

Purification and Functional Characterization of a Histone H3-Lysine 4-Specific Methyltransferase

Hengbin Wang,¹ Ru Cao,¹ Li Xia,¹
Hediye Erdjument-Bromage,² Christoph Borchers,¹
Paul Tempst,² and Yi Zhang^{1,3}

¹Department of Biochemistry and Biophysics
Lineberger Comprehensive Cancer Center
University of North Carolina at Chapel Hill
Chapel Hill, North Carolina 27599

²Molecular Biology Program
Memorial Sloan Kettering Cancer Center
New York, New York 10021

Summary

Methylation of histone H3 at lysine 9 by SUV39H1 and subsequent recruitment of the heterochromatin protein HP1 has recently been linked to gene silencing. In addition to lysine 9, histone H3 methylation also occurs at lysines 4, 27, and 36. Here, we report the purification, molecular identification, and functional characterization of an H3-lysine 4-specific methyltransferase (H3-K4-HMTase), SET7. We demonstrate that SET7 methylates H3-K4 *in vitro* and *in vivo*. In addition, we found that methylation of H3-K4 and H3-K9 inhibit each other. Furthermore, H3-K4 and H3-K9 methylation by SET7 and SUV39H1, respectively, have differential effects on subsequent histone acetylation by p300. Thus, our study provides a molecular explanation to the differential effects of H3-K4 and H3-K9 methylation on transcription.

Introduction

Packaging of DNA into chromatin allows efficient storage of genetic information but also impedes access to DNA by protein factors. Therefore, dynamic changes in chromatin structure play important roles in DNA-templated processes such as DNA replication, repair, recombination, and transcription. Studies in the past several years have identified at least two types of protein complexes that are capable of altering chromatin structure to allow protein factors access to nucleosomal DNA. One involves multiprotein complexes that utilize energy derived from ATP hydrolysis to “remodel” nucleosomes (Kingston and Narlikar, 1999); the other involves covalent modification of core histone tails (Spencer and Davie, 1999). Of the different covalent modifications identified so far, lysine acetylation is the most intensively studied due to its fundamental role in transcription regulation (Roth et al., 2001). Other prominent histone covalent modifications include methylation, phosphorylation, and ubiquitination (Spencer and Davie, 1999).

Although discovered three decades ago (Murray, 1964), histone methylation remains one of the least understood forms of posttranslational modification. Early studies using metabolic labeling followed by sequencing of bulk histones have shown that several lysine residues,

including lysines 4, 9, 27, and 36 of H3 and lysine 20 of H4, are preferred sites of methylation (reviewed in van Holde, 1988; Strahl et al., 1999). Based on the observation that acetylated H3 and H4 are often the targets of ongoing methylation (Annunziato et al., 1995; Hendzel and Davie, 1989; Strahl et al., 1999), it has been suggested that histone methylation may have a role in transcriptional regulation. However, direct evidence linking histone methylation to gene activity was not available until recently. One major obstacle in studying the function of histone methylation is the lack of information regarding the responsible enzymes.

Great progress has been made in the identification and characterization of HMTases recently (Jenuwein and Allis, 2001; Zhang and Reinberg, 2001). Importantly, methylation of histones at different residues seems to play distinctive roles in transcriptional regulation (Jenuwein and Allis, 2001; Zhang and Reinberg, 2001). It was demonstrated that a nuclear receptor coactivator-associated protein, CARM1 (or PRMT4), is an H3-specific arginine methyltransferase and that a mutation that cripples the methyltransferase activity also cripples its coactivator activity (Chen et al., 1999). The role of histone arginine methylation on nuclear hormone receptor-mediated transcriptional activation was further supported by the recent demonstration that H4-Arg3 methylation by PRMT1 facilitates subsequent histone acetylation and nuclear hormone-dependent transcriptional activation (Strahl et al., 2001; Wang et al., 2001). In addition to the HMTases specific for arginine residues, enzymes that methylate histones at lysine residues have also been identified (Rea et al., 2000; Tachibana et al., 2001; Yang et al., 2001). For example, SUV39H1, the human homolog of the *Drosophila* heterochromatic protein Su(var)3-9, was shown to be an H3-lysine 9-specific methyltransferase (Rea et al., 2000). Importantly, methylation of H3-lysine 9 by SUV39H1 creates a recognition site for subsequent binding of the heterochromatin protein HP1, providing substantial evidence linking H3-K9 methylation to heterochromatin formation and transcriptional silencing (Bannister et al., 2001; Lachner et al., 2001; Nakayama et al., 2001).

In addition to lysine 9, at least three more lysine residues on histone H3, K4, K27, and K36, can be methylated. To understand the function of histone methylation on these residues, we set out to purify the responsible enzymes. Here, we report the purification, molecular identification, and functional characterization of an H3-specific methyltransferase. Our study uncovers a SET domain-containing protein, named SET7, as an H3-K4-specific methyltransferase. In addition, we provide evidence indicating that methylation of histone H3 at K4 and K9 by SET7 and SUV39H1, respectively, has differential effects on subsequent histone acetylation by p300.

Results

Purification of an H3-Specific HMTase

By following histone methyltransferase activity, we have previously purified an H4-R3-specific methyltransferase

