Set2 Is a Nucleosomal Histone H3-Selective Methyltransferase That Mediates Transcriptional Repression

Brian D. Strahl,[1]† Patrick A. Grant,[1] Scott D. Briggs,[1] Zu-Wen Sun,[1] James R. Bone,[1]
Jennifer A. Caldwell,[2] Sahana Mollah,[2] Richard G. Cook,[3] Jeffrey Shabanowitz,[2] Donald F. Hunt,[2,4]
and C. David Allis[1]*

*Department of Biochemistry and Molecular Genetics[1] and Department of Pathology,[4] University of Virginia Health System, and
Department of Chemistry, University of Virginia,[2] Charlottesville, Virginia 22908, and Department of Immunology,
Baylor College of Medicine, Houston, Texas 77030[3]*

Recent studies of histone methylation have yielded fundamental new insights pertaining to the role of this modification in gene activation as well as in gene silencing. While a number of methylation sites are known to occur on histones, only limited information exists regarding the relevant enzymes that mediate these methylation events. We thus sought to identify native histone methyltransferase (HMT) activities from *Saccharomyces cerevisiae*. Here, we describe the biochemical purification and characterization of Set2, a novel HMT that is site-specific for lysine 36 (Lys36) of the H3 tail. Using an antiserum directed against Lys36 methylation in H3, we show that Set2, via its SET domain, is responsible for methylation at this site in vivo. Tethering of Set2 to a heterologous promoter reveals that Set2 represses transcription, and part of this repression is mediated through the HMT activity of the SET domain. These results suggest that Set2 and methylation at H3 Lys36 play a role in the repression of gene transcription.

Eukaryotic DNA is complexed in cells by histone proteins to form the fundamental repeating unit of chromatin, the nucleosome. Stretches of nucleosomes are further folded upon themselves to create higher-order chromatin structures that are currently not well defined. Compaction of DNA in this manner imposes a severe impediment to proteins that require access to the DNA template. Clear examples of this impediment have been shown to exist for the machinery that drives DNA transcription (28, 38, 41). However, this same impediment faces all aspects of DNA metabolism, including replication, repair and recombination (18, 40).

Posttranslational modifications of histone amino termini are recognized to play a central role in the control of chromatin structure and function. A diverse array of covalent histone modifications have been documented that take place on the tail domains of histones which protrude away from the nucleosome (9, 39). We and others have proposed that these modifications form a histone code which directly regulates chromatin function either by altering the specific structure of the chromatin polymer itself and/or by recruiting proteins or protein complexes that uniquely recognize a single or combinatorial set of modifications on one or more histone tails (14, 35, 37). For example, recent evidence showing that the bromodomains of various histone acetyltransferases, including PCAF, GCN5 and TAF$_{II}$250, bind to acetylated lysines in the histone tails suggests that specific recruitment of the transcriptional apparatus to promoters is one likely mechanism to explain how histone modifications influence transcription (8, 22). It appears that other histone modifications, including methylation, function in the same manner (see below).

Histone methylation is a posttranslational modification that occurs on lysine and arginine residues in the H3 and H4 tail domains (reviewed in reference 42). In histone H3, lysines 4, 9, 27, and 36 are well-documented sites of methylation, while in histone H4, lysine methylation is restricted to lysine 20 (36, 39). Recent work has identified the Su(var)3-9 family as being H3 lysine 9-specific histone methyltransferases (HMTs). These proteins, which play an important role in the establishment and maintenance of heterochromatin, mediate their methylation activity through the SET domain. Given that this methylation mark is specifically recognized by the chromo domains of HP1 and its *Schizosaccharomyces pombe* homologue Swi6 (2, 7, 12, 19, 25), it is proposed that specific recruitment of these proteins in a lysine 9 (Lys9)-dependent manner mediates the establishment of heterochromatin (13, 14, 32). Outside of heterochromatin-mediated silencing, transcriptional repression by retinoblastoma has recently been shown to require both SUV39H1 and HP1 (26), thus providing a novel mechanism by which the general repression of euchromatic genes may be established (14, 32).

In this study, we sought to identify and characterize histone-specific methyltransferases from *Saccharomyces cerevisiae*. Using a biochemical approach, we identified two separate nucleosome-specific HMT activities, one of which contained the SET family member Set2. The HMT activity of Set2, mediated by its SET domain, is specific for nucleosomal substrates and is site-selective for lysine 36 (Lys36) in histone H3. Generation of an H3 Lys36 antiserum reveals that Set2 is the only SET domain-containing protein in *S. cerevisiae* responsible for mediating methylation at this site. Finally, tethering experiments show that in vivo, Set2 functions as a transcriptional repressor that

* Corresponding author. Mailing address: Department of Biochemistry and Molecular Genetics, University of Virginia Health System, Charlottesville, VA 22908. Phone: (804) 243-6048. Fax: (804) 924-5069. E-mail: allis@virginia.edu.
† Present address: Department of Biochemistry and Biophysics, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599.

mediates its repression partially through the SET domain. These results suggest a role for Set2 and its methylation activity in regulating the repression of gene transcription.

MATERIALS AND METHODS

S. cerevisiae strains. Purification of Set2 was carried out using the *S. cerevisiae* strain CY396, which is described elsewhere (29). The SET domain deletion strains and isogenic wild-type (WT) strain (in the BY4742 background) were obtained from Research Genetics. The *set1Δ* and isogenic wild-type strains are described elsewhere (3, 3a). For in vivo repression assays, the wild-type *S. cerevisiae* strain YMH171, a strain previously employed for repression assays (43), was used.

Protein purification and chromatography. Whole-cell extracts were prepared from 20 liters of the *S. cerevisiae* strain CY396 grown to mid-log phase as described (10). This extract (160 ml total) was bound batchwise with 20 ml of Ni^{2+}-nitrilotriacetic acid (NTA)-agarose (Qiagen), and the resin was washed in a column with 20 mM imidazole-containing extraction buffer (10). Proteins were eluted with 300 mM imidazole and passed through a Mono Q HR 5/5 column (Amersham Pharmacia Biotech). The flowthrough was then collected and loaded directly onto a Mono S HR 5/5 column (Amersham Pharmacia Biotech) before bound proteins were eluted using a 25-ml linear gradient from 0.1 to 1 M NaCl. Three microliters from every other fraction was assayed for HMT activity as described below.

Peak 1 and peak 2 HMT activities were pooled, concentrated to 0.5 ml using a Centriprep-30 (Amicon), and then loaded on a Superose 6 HR 10/30 column (Amersham Pharmacia Biotech) equilibrated with extraction buffer plus 0.5 mM dithiothreitol (DTT). The chromatographic system was calibrated using the following standard proteins: thyroglobulin (669 kDa), apoferritin (440 kDa), aldolase (158 kDa), and ovalbumin (45 kDa). Peak 1 and peak 2, which comigrated between 100 and 200 kDa, were pooled and diluted to 0.1 M NaCl prior to loading on a 0.5-ml heparin-Sepharose column (Sigma). Bound proteins were eluted with a 5 ml of a 0.1 to 1 M NaCl gradient. Peak HMT activity was pooled and then loaded on the Mono S HR 5/5 column as described above. Peak 1 and 2 activities were pooled and loaded on a Mini S PC 3.2/3 column (Amersham Pharmacia Biotech) attached to a SMART System (Amersham Pharmacia Biotech).

