2 were not available, FLAG-tagged expression vectors for each of these factors were stably introduced into HEK-293 cells. All six components of the CFP1 complex were detected by immunoblot following immunoprecipitation with antisera directed against CFP1, Ash2, or Rbbp5 (Fig. 2A). Similarly, FLAG immunoprecipitation of cells expressing FLAG-Set1, FLAG-Wdr5, or FLAG-hSwd2 led to recovery of CFP1, Set1, Ash2, and Rbbp5 (Fig. 2B). No signal was detected when immunoprecipitations were performed on HEK-293 cells transfected with an empty expression vector. The Set1/CFP1 complex was further analyzed by Western analysis following sucrose gradient fractionation. All six components of the Set1/CFP1 complex were detected in the ~450 kDa size range, although the majority of the FLAG-Wdr5 and FLAG-hSwd2 proteins was found in a smaller size range (Fig. 2C).

Previous work documented that CFP1 localizes to heterochromatin nuclear speckles (20). Co-localization studies using confocal microscopy demonstrated that CFP1 exhibits a subnuclear distribution that is entirely distinct from heterochromatin markers such as HP1α and McCP2 but partially overlapping with euchromatin markers such as acetylated histones, sites of transcription, and RNA splicing (20). Confocal microscopy was performed to determine the subnuclear localization of green fluorescent protein (GFP)-Set1 fusion protein in HEK-293 cells. GFP-Set1 fusion protein is found in nuclear speckles that are largely excluded from DAPI-bright regions (heterochromatin), indicat-
CFP1 Is a Component of the Set1 Complex

ing a euchromatin localization (Fig. 3A). Furthermore, consistent with the biochemical data presented above, FLAG-CFP1 and GFP-Set1 co-localize to sub-nuclear regions that are distinct from regions of DAPI-bright heterochromatin (Fig. 3B).

The Set1/CFP1 Complex Exhibits Histone H3-Lys4 Methyltransferase Activity—Because other mammalian complexes containing Set1-like proteins exhibit histone methyltransferase activity (15, 39–41), experiments were performed to assess the enzymatic activity and specificity of the Set1/CFP1 complex. Similar to the yeast Set1/COMPASS complex, the mammalian Set1/CFP1 complex exhibits histone H3 histone methyltransferase activity (Fig. 4). The purified Set1/CFP1 complex transfers tritiated methyl groups onto both purified core histones and recombinant histone H3 (Fig. 4A) (20). The Set1/CFP1 complex fails to catalyze methylation of histone H4 (Fig. 4A). The faint signal observed for histone H4 methylation is also present in the vector control lane and may be due to nonspecific immunoprecipitation of the arginine methyltransferase PRMT5. PRMT5 was detected by mass spectrometry as a co-migrating band with Ash2 and CFP1 following FLAG immunoprecipitation (data not shown). PRMT5 has previously been shown to immunoprecipitate non-specifically with anti-FLAG antibody (42) and to exhibit methyltransferase activity toward histone H4 (43).

Western analysis of reaction products following incubation of the Set1/CFP1 complex with recombinant histone H3 detects H3K4me1, H3K4me2, and H3K4me3, but not H3K9me2 (Fig. 4B), similar to what was reported for the yeast Set1/COMPASS complex (29–31, 44) and other mammalian Set1-like complexes (15, 39–41). This enzymatic specificity for histone H3-Lys4 methylation is also consistent with the observed localization of the Set1 and CFP1 proteins to euchromatin (Fig. 3) Comparison of the components of the human Set1/CFP1 complex with the yeast Set1/COMPASS complex shows a high degree of conservation (Fig. 5), including the presence of six of the seven Set1/COMPASS components within the Set1/CFP1 complex. Summation of the predicted size of the six complex components (∼470 kDa) agrees well with the observed size of the Set1/CFP1 complex (Fig. 2C) and suggests a 1:1 stoichiometry for the subunits, similar to the yeast Set1/COMPASS complex (2).

ES Cells Lacking CFP1 Exhibit Aberrant Histone Modifications—ES cells lacking CFP1 were examined in an attempt to determine the function of this factor within the Set1 complex. Murine ES cells that lack CFP1 (CXXC1−/−) express normal levels of Set1, Ash2, and Rbbp5 (Fig. 6A). Importantly, the Set1 complex derived from these cells migrates at a slightly smaller mobility in a sucrose gradient (Fig. 6B), providing further evidence that CFP1 is an integral component of the Set1 complex and indicating that an intact Set1 complex persists in the absence of CFP1.