³Correspondence: yi_zhang@med.unc.edu

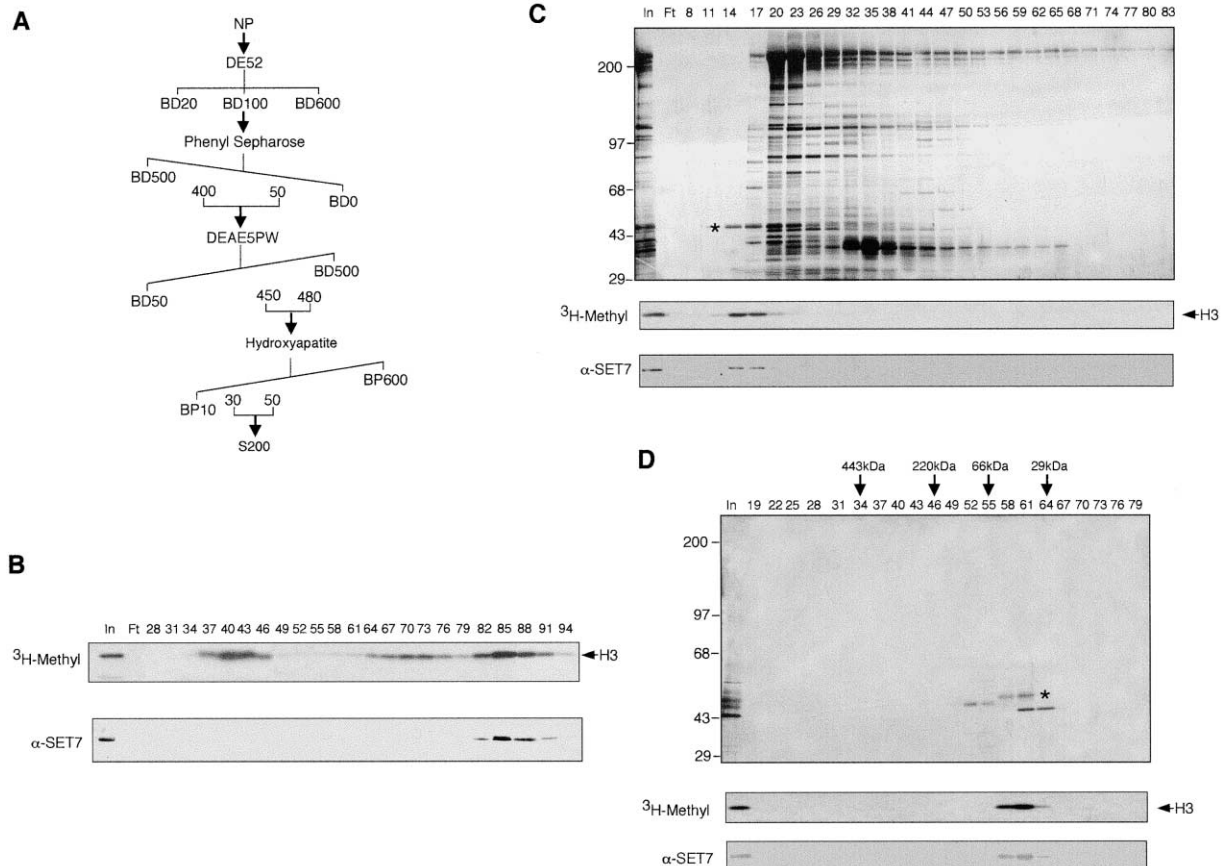


Figure 1. Purification of an H3-Specific Methyltransferase

(A) Schematic representation of the steps used to purify the H3-specific HMTase. Numbers represent the salt concentrations (mM) at which the HMTase activity elutes from the columns.

(B) Histone methyltransferase activity assay (upper panel) and Western blot analysis (lower panel) of column fractions derived from the DEAE5PW column. HMTase assay conditions were described in Experimental Procedures. The antibody used for Western blot is indicated on the left side of the panel. In and Ft represent Input and Flow-through, respectively.

(C) Silver staining of a polyacrylamide-SDS gel (top panel), HMTase activity assay (middle panel), and Western blot analysis (bottom panel) of the fractions derived from the hydroxyapatite column. The polypeptide that coeluted with the HMTase activity is indicated by a *.

(D) Silver staining of a polyacrylamide-SDS gel (top panel), HMTase activity assay (middle panel), and Western blot analysis (bottom panel) of the fractions derived from the S200 column. The elution profile of the protein markers is indicated on top of the panel. The polypeptide that coeluted with the HMTase activity is indicated by a *.

and identified it as the protein arginine methyltransferase PRMT1 (Wang et al., 2001). This biochemical approach also revealed that multiple HMTase activities specific for histone H3 or H4 exist in HeLa cells (Wang et al., 2001). Since multiple lysine residues on histone H3 are subjected to methylation, yet only H3-K9 methyltransferases are identified, we set out to identify proteins responsible for other site-specific histone methylation. Using a five-column purification scheme (Figure 1A), we have purified an H3-specific HMTase activity to homogeneity from HeLa nuclear pellet. After fractionation on DE52 and Phenyl Sepharose columns, the H3-specific HMTase activity split into three peaks on the DEAE5PW column (Figure 1B, top panel). Further fractionation of the third peak on a hydroxyapatite column followed by silver staining and enzymatic assay allowed us to correlate the HMTase activity with a 45 kDa protein (Figure 1C, top and middle panels). To confirm this result, the active fractions were pooled and loaded onto a gel-filtration Superose-200 column. Analysis of the column

fractions indicated that the peak of the enzymatic activity eluted between 29 and 66 kDa (Figure 1D, middle panel). Silver staining of an SDS-polyacrylamide gel containing the column fractions revealed again that the 45 kDa polypeptide coeluted with the enzymatic activity (Figure 1D, top panel). These results collectively support that the 45 kDa protein is likely responsible for the H3-specific HMTase activity and that it functions as a monomer.

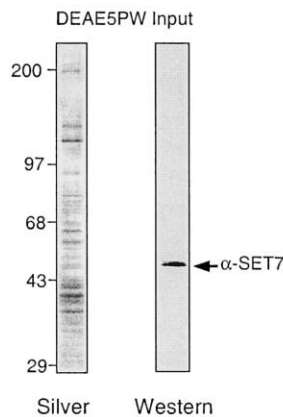
Identification of a SET Domain-Containing Protein as an H3-Specific HMTase

To identify the 45 kDa putative HMTase, we excised the band and subjected it to in-gel tryptic digestion. The resulting peptides were analyzed by mass spectrometry (Erdjument-Bromage et al., 1998). All masses obtained match a conceptual translated cDNA clone, KIAA1717 (accession number AB051504), which was deposited in the GenBank as a result of a sequencing project (Nagase et al., 2000). Although the cDNA is about 7 kb, the coding region is only 1.1 kb. The function of the large 3' UTR

A

	10	20	30	40	50	60	70	80
hSET7	MDSDDDEVVEE	AVEGHLDDDG	LPHGFCTVTY	SSTDRFEGNF	VHGEKNGRGK	FFFDGSLTE	GYVDDALQG	QGVVYYEDGG
mSET7	-----V---							
	90	100	110	120	130	140	150	160
hSET7	VLQGTIVDGE	LNGPAQEYDT	DGRLIFKQGY	KDNIRHGVCW	IYYPDGGSLV	GEVNEDGEMT	GEKIAYVYPD	ERTALYGKFI
mSET7	-----S-----		-----N-----		-H-----			Q-----
	170	180	190	200	210	220	230	240
hSET7	DGEMIEGKLA	TLMSTEEGRP	HFELMPGNSV	YHFDKSTSSC	ISTNALLPDP	YESERVYVAE	SLISSAGEGL	FSKVAVGPNT
mSET7	-----A-----		---VTS-S---		--SD-----		-----D-----	
	250	260	270	280	290	* 300	310	320
hSET7	VMSFYNGVRI	THQEVDSRDW	ALNGNTLSLD	EETVIDVPEP	YNHVSRYCAS	LGHKANHSFT	PNCIYDMFVH	PRFGPIKCIK
mSET7	-----						---V--L---	
	330	340	350	360				
hSET7	TLRAVEADEE	LTVAAYGYDHS	PPSKSGPEAP	EWYQVELKAF	QATQQK			
mSET7	-----E---							

B



C

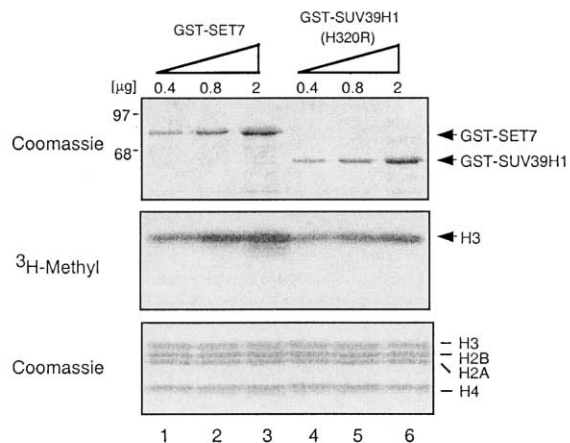


Figure 2. Identification of the H3-Specific HMTase SET7

(A) Amino acid sequences of the SET domain-containing H3-specific HMTase from human (accession number AF448509), and mouse (accession number AF448510). Amino acids that are identical between human and mouse SET7 are indicated by -. Peptides obtained from mass spectrometric analysis are underlined. Highlighted sequences represent SET domain. * indicates the amino acid that, when mutated, destroys the SET7 HMTase activity.