To further purify peak 1 for mass spectrometry identification, peak 1 activity from the Mini S column was pooled, concentrated to ≤50 μl as described above, and then loaded onto a Superose 12 PC 3.2/30 column (Amersham Pharmacia Biotech) equilibrated to 300 mM NaCl. Then 5 μl of each fraction (50 μl) was analyzed for HMT activity as described below. The Superose 12 column was calibrated using the following standard proteins: thyroglobulin (669 kDa), apoferritin (440 kDa), β-amylase (200 kDa), alcohol dehydrogenase (150 kDa), albumin (66 kDa), ovalbumin (45 kDa), and carbonic anhydrase (31 kDa). We note that attempts to resolve bacterially produced Set2-Flag (see below) on the Superose 12 column to determine the homo- or heterodimeric nature of this protein have been unsuccessful due to significant levels of aggregation observed under the isolation conditions used to purify Set2 (data not shown).

Cloning and expression of Set2. The full-length Set2 gene was amplified from *S. cerevisiae* genomic DNA by PCR using an oligonucleotide primer that added a *Bam*HI site in front of the ATG start site (5′-CGCGGATCCATGTCGAAG AACCAAAGTGTGAG-3′) and an oligonucleotide primer that added a *Sal*I site that followed a Flag coding sequence and stop site (5′-ACGCGTCACTTACT TGTCATCGTCGTCCTTGTAGTCTGATGATGTTGAAGGTGGAGGAG-3′). Set2-Flag PCR product was spin column purified using the QIAquick PCR purification kit (Qiagen), digested with *Bam*HI and *Sal*I, agarose gel purified, and then cloned into the *Escherichia coli* expression vector pCAL-n (Stratagene). The Set2 coding region was sequenced for accuracy. The resulting expression plasmid was transformed and expressed in the BL21-Gold(DE3) *E. coli* strain (Stratagene).

Cells containing vector only or Set2-Flag (500 ml each) were grown to an optical density at 600 nm (OD_{600}) of 0.8, followed by addition of 0.1 mM isopropyl-ᴅ-thiogalactoside for 3 h at 30°C. Harvested cells were resuspended in 30 ml of lysis buffer (50 mM Tris-HCl [pH 8.0], 150 mM NaCl, 1 mM DTT, 1.0 mM magnesium acetate, 2 mM $CaCl_2$, 0.1% NP-40, 5 mg/ml lysozyme, 1 mM phenylmethylsulfonyl fluoride [PMSF], and 2 μg of leupeptin, aprotinin, and pepstatin per ml each) and sonicated. Cell debris was centrifuged at $14,000 \times g$, and the supernatant was incubated batchwise with 200 μl of the calmodulin affinity resin (Stratagene). Bound proteins were eluted in a column according to the manufacturer's instructions except that the elution was done with stepwise increases in NaCl (150 mM, 300 mM, and 500 mM; 1 ml at each elution). Protein

concentration was determined by Coomassie staining of sodium dodecyl sulfate (SDS)-polyacrylamide gels using bovine serum albumin (BSA) as the standard.

Mutations in the Set domain of Set2 (at R195 and C201) were made by standard PCR mutagenesis methods using the Set2-Flag plasmid. After sequencing through the N terminus and SET domain for accuracy, a *Bam*HI-*Nsi*I fragment encompassing the N terminus and SET domain of Set2 was recloned into the Set2-Flag construct, generating the final mutant expression plasmids. Expression and purification of these proteins were performed as described above. For *S. cerevisiae* expression, the full-length Set2-Flag and Set2 SET domain mutant constructs described for bacterial expression were digested with *Bam*HI and *Sal*I, and the purified full-length DNAs were subcloned into the p416 alcohol dehydrogenase (ADH) expression plasmid that is driven by the *ADH1* promoter (24).

Methyltransferase activity assays. For HMT assays, partially purified *S. cerevisiae* extracts or 0.2 μg of recombinantly expressed Set2 or SET domain mutants of Set2 were incubated with either 2 μg of chicken core histones or mononucleosomes (23) along with 0.55 μCi of *S*-adenosyl-ʟ-[*methyl*-³H]methionine (³H-AdoMet; 72 Ci/mmol; NEN Life Science Products) in methyltransferase buffer (final concentrations being 50 mM Tris [pH 9.0], 1 mM PMSF, and 0.5 mM DTT) for 30 min at 30°C in a total volume of 10 μl. Then 2 μl of the reaction was spotted on Whatman P-81 for liquid scintillation counting to monitor reactions, while the remainder was analyzed by SDS-polyacrylamide gel electrophoresis (PAGE) followed by Coomassie staining and fluorography.

Set2 site mapping and protein microsequencing. For labeling studies involving recombinant Set2 for microsequencing, the above HMT reaction volume, using chicken nucleosomes as the substrate, was scaled up 20-fold. Histones from this reaction were acid extracted with 0.4 N H_2SO_4 and precipitated with trichloroacetic acid (TCA), and the H3 was purified by reverse-phase high-pressure liquid chromatography (RP-HPLC) using a C_8 column. ³H-methylated H3 (≈10 μg) was digested with 0.05 μg of V8 protease (Sigma) in 50 mM NH_4HCO_3 in a 10-μl volume for 3 h at 37°C; digestion products were then separated by RP-HPLC. The H3 1-50 fragment, identified by mass spectrometry as being the only V8 product strongly methylated by Set2, was further digested with 0.5 U of Arg-C protease (Roche Molecular Biochemicals) in a 10-μl reaction volume for 3 h at 37°C; reaction products were again separated by RP-HPLC. The ³H-methylated H3 Arg-C fragment was sequenced in an Applied Biosystems model 477A protein sequencer with an in-line 120A phenylthiohydantoin (PTH) analyzer (Applied Biosystems) using optimized cycles. After conversion, 50% of the sample was transferred to the RP-HPLC for PTH-amino acid identification, and the other 50% was collected for determination of radioactivity by scintillation counting.

Electrophoresis and Western blotting. SDS-PAGE and Western blot analyses were performed using procedures and reagents from Amersham Pharmacia Biotech. Antiserum detection was made using the ECL+plus chemiluminescent kit (Amersham Pharmacia Biotech). Rabbit anti-LexA (Upstate Biotechnology Inc.) was used at a dilution of 1:3,000, and mouse monoclonal anti-Flag antiserum (M2; Sigma) was used at 1 μg/ml. The anti-H3(Lys36)Me antiserum was routinely used at a dilution of 1:10,000.

Development of an antiserum specific to Lys36 methylation of histone H3. A synthetic peptide coding for sequence 30 to 38 of the human H3.1 amino terminus (NH₂-C*PATGGVKKP), in which Lys36 (underlined) was made dimethyllysine (Bachem), was conjugated by standard protocols to keyhole limpet hemocyanin via an N-terminal artificial cysteine (C*) prior to rabbit immunization. Enzyme-linked immunosorbent assays (ELISAs) were performed as previously described (4) using various amounts of unmodified or Lys36-methylated H3 30-38 peptides (data not shown).