CXXC1−/− ES cells were examined for perturbations in the levels of global histone modifications (Fig. 7). The level of the euchromatin histone modification H3K4me2 is increased ∼43% in CXXC1−/− ES cells.
CFP1 Is a Component of the Set1 Complex

Compared with wild type cells (Fig. 7A), introduction of a CFP1 expression vector into CXXC1−/− cells (CXXC1−/−/cDNA), which restores CFP1 expression to ~50% of wild type levels (25), results in H3K4me2 levels decreasing to wild type levels, thus indicating causality with CFP1 deficiency. Conversely, the level of the heterochromatin histone modification H3K9me2 is decreased ~28% in CXXC1−/− ES cells compared with wild type cells, and this defect is also rescued by the introduction of the CFP1 expression vector (CXXC1−/−/cDNA). Thus, ES cells lacking CFP1 exhibit decreased cytosine methylation (25) and histone H3-Lys9 methylation, indicating reduced levels of heterochromatin. These data implicate CFP1 as an important epigenetic regulator of histone modifications and chromatin structure.

We previously demonstrated that dynamic modulation of global histone modification is required for ES cell differentiation (22). Experiments were performed to assess the modulation of global histone modifications in ES cells lacking CFP1. The level of histone H3K4me3 declines slightly in wild type ES cells during the 6 days following removal of leukemia inhibitory factor (LIF) from the grown media and induction of differentiation but increases nearly 4-fold in CXXC1−/− ES cells following removal of LIF (Fig. 7B), consistent with an inhibitory role for CFP1 in the Set1 complex. Similar results were obtained for the levels of the H3K4me1 and H3K4me2 modifications (Fig. 7B). Defects of histone modulation were rescued in the CXXC1−/− ES cells into which the CFP1 cDNA was introduced (CXXC1−/−/cDNA). In contrast, global modulation of histone H3-Lys9 methylation appears normal in CXXC1−/− ES cells following induction of differentiation.

DISCUSSION

Previous work demonstrated that CFP1 is essential for early mammalian embryogenesis, ES cell differentiation, and normal cytosine methylation patterns (21, 25). However, the severity of the phenotype observed in the absence of CFP1 cannot be explained solely by deficient cytosine methylation. Furthermore, a CFP1 function distinct from regulation of cytosine methylation is suggested by the finding that the yeast homologue Spp1 is a component of the Set1/COMPASS histone
The finding of CFP1 as a component of the Set1 histone H3-Lys\(^4\) methyltransferase complex provides insight into how CFP1 intersects with the cellular epigenetic machinery. In addition to the previously reported deficiency in cytosine methylation (25), this study also documents that ES cells lacking CFP1 exhibit alterations in histone modifications, including reduced steady state levels of histone H3K9me2 and elevated steady state levels of histone H3K4me2, and a dramatic 4-fold elevation of histone H3-Lys\(^4\) methylation upon the induction of differentiation. These data indicate that CFP1 functions to restrict the activity of the Set1/CFP1 complex.

It remains to be determined whether CFP1 is a component of additional mammalian H3-Lys\(^4\) methyltransferase complexes, as is the case for other components of the Set1 complex such as Ash2, Rbbp5, and Wdr5 (15, 39–41, 45). We previously reported that human CFP1 co-localizes to an identical set of nuclear speckles with the HRX/MLL histone H3-Lys\(^4\) methyltransferase (20). The composition of the megadalton HRX/MLL histone H3-Lys\(^4\) methyltransferase complex has been reported, but CFP1 was not detected (40, 45). Similarly, a histone methyltransferase complex containing mammalian Set1 has been reported but lacks CFP1 (41). However, these studies reported the composition of soluble complexes. CFP1 is localized nearly exclusively in the nuclear matrix, and hence might not have been recovered by the extraction methods utilized.

The phenotype exhibited by ES cells lacking CFP1 is strikingly similar to that of cells lacking Dicer (46), a protein in the small interfering RNA pathway that is implicated in the initiation of heterochromatin formation via the RITS complex (47). ES cells lacking either Dicer or CFP1 are unable to differentiate \textit{in vitro} and exhibit a dramatic decrease in cytosine methylation and heterochromatin histone modifications. Rather than lacking a factor needed for heterochromatin formation such as Dicer, however, cells lacking CFP1 lack an inhibitory component of a euchromatin-associated histone modifying complex. This convergence of phenotype caused by mutations in diverse epigenetic regulators illustrates the competitive balance that must be achieved between enzyme complexes that facilitate heterochromatin and euchromatin formation.

Cytosine methylation levels can decline as a secondary effect of altered histone modifications. For example, cytosine methylation in \textit{Neurospora crassa} is dependent on methylation of histone H3, and inhibition of histone deacetylase activity results in a loss of cytosine methylation (48–50). Furthermore, the chromatin remodeling protein DDM1 in \textit{Arabidopsis thaliana} and the related factor LSH in mammals are required for normal cytosine methylation (51–53), and disruption of the Suv39H1 histone H3-Lys\(^4\) methyltransferase gene in murine ES cells leads to altered localization of Dnmt3b and decreased cytosine methylation at peri-centric satellite repeats (54). Thus, it remains possible that decreased cytosine methylation in ES cells lacking CFP1 is a secondary effect of altered histone modifications. However, several lines of evidence indicate a more direct involvement of CFP1 with the regulation of cytosine methylation. CFP1 exhibits a binding specificity for DNA sequences containing CpG dinucleotides (1), the major site of cytosine methylation. Furthermore, the magnitude of the cytosine methylation deficiency in ES cells lacking CFP1 (~75% reduction) is greater than the extent of histone code perturbation. These cells exhibit reduced maintenance Dnmt activity and reduced Dnmt1 protein (25), which is the first example of altered Dnmt1 function in the absence of direct abro-
gation of Dnm1 gene expression. Last, CFP1 orthologues in lower eukaryotes that lack cytosine methylation, such as yeast and C. elegans, lack the CXXC DNA-binding domain. This suggests that CFP1 acquired the ability to bind to DNA in organisms that utilize cytosine methylation as an epigenetic regulator of gene expression and chromatin structure. Taken together, these data indicate that CFP1 is a critical epigenetic regulator of both cytosine methylation and histone modifications.