(B) SET7 antibody is highly specific. Silver staining (left) and Western blot (right) analysis of an aliquot (40 μ l) of the DEAE5PW input samples. Western blot was probed with the SET7 antibody.

(C) Comparison of the HMTase activity of SET7 with that of the hyperactive SUV39H1 mutant. Different amounts of enzymes indicated on top of the panel (top panel) were compared for their enzymatic activities (middle panel) using equal amounts of core histone substrates (bottom panel).

is not known. Several lines of evidence indicate that the 366 amino acids encoded by the cDNA represent the full-length protein. First, the calculated molecular mass 41 kDa is close to the estimated 45 kDa apparent molecular mass. Second, the nucleotide sequence around the first methionine conforms to the Kozak initiator sequence (Kozak, 1980). Third, transfection of a mammalian expression vector containing the coding region of KIAA1717 into 293T cells generated a product of 45 kDa (data not shown).

KIAA1717 encodes an open reading frame of 366 amino acids (Figure 2A). Analysis of the amino acid sequences revealed a SET domain at its C terminus (Figure 2A). A search of the GenBank databases with the amino acid sequence revealed no apparent homolog in yeast, *Drosophila*, or *C. elegans* (no significant homology was found outside the SET domain). However, multiple mouse EST clones with significant amino acid identity to that derived from KIAA1717 exist. Assembly of these

mouse EST sequences allowed us to derive an amino acid sequence that is 96% identical to its human counterpart (Figure 2A). The high degree of conservation of the protein sequences between human and mouse, in combination with the lack of apparent homolog from yeast, *Drosophila*, or *C. elegans*, suggests that this protein is likely involved in vertebrate functions. Since the SET domain is the only recognizable motif found in this protein and the protein has no homology to any of the six yeast SET domain-containing proteins outside the SET domain, we named this protein SET7.

Previous studies on the SUV39H1 HMTase demonstrated that in addition to the SET domain, the adjacent cysteine-rich preSET and postSET domains are also required for HMTase activity (Rea et al., 2000). However, SET7 contains neither a preSET nor a postSET domain. To confirm that KIAA1717 indeed encodes the 45 kDa HMTase, antibodies against SET7 were generated. The fact that the antibody recognized a single band on West-

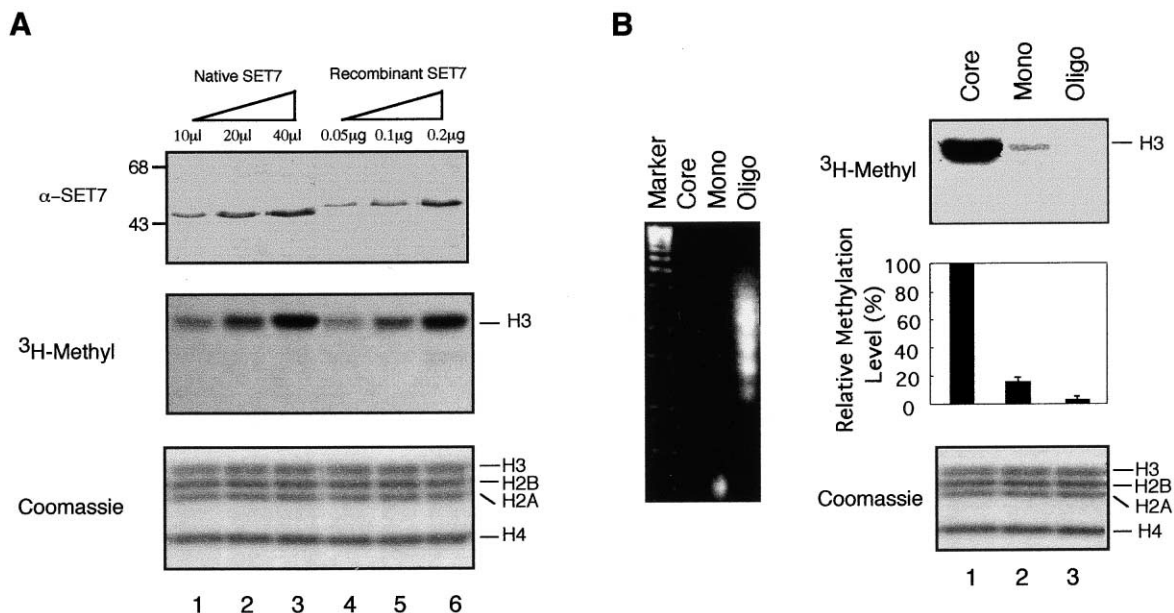


Figure 3. SET7 Prefers Nonnucleosomal Histone H3 Substrate

(A) SET7 purified from HeLa cells or expressed in *E. coli* has similar activity. Equal amounts of SET7 purified from HeLa cells (lanes 1–3) or *E. coli* (lanes 4–6) (top panel) are compared for their ability to methylate equal amounts of core histone substrates (bottom panel). The middle panel is an autoradiography of the bottom panel. Recombinant SET7 migrates a little slower due to the presence of a tag at its N terminus (top panel).

(B) SET7 prefers nonnucleosomal histone H3 substrate. The left panel is an ethidium bromide-stained agarose gel containing core histones, mononucleosomes, and oligonucleosomes as indicated on the top of the panel. Equal amounts of SET7 were used to methylate equal amounts of histones in octamer, mono-, and oligonucleosome forms (bottom panel). The top panel is an autoradiography of the bottom panel. Quantification of two independent experiments is shown in the middle panel. Variations of the two experiments were indicated by the error bars.

ern indicated that the antibody is highly specific (Figure 2B). Western blot analysis of the column fractions derived from the last three purification steps demonstrated that SET7 copurifies with the H3-specific HMTase activity (Figures 1B–D, bottom panels). These data, in combination with the mass spectrometry analysis data, confirm that KIAA1717 indeed encodes the 45 kDa HMTase. We noted that proteins other than SET7 must be responsible for the first two HMTase activity peaks detected in the DEAE5PW column fractions, because SET7 was not detectable in those fractions (Figure 1B).