Mass spectrometry. For the identification of proteins from gel slices, fractions from the Superose 12 size exclusion column containing peak 1 HMT activity were resolved on a SDS–10% PAGE gel followed by silver staining as described (30). Candidate gel bands were excised and digested with trypsin as described (33). Peptides were then extracted from the gel, and an aliquot of the resultant peptides was loaded on a microcapillary column packed with 5-μm C_{18} particles (YMC ODS-AQ; Waters). The peptides were gradient eluted into an LCQ DECA mass spectrometer (ThermoFinnigan) for data-dependent mass spectral analysis. The HPLC gradient was 0 to 60% B in 70 min, 60 to 100% B in 15 min. Solvents A and B were 0.1 M acetic acid in water and 0.1 M acetic acid in 70% acetonitrile, respectively. Peptide sequences identified from this analysis were used to search the *S. cerevisiae* database to identify respective candidate proteins.

Transcriptional repression assays. Full-length Set2 or the Set2 SET domain mutants were amplified by PCR and cloned into a C-terminal LexA DNA-binding domain fusion construct driven by the *ADH1* promoter (gift from Mike Hampsey). All constructs were confirmed by sequencing and shown to express full-length products when transformed into *S. cerevisiae* (see Fig. 6B). Repression assays were performed as previously described (43). In brief, the Set2-LexA
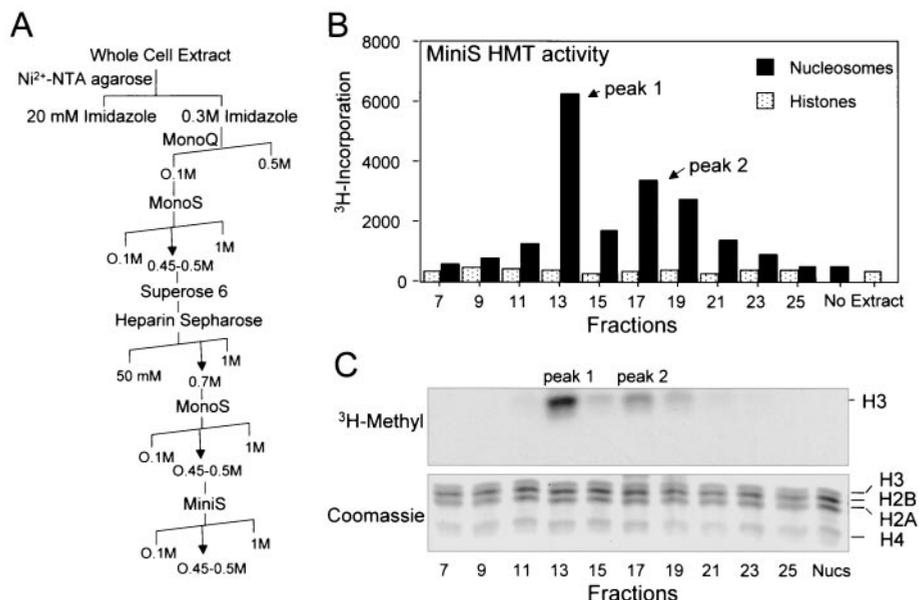
FIG. 1. Partial purification of two nucleosomal H3 HMT activities from *S. cerevisiae*. (A) Schematic representation of the chromatographic procedure used to partially purify two nucleosome-specific HMT activities from *S. cerevisiae* whole-cell extracts (see text for details). (B) Partially purified extracts fractionated by a Mini S column were assayed for HMT activity using chicken core histones or nucleosomes as the substrate. Incorporation of ³H-AdoMet was measured using the filter-binding assay, and identified peaks of activity are labeled. (C) A portion of the HMT assays from the above Mini S reactions involving nucleosomes were resolved on an SDS–15% PAGE gel and examined by Coomassie staining (lower) and fluorography (upper).

constructs, along with an Rpd3-LexA expression construct, were individually transformed into the wild-type *S. cerevisiae* strain YMH171 along with the *lacZ* reporter construct (pCK30) that contained one LexA operator site upstream of the *CYC1* promoter (17). β-Galactosidase assays were performed as described (1), and the values are reported as the fold repression using the vector-only (LexA DNA-binding domain only) as the comparison. All values represent the average for at least three independent transformants repeated in duplicate. Statistical differences between the level of repression of Set2-LexA, Set2(R195G)-LexA, and Set2(C201A)-LexA were determined using one-way analysis of variance followed by Tukey's post hoc test for individual comparisons (performed using Prism; Graphpad Software Inc.).

RESULTS

Purification of native HMT activities from *S. cerevisiae*. A number of studies have documented the existence of both lysine and arginine methylation on histones in vivo (42). However, the functional significance of histone methylation has largely remained obscure, due, in part, to the lack of knowledge regarding the identity of the enzymes that mediate histone methylation. To identify native *S. cerevisiae* HMTs, whole-cell extracts, prepared from asynchronously growing *S. cerevisiae* cells, were fractionated over a series of chromatographic steps (Fig. 1A) that have proven useful in the identification of native histone acetyltransferases (10). Column fractions were then analyzed for histone and nucleosomal HMT activity using chicken core histones or mononucleosomes as the substrate.

Preliminary surveys of the Mono S and Mono Q fractions from the unbound nickel-agarose material revealed no significant HMT activities (data not shown). In contrast, analysis of the 300 mM imidazole-eluted material from the nickel-agarose resin revealed the presence of two peaks of HMT activity specific to the Mono S column that were nucleosome and H3

selective (see Fig. 1C and 3A for examples). No significant HMT activity was observed from nickel-agarose-bound Mono Q fractions (data not shown).

To further characterize these HMT activities, both peaks of activity from the first Mono S fractionation were further fractionated by Superose 6 and heparin-Sepharose chromatography. Results showed that both peaks were resolved as a single peak of HMT activity on both columns, and size estimation on the Superose 6 column showed these activities to have an apparent native molecular mass of ≈100 to 200 kDa (data not shown). Subsequent purification of these activities through additional cation exchange columns restored the separation of the two individual peaks of nucleosomal H3 HMT activity (Fig. 1B and C; hereafter referred to as peak 1 and peak 2).

Given the greater abundance of HMT activity observed with peak 1 than with peak 2, we attempted to identify the polypeptide responsible for catalytic activity from HMT peak 1. Fractions 12 to 16 (corresponding to peak 1) from the Mini S column were pooled and then subjected to chromatography on a Superose 12 size exclusion column. Results from this separation revealed peak 1 to be approximately 175 kDa (Fig. 2A), with the majority of HMT activity restricted to fractions 9 and 10 (Fig. 2B). Comparison of the protein profiles revealed several candidate polypeptides that cofractionated closely with HMT activity. These bands were excised and digested with trypsin, and the resulting peptides were examined by nano-HPLC microelectrospray ionization tandem mass spectrometry. Of the proteins tentatively identified from this analysis, one of these was identified as Set2 (Fig. 2C).