Acknowledgments—We thank Winship Herr for generously providing a rabbit polyclonal Set1 antibody and Dr. Exing Wang of the Indiana Center for Biological Microscopy for expert assistance with confocal microscopy. Anti-MEF50 antibody was a kind gift of Gideon Dreyfuss (University of Pennsylvania School of Medicine).

REFERENCES
1. Voo, K. S., Carlone, D. L., Jacobsen, B. M., Flodin, A., and Skalnik, D. G. (2000) Mol. Cell. Biol. 20, 2108–2121
2. Miller, T., Krogan, N. J., Dover, J., Erdjument-Bromage, H., Tempst, P., Johnston, M., Greenblatt, J. F., and Shilatifard, A. (2001) Proc. Natl. Acad. Sci. U. S. A. 98, 12902–12907
3. Lee, J.-H., Voo, K. S., and Skalnik, D. G. (2001) Mol. Cell. Biol. 21, 3080–3089
4. Hendrich, B., and Bird, A. (1998) J. Biol. Chem. 273, 1714–1725
5. Cross, S. H., Meehan, R. R., Nan, X., and Bird, A. (1997) J. Biol. Chem. 272, 21457–21462
6. Miller, T., Krogan, N. J., Dover, J., Erdjument-Bromage, H., Tempst, P., Johnston, M., Greenblatt, J. F., and Shilatifard, A. (2001) Proc. Natl. Acad. Sci. U. S. A. 98, 12902–12907
7. Lee, J.-H., Voo, K. S., and Skalnik, D. G. (2001) Mol. Cell. Biol. 21, 3080–3089
8. Prasad, R., Yano, T., Sorio, C., Nakamura, T., Rallapalli, R., Gu, Y., Leshkowitz, D., Nakamura, T., Canaani, E., Ma, Q., Alder, H., Nelson, K. K., Chatterjee, D., Gu, Y., Kaname, T., Canaani, O., Cimino, G., Croce, C. M., and Canaani, E. (1992) Cell 71, 701–708
9. Ma, Q., 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
10. Prasad, R., Yano, T., Sorio, C., Nakamura, T., Rallapalli, R., Gu, Y., Leshkowitz, D., Croce, C. M., and Canaani, E. (1995) Proc. Natl. Acad. Sci. U. S. A. 92, 12160–12164
11. Hendrich, B., and Bird, A. (1998) J. Biol. Chem. 273, 1714–1725
12. Cross, S. H., Mehan, R. R., Nan, X., and Bird, A. (1997) Nat. Genet. 16, 256–259
13. Voo, K. S., Carlone, D. L., Jacobsen, B. M., Flodin, A., and Skalnik, D. G. (2000) Mol. Cell. Biol. 20, 2108–2121
14. Miller, T., Krogan, N. J., Dover, J., Erdjument-Bromage, H., Tempst, P., Johnston, M., Greenblatt, J. F., and Shilatifard, A. (2001) Proc. Natl. Acad. Sci. U. S. A. 98, 12902–12907
15. Lee, J.-H., Voo, K. S., and Skalnik, D. G. (2001) Mol. Cell. Biol. 21, 3080–3089
16. Aasland, R., Gibson, T. J., and Stewart, F. (1995) Trends Biochem. Sci. 20, 56–59
17. Fukis, F., Hurd, P. J., Deplus, R., and Kouzarides, T. (2003) Nucleic Acids Res. 31, 2305–2312
18. Goshinski, N., Karumanchi, A. S., Bhatnagar, R., Parikh, P., and Anderson, D. (2003) J. Biol. Chem. 278, 39532–39537
19. Voo, K. S., Carlone, D. L., Jacobsen, B. M., Flodin, A., and Skalnik, D. G. (2000) Mol. Cell. Biol. 20, 2108–2121
20. Lee, J.-H., and Skalnik, D. G. (2002) J. Biol. Chem. 277, 42259–42267
21. Carlone, D. L., and Skalnik, D. G. (2001) Mol. Cell. Biol. 21, 7601–7606
22. Lee, J.-H., and Skalnik, D. G. (2002) J. Biol. Chem. 277, 42259–42267
23. Li, E. (2002) Nat. Rev. Genet. 3, 662–673