Having established that SET7 copurifies with the H3-specific HMTase activity (Figure 1), we asked whether recombinant SET7 possesses HMTase activity. To this end, GST-SET7 fusion protein was generated, and its ability to methylate histone H3 was compared with that of a hyperactive SUV39H1 mutant, GST-SUV39H1(H320R) (Rea et al., 2000). Quantification of the results presented in Figure 2C indicate that GST-SET7 is more than twice as active as the hyperactive SUV39H1 mutant when equal amounts of enzymes are compared. Since the hyperactive SUV39H1 mutant was estimated to be more than 20-fold more active than the wild-type SUV39H1 (Rea et al., 2000), we conclude that SET7 is at least 40-fold more active than SUV39H1 in methylating histone H3 when equal amounts of proteins are compared.

SET7 Prefers Nonnucleosomal Histone H3 Substrate

Having established that recombinant SET7 possesses HMTase activity, we asked whether it is as active as that purified from HeLa cells. To this end, equal amounts

of SET7 purified from HeLa cells or expressed in *E. coli* were compared for their ability to methylate core histones. Results shown in Figure 3A indicate that SET7 from HeLa cells or expressed in *E. coli* has no significant difference in its ability to methylate histone H3.

Since core histones in the cell nucleus exist in nucleosomal form, we asked whether SET7 by itself is capable of methylating H3 assembled into nucleosomes. To this end, mono- and oligonucleosomes (Figure 3B, left panel) were prepared and were compared with core histone octamers for their ability to serve as substrates for SET7. Results shown in Figure 3B indicate assembly of histone H3 into nucleosomes significantly reduced the ability of SET7 to methylate H3, although SET7 still has detectable activity toward H3 assembled into mononucleosomes. Previous studies indicate that histone acetyltransferases have different substrate specificity depending on whether they function alone or together with other associated proteins. For example, GCN5 preferentially acetylates free histone when used as a monomeric form (Kuo et al., 1996). However, it is capable of methylating nucleosomal histones when incorporated into the SAGA complex (Grant et al., 1997). Given that SET7 is able to methylate H3-K4 *in vivo* (see below), it is likely that SET7 may function in concert with other proteins in cells.

The SET Domain and Its Adjacent Sequences on SET7 Are Required for Its HMTase Activity

Having demonstrated that SET7 is a potent H3-specific HMTase (Figures 2 and 3), we attempted to determine the minimum amino acid sequences required for its

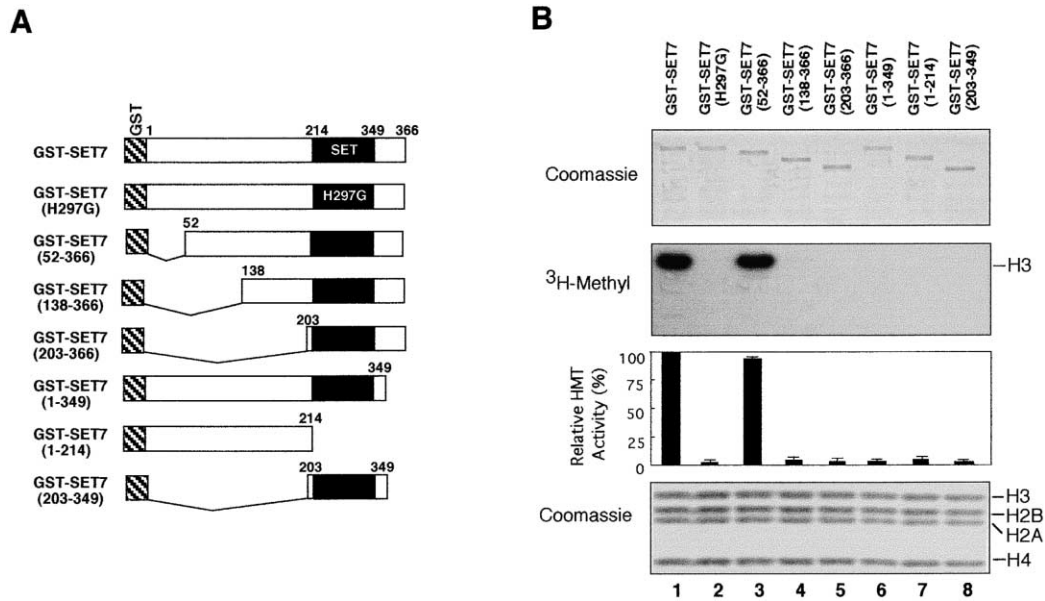


Figure 4. Identification of the Minimum Amino Acid Sequences Required for SET7 HMTase Activity

(A) Schematic representation of the different SET7 deletion mutants. The SET domain and the point mutation are indicated. The amino acid numbers for each construct are indicated.

(B) Equal amounts of wild-type and mutant GST-SET7 proteins (top panel) were compared for their HMTase activities (middle panels) using equal amounts of core histone substrates (bottom panel). The top and bottom panels are from the same 18% polyacrylamide-SDS gel. The relative HMTase activity represents the average of two independent experiments. The variations are indicated by error bars.

HMTase activity. Studies on SUV39H1 indicated that the SET domain is required for its HMTase activity (Rea et al., 2000). To determine whether the same is true for SET7, a SET7 mutant protein that contains a single amino acid change (H297G) was generated. HMTase activity assay demonstrated that the mutant protein is devoid of detectable enzymatic activity (Figure 4B, middle two panels, compare lanes 1 and 2). Therefore, the SET domain in SET7 is critical for its HMTase activity. In addition to the SET domain, the adjacent cysteine-rich preSET and postSET domains of SUV39H1 were also demonstrated to be required for HMTase activity (Rea et al., 2000). However, neither a preSET nor post-SET domain was identified in SET7 (Figure 2A). Therefore, we asked whether any sequences in SET7 other than the SET domain are required for its HMTase activity. To this end, a series of N-terminal and C-terminal deletion mutant SET7 proteins were generated and tested for HMTase activity (Figure 4A). Results shown in Figure 4B indicate that amino acids flanking the SET domain on both sides are required for SET7 HMTase activity. The only sequences that are dispensable for the HMTase activity are the extreme N-terminal amino acids (Figure 4B, middle two panels, compare lanes 1 and 3). It is possible that sequences adjacent to the SET domain, regardless of whether they are cysteine rich or not, might play an important role in proper folding of the SET domain to accommodate histone tails.

SET7 Methylates H3 at Lysine 4 In Vitro and In Vivo

Having established that SET7 is an H3-specific HMTase, we attempted to determine the methylation site. It was previously reported that the N terminus of histone H3 contains at least three lysine residues (K4, K9, and K27)

that can be methylated. To determine which of the three sites is methylated by SET7, core histone octamers were subjected to methylation by SET7 in the presence of S-adenosyl-L-[methyl-³H]methionine (³H-SAM). After separation by SDS-PAGE, methylated H3 was recovered and microsequenced by automated Edman chemical sequencing. Sequentially released amino acid derivatives were collected and counted by liquid scintillation. Results shown in Figure 5A reveal that Lysine 4 is the only methylated residue.