Set2 is required for peak 1 HMT activity. To further determine if Set2 is necessary for the establishment of peak 1 and/or
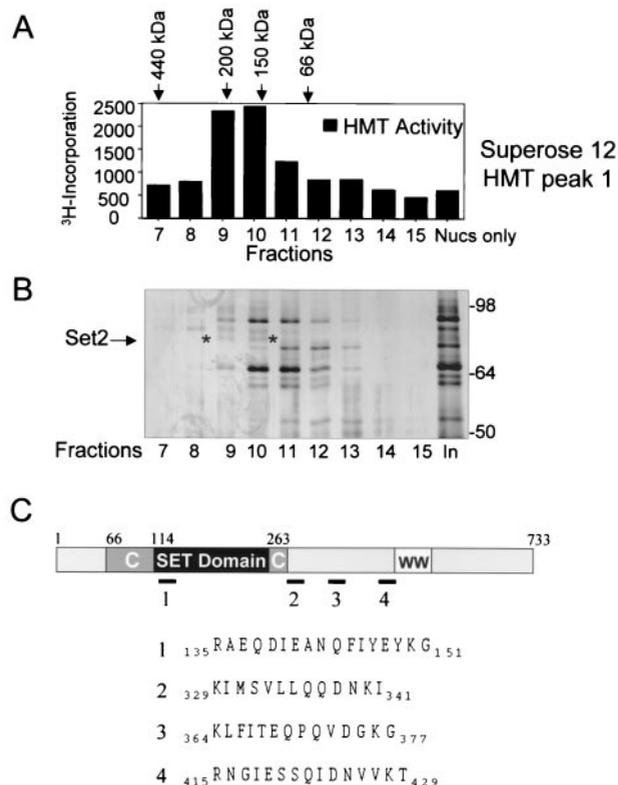
FIG. 2. Set2 is identified as a component of the first peak of nucleosomal HMT activity. (A) Peak 1 HMT activity from the Mini S column was fractionated using a Superose 12 size exclusion column, and collected fractions were analyzed for nucleosome-specific HMT activity by the filter-binding assay. Size estimation of peak 1 activity was determined to be ≈175 kDa; calibration sizes are shown. (B) Silver stain gel of the Superose 12 fractions containing HMT peak 1 activity. Candidate bands which correlated well with the peak of HMT activity were excised, and the proteins were identified by mass spectrometry. Asterisks surround the candidate band that was identified as Set2. (C) A schematic representation of the domain structure of Set2. The SET domain and flanking cysteine-rich regions (C) are shown. These flanking Cys-rich regions are also denoted as the pre- and post-SET domains, and we note that the pre-SET domain of Set2 is noncanonical (see reference 11 for alignment). A putative proline-rich-binding motif (WW motif) is also shown. Peptide fragments identified by mass spectrometry and their relative locations in Set2 are shown.
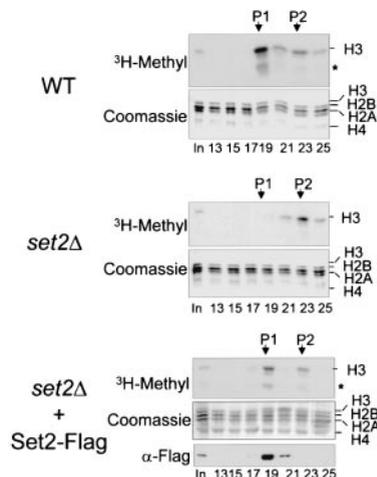

FIG. 3. Peak 1 HMT activity in *S. cerevisiae* is dependent on Set2. Whole-cell extracts prepared from *S. cerevisiae* strains which were either wild type, *set2Δ*, or *set2Δ* containing a Set2-Flag expression construct were bound and eluted from nickel-agarose, passed through a Mono Q column, and then fractionated on a Mono S column as outlined in Fig. 1. Fractions were assayed for nucleosome-specific HMT activity, and reaction products were resolved on SDS–15% PAGE gels followed by Coomassie staining (lower) and fluorography (upper). Western blot analysis of the expression and location of Set2-Flag on the Mono S column is shown. Asterisks indicate partial H3 breakdown products that are typically observed; P1 and P2 indicate peak 1 and peak 2, respectively.

peak 2 HMT activity, whole-cell extracts were prepared and fractionated as described above except from a yeast strain in which *SET2* was disrupted (*set2Δ*). In parallel, an isogenic WT strain was examined. Both extracts were then analyzed for the presence of the two nucleosome-specific HMT peaks of activity. As shown in the top panels of Fig. 3, fractions prepared from the WT strain showed the presence of both peaks of nucleosomal HMT activity, as expected. In contrast, extracts prepared from *set2Δ* lacked peak 1 but not peak 2 HMT activity, suggesting that Set2 is responsible for peak 1 HMT activity (Fig. 3, middle panels).

To demonstrate the ability of Set 2 to restore peak 1 HMT activity, an expression plasmid in which a Flag epitope was fused to the C terminus of Set2 was constructed and introduced into the *set2Δ* strain. Results clearly showed that expression of Set2 restores the HMT activity of peak 1 (Fig. 3, lower panels). In agreement, Set2 comigrated precisely in the Mono S fractions that corresponded to peak 1 but not peak 2 HMT activity (Fig. 3, lower panels). These results demonstrate not only that Set2 is required for the HMT activity of peak 1, but also that Set2 is not responsible for peak 2 activity.

Set2 is a nucleosomal HMT highly selective for H3. To determine if Set2 has intrinsic HMT activity, recombinant Set2 containing an N-terminal calmodulin-binding protein (CBP) and a C-terminal Flag tag was expressed in bacteria. In addition, similar constructs were generated in which two independent amino acid substitutions were generated in the SET domain (R195G and C201A; see Fig. 4A). R195 and C201 are highly conserved residues in the SET domain of Set2, and recently, homologous residues were shown to be important for the HMT activity of SUV39H1 (31). Equivalent amounts of recombinant proteins were purified using calmodulin resin (determined by Coomassie staining and by Western blot analysis, Fig. 4B) and used in standard HMT assays. In agreement with our earlier findings (see Fig. 3), WT Set2 protein contained strong nucleosome-specific HMT activity specific for H3 (Fig. 4C and D). In contrast, both SET domain point mutants, R195G and C201A, showed dramatically reduced HMT activity towards nucleosomes (Fig. 4C). In addition, no significant HMT activity was observed by Set2 on recombinant H3 in either the presence or absence of sheared calf thymus DNA or histone octamers, indicating that a DNA-bound nucleosomal structure is required for Set2 HMT activity (data not shown).

Set2 is responsible for mediating H3 lysine 36 methylation in vivo. We next sought to determine the target site(s) of Set2 in H3 both in vitro and in vivo. Unexpectedly, preliminary
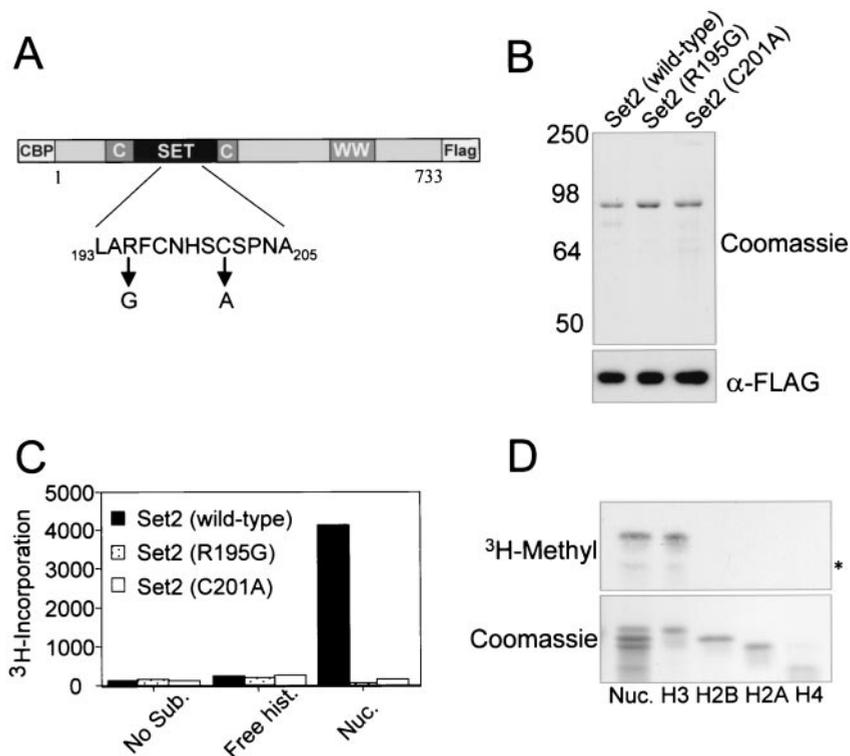
FIG. 4. Set2 methylation activity is dependent upon the SET domain. (A) Schematic representation of the domain structure of the recombinant Set2 protein expressed in bacteria. Substitutions made of highly conserved amino acids in the SET domain are shown. (B) Coomassie and Western blot analysis of affinity-purified recombinant Set2-Flag and Set2-Flag SET domain mutants. (C) Affinity-purified recombinant Set2-Flag and Set2-Flag SET domain mutants were incubated in the presence or absence of either chicken core histones (Free hist.) or nucleosomes (Nuc.) along with ^{3}H-AdoMet. Reaction products were analyzed by the filter-binding assay. (D) RP-HPLC-purified histones isolated from HMT reactions involving recombinant Set2-Flag and nucleosomes were resolved on an SDS–15% PAGE gel and examined by Coomassie staining (lower panel) and fluorography (upper panel). Asterisks indicate partial H3 breakdown products that were observed.