It is well known that the recovery rate drops as the number of microsequencing cycles increases. To avoid potential missing of methylation sites due to this microsequencing-associated problem, we generated GST-H3 fusion proteins with or without mutations at the three methyl-able lysine residues and compared their ability to serve as substrates for SET7. Results shown in Figure 5B indicate that while mutation on lysine 4 completely eliminated the ability of H3 to serve as SET7 substrate (Figure 5B, compare lanes 2, 5, 7, and 8 with 1), mutation on lysine 9 and lysine 27, either individually or in combination, has little effect on the ability of H3 to serve as SET7 substrate (Figure 5B, compare lanes 3, 4, and 6 with 1). This result, in combination with the radioactive mapping result, allows us to conclude that H3-K4 is the only lysine residue on histones that can be methylated by SET7.

Having established that SET7 specifically methylates H3-K4 in vitro, we asked whether SET7 is capable of methylating H3-K4 in vivo. To this end, a methyl-K4-specific antibody (Nakayama et al., 2001) was used to monitor the H3-K4 methylation status. As expected, the antibody recognized endogenous histone H3 purified from 293T cells transfected with an empty mammalian expres-

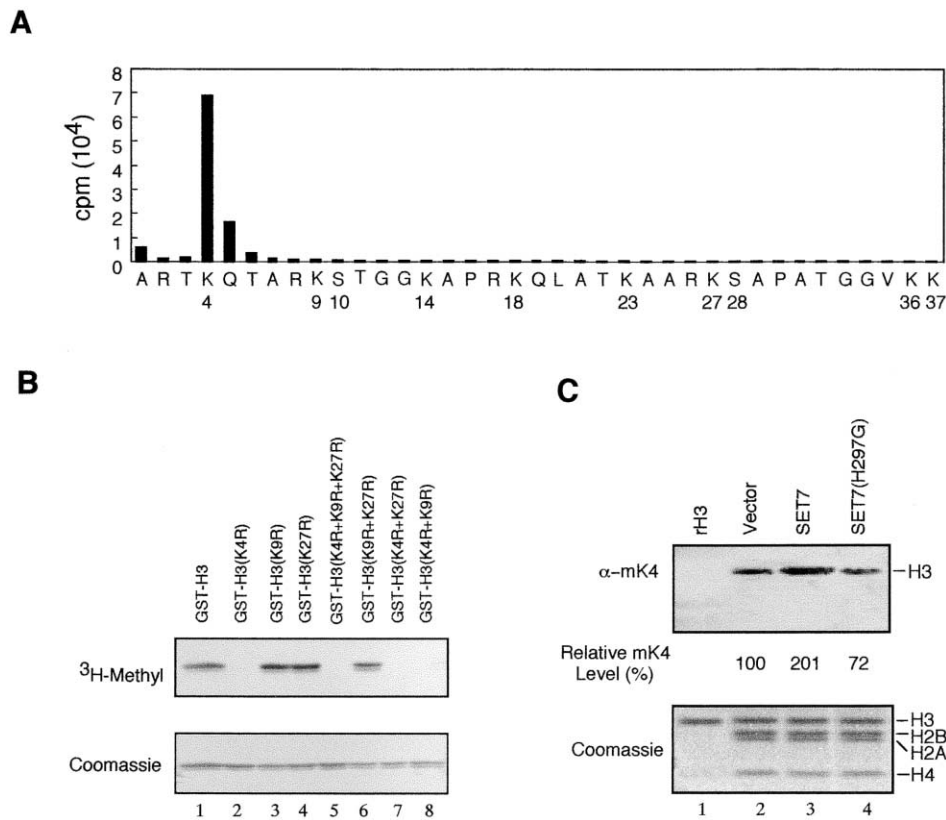


Figure 5. SET7 Methylates Histone H3-Lysine 4 In Vitro and In Vivo

(A) Lysine 4 of H3 is the major methylation site of SET7. In vitro SET7-methylated H3 was blotted onto PVDF membrane and subjected to N-terminal automated sequencing. ³H-radioactivity eluted from each cycle was counted. The amino acids identified at each cycle of micro-sequencing are listed.

(B) HMTase assay using different GST-H3 tail fusion proteins as substrates. The lysines that were mutated are indicated on top of the panel. The bottom panel indicates that equal amounts of substrates were used in each reaction. The top panel is an autoradiogram of the same gel shown in the bottom panel.

(C) SET7 methylates H3-K4 in vivo. Recombinant histone H3 (200 ng) and equivalent amounts of core histones (bottom panel) purified from 293T cells transfected with empty vector (lane 2), a vector that expresses wild-type SET7 (lane 3), or a vector that expresses mutant SET7 (lane 4), respectively, were analyzed by Western blot using the methyl-K4-specific antibody (top panel). The numbers between the two panels were derived from the top panel, and they represent the relative H3-K4 methylation levels.

sion vector, while it did not recognize an equal amount of recombinant H3 purified from *E. coli* (Figure 5C, compare lanes 1 and 2), indicating that the antibody is methyl-K4 specific. Importantly, transfection of a mammalian expression vector encoding SET7 doubled the H3-K4 methylation level (compare lanes 2 and 3). In contrast, transfection of a vector expressing the HMTase-defective SET7(H297G) mutant reduced the endogenous H3-K4 methylation level (compare lanes 4 and 2), indicating that this mutant is likely function in a dominant-negative fashion. Collectively, these data suggest that SET7 is at least partially responsible for H3-K4 methylation in vivo.

H3-K4 and -K9 Methylation Inhibit Each Other

Accumulating evidence indicates that different modifications at the histone N-terminal tails affect each other (reviewed in Zhang and Reinberg, 2001). Since SET7 specifically methylates lysine 4 of H3, we attempted to address the effect of H3-K4 methylation on subsequent histone modifications at other residues. To this end, core histones were methylated at K4 or K9 using SET7 or SUV39H1 in the presence of nonradioactive SAM. After purification and verification of the methylation

(cannot be further methylated in the presence of ³H-SAM), the histones that were methylated on K4 or K9 were subjected to methylation by SUV39H1 or SET7 in the presence of ³H-SAM. Parallel mock-methylated histones were used as a control. Results shown in Figure 6A indicate that methylation on H3-K4 by SET7 inhibits subsequent K9 methylation by SUV39H1 (compare lanes 1 and 2). Similarly, methylation on K9 by SUV39H1 inhibits subsequent K4 methylation by SET7 (Figure 6A, compare lanes 3 and 4). To confirm the above results, H3 peptides that are unmodified, K4 dimethylated, or K9 dimethylated, respectively, were subjected to methylation using SET7 or SUV39H1 in the presence of ³H-SAM. Results shown in Figure 6B confirm the results obtained using core histones as substrates. Collectively, the above data allow us to conclude that H3-K4 and -K9 methylation inhibit each other.