microsequencing analyses, using H3 from reactions involving chicken nucleosomes and recombinant Set2, revealed no evidence for Set2-mediated methylation in the extreme 1-27 amino terminus of H3 (data not shown). We reasoned that methylation by Set2 might occur elsewhere in the histone tail domain or in the C terminus of H3. To that end, a combination of endoproteinases, Glu-C and Arg-C, were employed to identify H3 peptide fragments that were methylated preferentially by Set2 under our assay conditions. From these stepwise digestions (see Materials and Methods for details), one predominant methylated fragment was isolated and determined by mass spectrometry to be amino acid residues 27 to 40 (data not shown). As shown in Fig. 5A, microsequence analysis of the radiolabeled Arg-C H3 peptide showed clearly that Lys36 is the exclusive site of Set2 methylation under these in vitro conditions.

To determine if Set2 methylation is responsible for H3 Lys36 methylation in vivo, and to determine if any other SET domain-containing proteins in *S. cerevisiae* mediate methylation at this site, a novel H3 Lys36 methyl-specific antiserum was generated, α-Me(Lys36)H3, by standard procedures (see Materials and Methods). Using ELISAs, we found the resulting antiserum to be highly selective for the Lys36 methyl-containing H3 peptide, but not for a corresponding unmethylated peptide or H3 peptides methylated at lysine 4 or 9 (data not shown). To

further characterize the specificity of the α-Me(Lys36)H3 antiserum, chicken mononucleosomes were incubated with or without recombinant Set2 in the presence of AdoMet, and the products were analyzed by immunoblotting.

As shown in Fig. 5B, nucleosomal H3 from reactions including Set2 was strongly recognized by α-Me(Lys36)H3 compared to control (minus Set2) reactions. Moreover, recognition of nucleosomal H3 by the α-Me(Lys36)H3 antiserum in Set2 reactions was completely abolished by preincubation of the antiserum with the Lys36 methyl-containing peptide (Fig. 5B). In contrast, preincubation of α-Me(Lys36)H3 with an unmethylated H3 peptide or H3 peptides methylated at lysine 4 or 9 did not prevent the recognition of Set2-dependent Lys36 methylation, demonstrating that this antiserum is selective for H3 Lys36 methylation. In support, the α-Me(Lys36)H3 antiserum did not recognize recombinant *Xenopus* H3 (Fig. 5B). Weak recognition of chicken H3 by the α-Me(Lys36)H3 antiserum in control reactions suggests a low level of Lys36 methylation in chickens, a result confirmed in *S. cerevisiae*, *Tetrahymena*, and human 293T cells (data not shown).

Next, we sought to identify the enzyme that mediates Lys36 methylation in *S. cerevisiae*. Using the α-Me(Lys36)H3 antiserum, we screened *S. cerevisiae* whole-cell lysates prepared from strains with deletions of each of the SET domain-containing genes in *S. cerevisiae*. As shown in Fig. 5C, deletion of
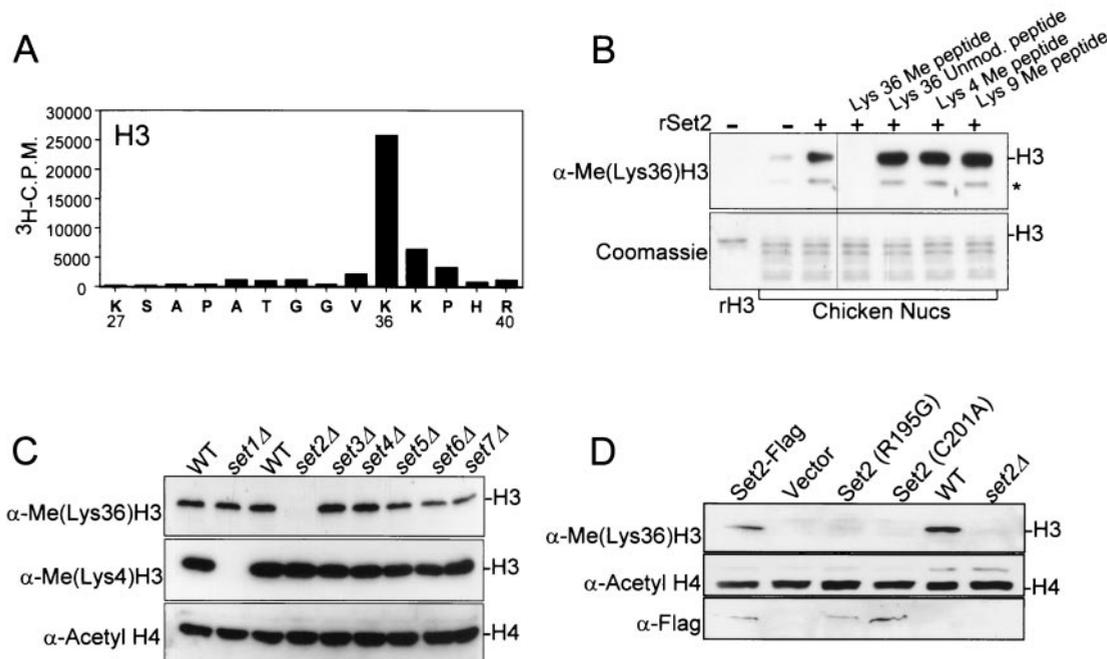
FIG. 5. Methylation of H3 at lysine 36 is mediated by Set2. (A) RP-HPLC-purified H3 isolated from Set2 reactions involving ^{3}H-AdoMet and chicken nucleosomes was digested with endoproteinase Glu-C, and the resulting peptide fragments were separated by RP-HPLC. A labeled H3 fragment, determined to be amino acids 1 to 50 by mass spectrometry (data not shown), was further digested with the endoproteinase Arg-C followed by RP-HPLC fractionation. From this digestion, a ^{3}H-methylated peptide fragment identified as being amino acids 27 to 38 in H3 by mass spectrometry (data not shown) was subjected to N-terminal automated sequencing, and ^{3}H radioactivity eluted from each cycle was counted. Amino acids identified at each cycle of microsequencing are listed; numbers correspond to the known positions of H3 residues. (B) Recombinant H3 or chicken nucleosomes reacted in the presence or absence of purified recombinant Set2 along with AdoMet were immunoblotted with the α-Me(Lys36)H3 antiserum (upper panel). Identical reactions were performed and used in peptide competition assays in which the α-Me(Lys36)H3 antiserum was incubated for 2 h at 25°C with the various peptides shown (peptides were at 10 μg/ml). Parallel reactions were performed and examined by Coomassie staining to monitor loading (lower panel). Asterisk indicates the partial H3 breakdown product observed. (C) Set2 mediates H3 Lys36 methylation in vivo. Whole-cell extracts from wild-type *S. cerevisiae* strains or strains with a deletion of the SET domain were probed with either the α-Me(Lys36)H3 antiserum, the α-Me(Lys4)H3 antiserum, or the α-acetyl-H4 antiserum (used for a loading control). Details of the *S. cerevisiae* strains used are as follows (from left to right as shown in the panel): WT (MBY1198), *set1Δ* (*YHR119w*), WT (BY4742), *set2Δ* (*YJL168c*), *set3Δ* (*YKR029c*), *set4Δ* (*YJL105w*), *set5Δ* (*YHR207c*), *set6Δ* (*YPL165c*), and *set7Δ* (*YDR257c*). MBY1198 and BY4742 are the isogenic wild-type strains for the *set1Δ* and *set2Δ* to *set7Δ* strains, respectively. (D) The SET domain of Set2 is essential for mediating H3 Lys36 methylation in vivo. Whole-cell extracts from wild-type, *set2Δ*, and *set2Δ* strains with either the vector control plasmid (Vector), wild-type Set2 expression plasmid (Set2-Flag), or either of the SET domain mutants [Set2 (R195G) and Set2 (201A)] were probed with the α-Me(Lys36)H3 antiserum. For a loading control, these extracts were probed with the α-acetyl-H4-specific antiserum. Expression of wild-type or mutant Set2 proteins was monitored using the anti-Flag monoclonal antiserum.