H3-K4 and -K9 Methylation Differentially Affect Subsequent Histone Acetylation by p300

Having established the antagonistic relationship between H3-K4 and -K9 methylation, we asked whether K4 and K9 methylation have differential effects on sub-

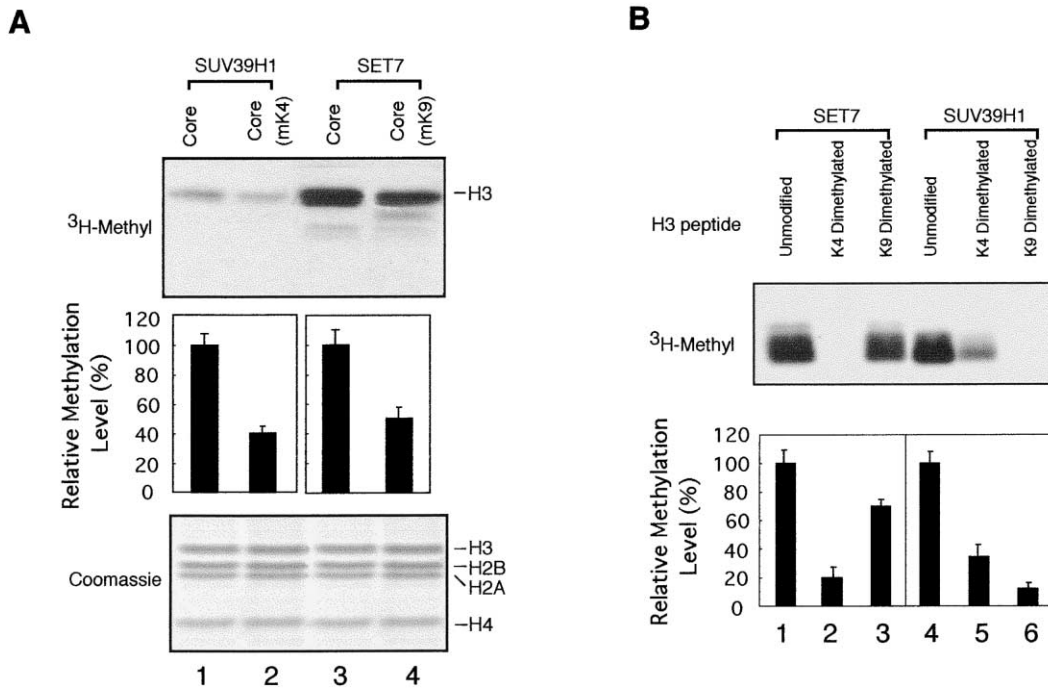


Figure 6. H3-K4 and H3-K9 Methylation Inhibit Each Other

(A) Core histones in which H3 is methylated at K4 or K9 were used as substrates for methylation by SUV39H1 or SET7, respectively (top panel). Quantification of H3 methylation by the two enzymes is shown relative to that when the unmodified core histones were used as substrates. Variations between two independent experiments are depicted by error bars. The bottom panel indicates that equal amounts of histone substrates were used in each reaction.

(B) Peptides corresponding to human H3 N-terminal 19 amino acids that are unmodified, K4 dimethylated, and K9 dimethylated, respectively, were subjected to methylation by SET7 (lanes 1–3) or SUV39H1 (lanes 4–6). Quantification of the methylation level (bottom panel) was based on when unmodified peptide was used as a substrate. Variations between two independent experiments are depicted by error bars.

sequent histone acetylation by p300. To this end, recombinant histone H3 that are mock-, K4-, or K9 methylated are subjected to acetylation by p300 in the presence of ³H-acetyl CoA. Results shown in Figure 7A indicated that K4 methylation facilitates subsequent acetylation by p300 (compare lanes 1, 3, and 5 with 2, 4, and 6). In contrast, methylation on K9 inhibits subsequent acetylation by p300 (compare lanes 7, 9, and 11 with 8, 10, and 12). To confirm the above results with a more physiological substrate, core histone octamers that were K4- or K9 methylated by SET7 or SUV39H1, respectively, were subjected to acetylation by p300 in the presence of ³H-acetyl CoA. Parallel experiments using mock-methylated core histones as substrates were also performed. Results shown in Figure 7B not only confirmed the positive/negative effect of K4/K9 methylation on subsequent H3 acetylation by p300 but importantly also revealed that methylation on H3 affects subsequent acetylation on H4 by p300 (compare lane 1 with 2, 3 with 4). The demonstration that covalent modification on one histone molecule can affect subsequent modification on a different histone molecular has important implications for the “histone code” hypothesis (Strahl and Allis, 2000). It significantly increases the complexity of the histone code. Given that p300-mediated histone acetylation correlates with gene activation, our data may provide a molecular basis for the observed correlation between H3-K9/H3-K4 methylation and transcriptional repression/activation (Litt et al., 2001; Noma et al., 2001).

Discussion

By following histone H3-specific methyltransferase activity, we isolated a new HMTase that we named SET7. SET7 is capable of methylating histone H3 at lysine 4 in vitro and in vivo. In addition, methylation of H3-K4 by SET7 inhibits subsequent methylation on K9 by SUV39H1. In contrast to its inhibitory effect on K9 methylation, H3-K4 methylation facilitates subsequent acetylation of H3 and H4 by p300.

SET Domain, a Lysine-Specific Methyltransferase Signature

The SET domain is an evolutionarily conserved sequence motif initially identified in the *Drosophila* PEV (position effect variegation) suppressor SU(VAR)3-9 (Tschiersch et al., 1994), the *Polycomb*-group protein Enhancer of zeste (Jones and Gelbart, 1993), and the *trithorax*-group protein Trithorax (Stassen et al., 1995). A major function of the SET domain-containing proteins is to modulate gene activity (Jenuwein et al., 1998). However, the underlying mechanism is just beginning to be revealed. The demonstration that SUV39H1 and its fission yeast homolog Clr4 possess intrinsic H3-specific HMTase activity suggests that SET domain-containing proteins might modulate gene activity through histone methylation (Rea et al., 2000). Mutational studies with SUV39H1 revealed that the SET domain as well as the two adjacent cysteine-rich preSET and postSET do-

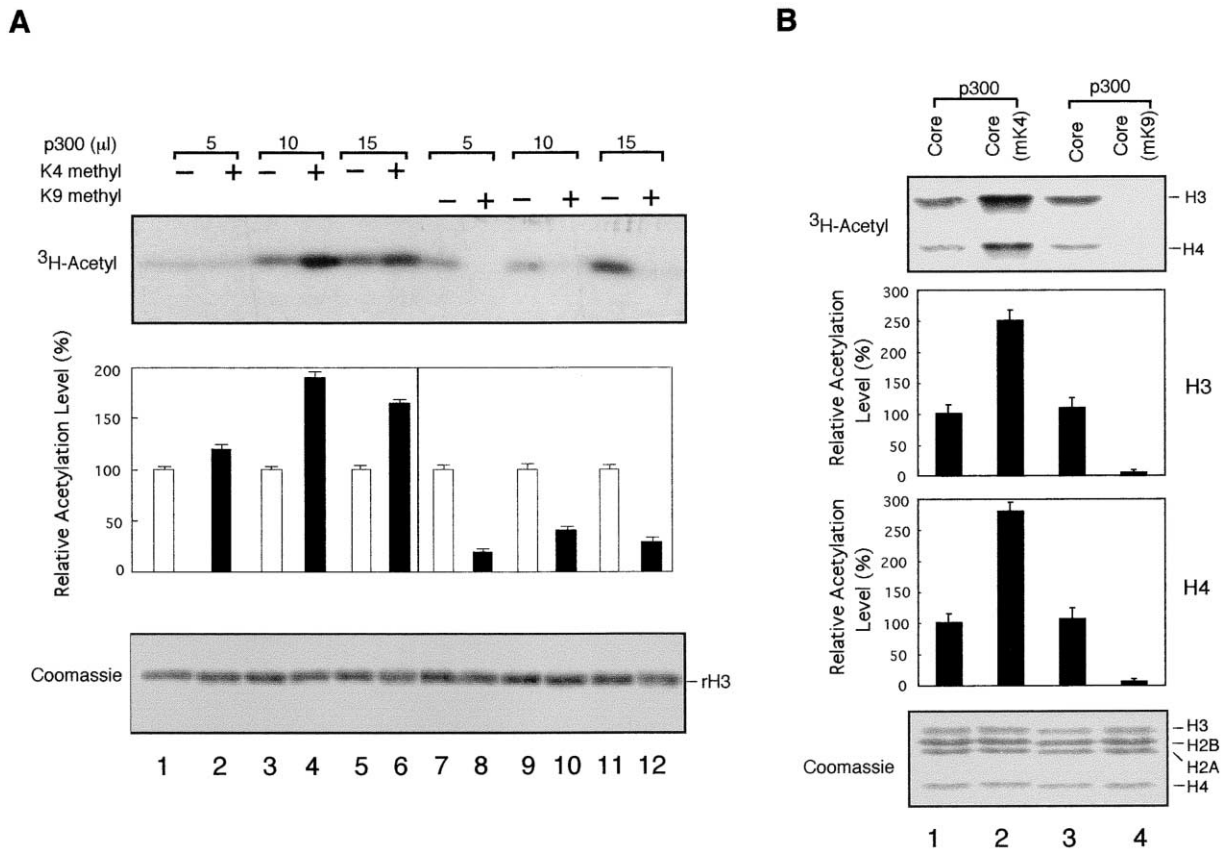


Figure 7. Differential Effects of H3-K4 and H3-K9 Methylation on Subsequent Lysine Acetylation
 (A) H3-K4/K9 methylation facilitates/inhibits subsequent H3 acetylation by p300. Recombinant histone H3 that was mock methylated (–) or K4- or K9-methylated (+) was used as substrates for acetylation by p300. The amount of p300 used in each reaction is indicated on top of the panel. After exposure to X-ray film, the corresponding gel slices were excised and counted by scintillation counting. All quantification is relative to that when the unmodified H3 was used as a substrate. Variations between two independent experiments are depicted by error bars. The bottom panel indicates that an equal amount of recombinant H3 was used in each reaction.
 (B) H3-K4/K9 methylation facilitates/inhibits subsequent H3 and H4 acetylation by p300. Core histones that were methylated on K4 or K9 by SET7 or SUV39H1, respectively, were compared with mock methylated core histone to serve as substrates for acetylation by p300 (10 μl, top panel). The corresponding gel slices were excised and counted by scintillation counting for both H3 and H4. Quantification is relative to that when unmodified core histones were used as substrates (middle panels). Variations between two independent experiments are depicted by error bars. Equal amounts of substrates are verified by Coomassie staining (bottom panel).

mains are required for its enzymatic activity (Rea et al., 2000). Consistent with the requirement of these three regions for enzymatic activity, other SET domain-containing proteins that lack one or both cysteine-rich regions, for example recombinant EZH2 and TRX, do not contain histone methyl-transferase activity in vitro (Rea et al., 2000). However, the demonstration that SET7, although it lacks the preSET and postSET domains, possesses intrinsic HMTase activity argues against a general requirement for the preSET and postSET domains for HMTase activity. This argument is further supported by analyzing the SET domain-containing proteins from the budding yeast. Analysis of the *S. cerevisiae* genome database identified six SET domain-containing proteins, yet none of these proteins contains both the preSET and postSET domains. Nonetheless, histones in yeast are clearly methylated (Strahl et al., 1999). Therefore, it is likely that the SET domain is the only signature motif for lysine-specific HMTases. However, SET domain alone does not seem to be sufficient for HMTase activity

(Figure 4). Our data, as well as those from Jenuwein and colleagues (Rea et al., 2000), indicate that sequences flanking the SET domain, whether they are cycteine rich or not, seem to be required for HMTase activity. Whether these flanking sequences simply facilitate folding of the SET domain to accommodate histone tails or directly involve in the enzymatic reaction awaits to be determined. A clear answer to the above question likely requires determination of the structures of these enzymes.

Methylation on Lysine 4 and Lysine 9 of Histone H3 Antagonize Each Other

Histone tails are rich in covalent modifications (van Holde, 1988). One of the most important advances in the studies of these covalent modifications is the realization that these covalent modifications affect each other. As a result, the histone code hypothesis was proposed (Strahl and Allis, 2000). This hypothesis predicts that a preexisting modification affects subsequent modifications on histone tails and that these modifications serve

as marks for the recruitment of different proteins or protein complexes to regulate diverse chromatin functions. Several recent studies have demonstrated that acetylation, phosphorylation, and methylation on histone tails affect each other. For example, several transcription-related HATs, including GCN5, PCAF, and p300, prefer S10 phosphorylated H3 as a substrate, since S10 phosphorylation facilitates subsequent H3-K14 acetylation (Cheung et al., 2000; Lo et al., 2000). In addition to the positive effect on K14 acetylation, S10 phosphorylation negatively regulates K9 methylation by SUV39H1 (Rea et al., 2000). Furthermore, we have recently demonstrated that acetylation on any of the four lysines on the histone H4 N-terminal tail inhibits subsequent methylation on H4-R3 by PRMT1. In contrast, methylation on H4-R3 by PRMT1 facilitates subsequent acetylation by p300 on histone H4-K8 and H4-K12 (Wang et al., 2001).

To understand the consequences of H3-K4 methylation on subsequent histone tail modifications, we took advantage of the fact that SET7 specifically methylates H3-K4 and analyzed the effect of H3-K4 methylation on subsequent H3-K9 methylation by SUV39H1. Our data indicate that methylation on H3-K4 and -K9 inhibit each other (Figure 6). This provides a molecular explanation to the recent demonstration that H3-K4 and -K9 have distinctive methylation patterns at the fission yeast heterochromatin domain boundaries as well as at the chicken β -globin locus (Litt et al., 2001; Noma et al., 2001). In addition to inhibiting each other, methylation on H3-K4 and -K9 has an opposite effect on subsequent histone acetylation by p300. While K9 methylation inhibited subsequent histone acetylation, K4 methylation increased subsequent histone acetylation. Interestingly, this not only occurs on histone H3 but also on histone H4, indicating that interplay between different modifications can occur among different histone molecules (Figure 7).