SET2 but not deletion of any other known SET domain-containing gene resulted in complete abolishment of Lys36 methylation, indicating that Set2 is likely to be the sole enzyme responsible for methylation at this site in budding yeast.

For comparison and as a control, a methyl-lysine 4-specific antiserum, α-Me(Lys4)H3, was also analyzed with this set of deletions. As previously observed (3, 3a), only deletion of *SET1* results in abolishment of Lys4 methylation, demonstrating that this modification is mediated specifically by Set1 (Fig. 5C). Taken together, these data suggest that Lys4 and Lys36 H3 methylation is catalyzed by Set1 and Set2, respectively. The putative site and substrate specificity of the other SET domain-containing proteins in *S. cerevisiae* (Set3 to Set7) are currently unknown.

To determine if the loss of H3 Lys36 methylation in the *SET2* deletion strain could be restored by catalytically active forms of Set2, the Set2 expression construct previously used to rescue peak 1 HMT activity (see Fig. 3) was assayed for the presence of Lys36 methylation. As shown in Fig. 5D, expres-

sion of Set2 but not the vector-only control or the two catalytically dead point mutations in Set2 (R195G and C201A; see Fig. 4A) restored H3 Lys36 methylation in this assay. These data strongly suggest that Set2 is responsible for bringing about Lys36 H3 methylation in vivo in *S. cerevisiae*.

Set2 functions as a transcriptional repressor in *S. cerevisiae*. To our knowledge, little information is known about the biological role of Set2, although one published abstract suggests that Set2 may function as a transcriptional repressor (L. Lutfiyya, T. Hesman, and M. Johnston, abstract, Yeast **11:**S220, 1995). To further investigate this possibility, an expression plasmid containing the DNA-binding domain of LexA fused to the C terminus of Set2 was constructed and transformed into *S. cerevisiae* along with a *lacZ* reporter construct driven by a heterologous *CYC1* promoter containing a single copy of the LexA operator (*lexA$_{op}$*). As a positive control for transcriptional repression, we also analyzed an Rpd3-LexA expression construct under identical assay conditions (43).
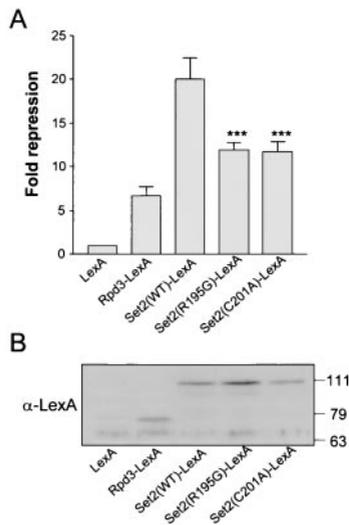
FIG. 6. SET domain point mutants in Set2 are partially defective in transcriptional repression when tethered to a heterologous promoter. (A) The wild-type *S. cerevisiae* strain YMH171 was transformed with either LexA vector-only, Rpd3-LexA, Set2(WT)-LexA, Set2(R195G)-LexA, or Set2(C201A)-LexA along with the *CYC1-lexA$_{op}$-lacZ* reporter plasmid (pCK30), and β-galactosidase assays were performed. Three independent transformants of each strain were assayed in duplicate, and the results are reported as the fold repression relative to the activity of the LexA vector alone. Asterisks indicate that the repression mediated by the two SET domain mutants of Set2 was significantly less ($P \leq 0.05$) than that mediated by wild-type Set2-LexA. (B) To monitor expression of the LexA constructs, extracts from the above experiment were immunoblotted with an anti-LexA polyclonal antiserum. Expression of the vector-only LexA DNA-binding domain (which is 24.5 kDa) was omitted from the figure.

As shown in Fig. 6A, tethering of Set2-LexA showed a dramatic 20-fold repression of transcription, indicating that Set2 can act as a potent transcriptional repressor under these assay conditions. In comparison, Rpd3-LexA induced a 6.6-fold reduction in basal *lacZ* gene expression compared to the LexA DNA-binding domain alone. This result is comparable to previously published reports on the repressive effects of Rpd3-LexA using a similar *lexA$_{op}$-lacZ* reporter construct (16). We note that in parallel experiments using an independent strain containing an integrated MFA2 promoter with a single copy of the *lexA* operator, similar levels of repression were observed by Set2-LexA (data not shown). To investigate the possibility that Set2 may activate transcription, we employed a strain that contained a *GAL1* promoter containing a *lexA* operator fused to *lacZ*. Results showed that Set2-LexA, as well as Rpd3-LexA, exhibited no ability to activate transcription, while LexA-VP16 efficiently activated transcription (data not shown).

To investigate whether the point mutations previously engineered in the SET domain that abolish catalytic activity in vitro and in vivo (see Fig. 4) affect the observed repression in our tethering experiments, Set2-LexA fusion constructs containing mutations at R195 and C201 [Set2(R195G)-LexA and Set2(C201A)-LexA] were generated and analyzed as described above. Results show a modest but significant ($P < 0.05$) reduction in the level of repression observed by Set2 when the SET domain mutants of Set2 were analyzed (both were reduced to ≈12-fold), suggesting that the HMT activity of Set2

plays a partial role in mediating transcriptional repression in this context (Fig. 6A). Using an anti-LexA polyclonal antiserum, all LexA-containing proteins were found to be expressed at similar levels (Fig. 6B). These results indicate that Set2 can act as a transcriptional repressor in this setting that functions, at least in part, through the catalytic activity of its SET domain (i.e., methylation of Lys36 in H3).

DISCUSSION

In this report, we have demonstrated that the *S. cerevisiae* protein Set2 is a nucleosomal H3 methyltransferase that is highly selective for Lys36, a lysine embedded in the more structured region of the H3 tail domain (21). We found that this site of methylation was highly conserved throughout eukaryotes in vivo and showed that Set2, via its SET domain, was responsible for Lys36 methylation in vitro and in vivo. In addition, we provide evidence that Set2 can function as a transcriptional repressor when tethered to a heterologous promoter in *S. cerevisiae*, a result consistent with its previous identification as a repressor for transcription (Lutfiyya et al., abstract). The observed repression appears to be partially mediated through the catalytic activity of the SET domain, indicating a role for Lys36 H3 methylation in downregulating gene expression, at least in this context.

Set2 and other SET domain-containing proteins in *S. cerevisiae*. In *S. cerevisiae*, there are seven unique SET domain-containing proteins, the majority of which are not functionally characterized (34). Of these, Set1 and Set2 were the first identified and characterized as containing SET domains (20, 27; Lutfiyya et al., abstract). In our biochemical survey of *S. cerevisiae* whole-cell extracts, we identified two nucleosomal H3 HMT activities, one of which was shown to be Set2. As well, Set1 has recently been shown to be responsible for H3(Lys4) methylation in budding yeast (3; this report). Initially, we hypothesized that the minor HMT peak 2 (see Fig. 1) would be catalyzed by Set1. However, deletion of Set1 does not result in a loss of either peak 1 or peak 2 HMT activity (data not shown), indicating that peak 2 represents a yet unidentified H3-specific, nucleosomal HMT. The site and substrate specificity of the remaining SET domain-containing proteins in *S. cerevisiae* are yet to be determined.

Set2 as a repressor of transcription. Recent studies have revealed histone methylation as being an important regulator of chromatin function, and to date, these studies document histone methylation as playing both positive and negative roles in the process of gene transcription (14). The initial identification of Set2 was made on the basis of identifying proteins in a genetic screen that, when mutated, relieved basal repression of the *GAL4* promoter (Lutfiyya et al., abstract). In agreement, our work shows that Set2 is a potent repressor of transcription when targeted to a heterologous promoter. While the observed repression is partially mediated by the activity of the SET domain, a significant amount of repression remains in our catalytically dead mutants.

Several possibilities may account for this discrepancy. First, since our tethering experiments were performed in WT strains, the remaining repression observed in the presence of the catalytically dead Set2-LexA mutants may be a result of endogenous Set2 activity. However, tethering experiments with these

LexA constructs in a *set2Δ* strain revealed similar results (data not shown). Alternatively, this remaining repression might be mediated by Set2's ability to recruit additional proteins that aid in the repression process, a result that has been observed for other corepressors such as Rpd3 (5, 6). Similar to what we have observed with catalytically defective Set2 mutants, inactive forms of Rpd3 can also mediate transcriptional repression, and these defective Rpd3 proteins still recruit additional repressive factors (16). Recruitment of additional factors by Set2 might occur through its SET domain, as this domain has been shown to be a protein-protein interaction motif (15). Set2 also contains a WW (proline-rich) binding motif (see Fig. 2C). Whether this motif plays any role in the ability of Set2 to bind other proteins is not known.

In consideration of Set2 as a regulator of transcription, how Set2 is targeted remains an intriguing question. Size estimation of the Set2-mediated peak 1 activity indicates a native molecular mass of approximately 175 kDa (Fig. 2A), suggesting that Set2 may function as either a homodimer or in association with another, yet unidentified factor(s). Given that Set2 shows no evidence for any kind of DNA-binding motif, it is likely that Set2 functions as a corepressor that is targeted through direct protein-protein interaction with other factors. In support, we find no evidence for the ability of Set2-Flag to coimmunoprecipitate Set2-LexA from whole-cell extracts from *S. cerevisiae* cells coexpressing Flag-tagged Set2 (see Fig. 3) and Set2-LexA (see Fig. 6) (data not shown).

Exclusive role for H3 lysine 36 methylation in gene silencing? While sites of methylation have long been identified, only recently has the functional significance of these modifications begun to be elucidated. Based on our tethering experiments (Fig. 6), we suggest that Lys36 methylation, mediated by Set2, plays a role in mediating gene repression, at least under some circumstances. However, we do not rule out the possibility that this modification also may play a positive role for transcription as well. In support, H3(Lys36) methylation, like H3(Lys4) methylation (36), occurs in transcriptionally active macronuclei in *Tetrahymena* (data not shown), suggesting that this mark may function to potentiate transcription. Recent data on H3(Lys4) methylation suggest that this modification functions as part of a long-range, epigenetic marking system that serves to remodel large chromatin domains of chromatin into a state more permissive for gene transcription (reviewed in reference 42). Interestingly, the same H3(Lys4) methyl mark also plays a positive role in ribosomal DNA gene silencing (3, 3a). Thus, it seems likely that each specialized chromatin environment may influence gene regulation through chromatin modifications in different ways. The extent to which Lys36 H3 methylation is a targeted or more global phenomenon, influencing gene expression in a positive and/or negative fashion, remains an important issue for future investigations.

REFERENCES

1. **Adams, A., D. E. Gottschling, C. A. Kaiser, and T. Stearns.** 1997. Methods in yeast genetics, p. 123–128. Cold Spring Harbor Laboratory Press, Plainview, N.Y.
2. **Bannister, A. J., P. Zegerman, J. F. Partridge, E. A. Miska, J. O. Thomas, R. C. Allshire, and T. Kouzarides.** 2001. Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. Nature **410:**120–124.
3. **Briggs, S. D., M. Bryk, B. Strahl, W. Cheung, J. Davie, S. Dent, F. Winston, and C. Allis.** 2001. Histone H3 lysine 4 methylation is mediated by Set1 and required for cell growth and rDNA silencing in *Saccharomyces cerevisiae*. Genes Dev. **15:**3286–3295.
3a. **Bryk, M., S. D. Briggs, B. D. Strahl, M. J. Curcio, C. D. Allis, and F. Winston.** Set1, a factor required for methylation of histone H3, regulates rDNA silencing in *Saccharomyces cerevisiae* by a Sir2-independent mechanism. Curr. Biol., in press.
4. **Cheung, P., K. G. Tanner, W. L. Cheung, P. Sassone-Corsi, J. M. Denu, and C. D. Allis.** 2000. Synergistic coupling of histone H3 phosphorylation and acetylation in response to epidermal growth factor stimulation. Mol. Cell **5:**905–915.
5. **Collingwood, T. N., F. D. Urnov, and A. P. Wolffe.** 1999. Nuclear receptors: coactivators, corepressors and chromatin remodeling in the control of transcription. J. Mol. Endocrinol. **23:**255–275.
6. **Davie, J. R., and V. A. Spencer.** 1999. Control of histone modifications. J. Cell. Biochem. **32-33:**141–148.
7. **Eissenberg, J. C., and S. C. Elgin.** 2000. The HP1 protein family: getting a grip on chromatin. Curr. Opin. Genet. Dev. **10:**204–210.
8. **Filetici, P., and P. Ballario.** 2001. The bromodomain: a chromatin browser? Front. Biosci. **6:**D866–D876.
9. **Grant, P. A.** 2001. A tale of histone modifications. Genome Biol. **2:**1–6.
10. **Grant, P. A., S. L. Berger, and J. L. Workman.** 1999. Identification and analysis of native nucleosomal histone acetyltransferase complexes. Methods Mol. Biol. **119:**311–317.
11. **Huang, N., E. vom Baur, J. M. Garnier, T. Lerouge, J. L. Vonesch, Y. Lutz, P. Chambon, and R. Losson.** 1998. Two distinct nuclear receptor interaction domains in NSD1, a novel SET protein that exhibits characteristics of both corepressors and coactivators. EMBO J. **17:**3398–3412.
12. **Jacobs, S. A., S. Taverna, Y. Zhang, S. Briggs, J. Li, J. C. Eissenberg, C. Allis, and S. Khorasanizadeh.** 2001. Specificity of the HP1 chromo domain for the methylated N terminus of histone H3. EMBO J. **20:**5232–5241.
13. **Jenuwein, T.** 2001. Re-SET-ting heterochromatin by histone methyltransferases. Trends Cell Biol. **11:**266–273.
14. **Jenuwein, T., and C. D. Allis.** 2001. Translating the histone code. Science **293:**1074–1080.
15. **Jenuwein, T., G. Laible, R. Dorn, and G. Reuter.** 1998. SET domain proteins modulate chromatin domains in eu- and heterochromatin. Cell Mol. Life Sci. **54:**80–93.
16. **Kadosh, D., and K. Struhl.** 1998. Histone deacetylase activity of Rpd3 is important for transcriptional repression in vivo. Genes Dev. **12:**797–805.
17. **Keleher, C. A., M. J. Redd, J. Schultz, M. Carlson, and A. D. Johnson.** 1992. Ssn6-Tup1 is a general repressor of transcription in yeast. Cell **68:**709–719.
18. **Kornberg, R. D., and Y. Lorch.** 1999. Twenty-five years of the nucleosome, fundamental particle of the eukaryote chromosome. Cell **98:**285–294.
19. **Lachner, M., D. O'Carroll, S. Rea, K. Mechtler, and T. Jenuwein.** 2001. Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. Nature **410:**116–120.
20. **Laible, G., A. Wolf, R. Dorn, G. Reuter, C. Nislow, A. Lebersorger, D. Popkin, L. Pillus, and T. Jenuwein.** 1997. Mammalian homologues of the Polycomb-group gene Enhancer of zeste mediate gene silencing in Drosophila heterochromatin and at S. cerevisiae telomeres. EMBO J. **16:**3219–3232.
21. **Luger, K., A. W. Mader, R. K. Richmond, D. F. Sargent, and T. J. Richmond.** 1997. Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature **389:**251–260.
22. **Marmorstein, R.** 2001. Protein modules that manipulate histone tails for chromatin regulation. Nat. Rev. Mol. Cell Biol. **2:**422–432.
23. **Mizzen, C. A., J. E. Brownell, R. G. Cook, and C. D. Allis.** 1999. Histone acetyltransferases: preparation of substrates and assay procedures. Methods Enzymol. **304:**675–696.
24. **Mumberg, D., R. Muller, and M. Funk.** 1995. Yeast vectors for the controlled expression of heterologous proteins in different genetic backgrounds. Gene **156:**119–122.
25. **Nakayama, J., J. C. Rice, B. D. Strahl, C. D. Allis, and S. I. Grewal.** 2001.

Role of histone H3 lysine 9 methylation in epigenetic control of heterochromatin assembly. Science **292:**110–113.

26. **Nielsen, S. J., R. Schneider, U. M. Bauer, A. J. Bannister, A. Morrison, D. O'Carroll, R. Firestein, M. Cleary, T. Jenuwein, R. E. Herrera, and T. Kouzarides.** 2001. Rb targets histone H3 methylation and HP1 to promoters. Nature **412:**561–565.

27. **Nislow, C., E. Ray, and L. Pillus.** 1997. SET1, a yeast member of the trithorax family, functions in transcriptional silencing and diverse cellular processes. Mol. Biol. Cell **8:**2421–2436.

28. **Orphanides, G., and D. Reinberg.** 2000. RNA polymerase II elongation through chromatin. Nature **407:**471–475.

29. **Peterson, C. L., A. Dingwall, and M. P. Scott.** 1994. Five SWI/SNF gene products are components of a large multisubunit complex required for transcriptional enhancement. Proc. Natl. Acad. Sci. USA **91:**2905–2908.

30. **Rabilloud, T., G. Carpentier, and P. Tarroux.** 1988. Improvement and simplification of low-background silver staining of proteins by using sodium dithionite. Electrophoresis **9:**288–291.

31. **Rea, S., F. Eisenhaber, D. O'Carroll, B. D. Strahl, Z. W. Sun, M. Schmid, S. Opravil, K. Mechtler, C. P. Ponting, C. D. Allis, and T. Jenuwein.** 2000. Regulation of chromatin structure by site-specific histone H3 methyltransferases. Nature **406:**593–599.

32. **Rice, J. C., and C. D. Allis.** 2001. Histone methylation versus histone acetylation: new insights into epigenetic regulation. Curr. Opin. Cell Biol. **13:**263–273.

33. **Rosenfeld, J., J. Capdevielle, J. C. Guillemot, and P. Ferrara.** 1992. In-gel digestion of proteins for internal sequence analysis after one- or two-dimensional gel electrophoresis. Anal. Biochem. **203:**173–179.

34. **Schultz, J., R. R. Copley, T. Doerks, C. P. Ponting, and P. Bork.** 2000. SMART: a web-based tool for the study of genetically mobile domains. Nucleic Acids Res. **28:**231–234.

35. **Strahl, B. D., and C. D. Allis.** 2000. The language of covalent histone modifications. Nature **403:**41–45.

36. **Strahl, B. D., R. Ohba, R. G. Cook, and C. D. Allis.** 1999. Methylation of histone H3 at lysine 4 is highly conserved and correlates with transcriptionally active nuclei in Tetrahymena. Proc. Natl. Acad. Sci. USA **96:**14967–14972.

37. **Turner, B. M.** 2000. Histone acetylation and an epigenetic code. Bioessays **22:**836–845.

38. **Tyler, J. K., and J. T. Kadonaga.** 1999. The "dark side" of chromatin remodeling: repressive effects on transcription. Cell **99:**443–446.

39. **van Holde, K. E.** 1989. Histone modifications, p. 111–148. Springer, New York, N.Y.

40. **Wolffe, A. P., and D. Guschin.** 2000. Review: chromatin structural features and targets that regulate transcription. J. Struct. Biol. **129:**102–122.

41. **Workman, J. L., and R. E. Kingston.** 1998. Alteration of nucleosome structure as a mechanism of transcriptional regulation. Annu. Rev. Biochem. **67:**545–579.

42. **Zhang, Y., and D. Reinberg.** 2001. Transcriptional regulation by histone methylation: interplay between different covalent modifications of the core histone tails. Genes Dev. **15:**2343–2360.

43. **Zhang, Y., Z. W. Sun, R. Iratni, H. Erdjument-Bromage, P. Tempst, M. Hampsey, and D. Reinberg.** 1998. SAP30, a novel protein conserved between human and yeast, is a component of a histone deacetylase complex. Mol. Cell **1:**1021–1031.