Given the general correlation between histone acetylation and transcriptional activation, we predict that K4 methylation is likely involved in transcriptional activation. Consistent with this prediction is the demonstration that H3-K4 is a preferred site of methylation in the transcriptionally active macronuclei of *Tetrahymena* (Strahl et al., 1999). In addition, chromatin immunoprecipitation (ChIP) experiments covering the fission yeast mating-type locus and the chicken β -globin locus have revealed a correlation between gene activation/silencing with K4/K9 methylation, respectively (Litt et al., 2001; Noma et al., 2001). However, data demonstrating that the enzymatic activities of H3-K4/K9 methylation are responsible for gene activation/repression are not available. To directly address the role of the HMTase activities of SET7 and SUV39H1 in transcription, we attempted to target wild-type and HMTase-inactivated SET7 and SUV39H1 mutants to reporter genes through the Gal4-DNA binding domain in both mammalian and replication-coupled *Xenopus* oocyte chromatin assembly systems. These efforts, however, have so far failed to demonstrate a direct role for SET7/SUV39H1 HMTase activities in transcription (data not shown). One explanation is that these enzymes might function in concert with other chromatin modifying enzymes, such as nucleosome remodeling factors. Consistent with this possibility is the finding that SET7 has marginal HMTase activity toward H3 in mononucleosomes while it is completely inactive toward

H3 in oligonucleosomes (Figure 3B), since the histone tails in mononucleosomes are more accessible than those in oligonucleosomes. Alternatively, these enzymes might function in a promoter-specific manner. In this case, demonstration of the direct involvement of the SET7/SUV39H1 HMTase activity in transcription would require the identification of their functioning partners and their target genes.

Experimental Procedures

Purification of the H3-K4 HMTase

HeLa nuclear proteins were separated into nuclear extract and nuclear pellet fractions using a previously described procedure (Dignam et al., 1983). Nuclear pellet was solubilized as described (LeRoy et al., 1998). The recovered proteins (7 g), resuspended in buffer D (50 mM Tris-HCl [pH 7.9], 0.1 mM EDTA, 2 mM DTT, 0.2 mM PMSF, and 25% glycerol) containing 20 mM ammonium sulfate (BD20), were then loaded onto a 900 ml DEAE52 column equilibrated with BD20. Proteins that bound to the DEAE52 column were step eluted with BD100 and BD600. The BD100 fraction was adjusted to BD500 using saturated ammonium sulfate and loaded to a 22 ml FPLC Phenyl Sepharose column (Pharmacia). The column was eluted with a 15 column-volume (cv) linear gradient from BD500 to BD0. The fractions that contain the HMTase activity were combined and dialyzed into BD50 and loaded onto a 45 ml HPLC-DEAE-5PW column (TosoHaas). The column was eluted with a 10 cv linear gradient from BD50 to BD500. HMTase activities were separated into three peaks on this column. The third peak fractions were pooled and dialyzed into buffer P (5 mM HEPES-KOH [pH 7.5], 40 mM KCl, 0.01% Triton X-100, 0.01 mM CaCl₂, 0.5 mM PMSF, 1 mM DTT, and 10% glycerol) containing 10 mM potassium phosphate (BP10) and loaded onto a 1 ml hydroxyapatite column. The bound proteins were eluted with 20 cv linear gradient from BP10 to BP600. The active fractions were pooled and concentrated on a 0.2 ml DEAE-52 column and fractionated on a Sepharose 200 gel-filtration column (Pharmacia). The HMTase activity elutes with a relative molecular mass between 66 and 29 kDa.

Protein Identification and Methylation Site Determination

To identify the 45 kDa protein that coelutes with the H3 HMTase activity, the 45 kDa protein was excised and digested with trypsin and processed for mass spectrometric analysis as described (Erdjument-Bromage et al., 1998). In brief, the peptide mixture was partially fractionated on a Poros 50 R2 RP microtip, and resulting peptide pools were analyzed by matrix-assisted laser-desorption/ionization reflectron time-of-flight (MALDI-reTOF) MS using a Reflex III instrument (Brüker Franzen; Bremen, Germany) and by electrospray ionization (ESI) tandem MS on an API 300 triple quadrupole instrument (PE-SCIEX; Thornhill, Canada), modified with an ultrafine ionization source as described (Geromanos et al., 2000). Selected mass values from the MALDI-TOF experiments were taken to search the NCBI nonredundant protein database using the PeptideSearch algorithm. The protein identification was verified by comparing the computer-generated fragment ion series of the predicted tryptic peptide with the experimental MS/MS data. For methylation site determination, core histones were methylated with the purified enzyme in the presence of ³H-SAM (15 Ci/mM; NEN Life Science Products). Histones were then separated by 18% SDS-PAGE and transferred to PVDF membrane. After Coomassie staining and destaining, the protein band corresponding to H3 was excised and subjected to automated Edman chemical sequencing using an Applied Biosystems (Foster City, CA) 477A automated instrument. Sequentially released amino acid derivatives were collected after each cycle for liquid scintillation counting as described (Frydman et al., 1999).

Histone Methyltransferase and Acetyltransferase Assays

HMTase and HAT assays were performed essentially as described (Wang et al., 2001). In brief, purified enzymes or column fractions (20 μ l) were incubated with core histones (4 μ l, 1 mg/ml, purified from HeLa cells), H3-tail peptides (ABCAM, Cambridge, UK), or different GST-H3 fusion proteins (Tachibana et al., 2001) in 32.5 μ l reactions

containing 20 mM Tris-HCl (pH 8.0), 4 mM EDTA, 1 mM PMSF, 0.5 mM DTT, and 1 μ l 3 H-SAM (15 Ci/mM; NEN Life Science Products) for 1 hr at 30°C. Reactions were stopped by addition of 8 μ l of 5 \times SDS loading buffer and proteins separated in an 18% SDS-PAGE. After Coomassie staining and destaining, gels were treated with Entensify (NEN Life Science Products), dried, and exposed to X-ray films. For quantification, the gel slices were excised and counted with scintillation counting. HAT assays were performed in 30 μ l volume containing 50 mM HEPES (pH 8.0), 5 mM DTT, 5 mM PMSF, 10 mM sodium butyrate, 10% glycerol, 2 μ l 3 H-acetyl-CoA, and varying amounts of p300. Reaction mixture was incubated for 1 hr at 37°C and terminated by the addition of SDS sample buffer. Histone and nucleosome substrates were purified from HeLa cells as previously described (Zhang et al., 1998). Histone tail peptides were purchased from ABCAM (Cambridge, UK).

Plasmids, Antibody, and Transfection

Plasmid encoding GST-SET7 was constructed by digestion of the KIAA1717 cDNA with NcoI and HindIII and cloned into the NcoI/HindIII sites of the pGEX-KG vector. The mammalian expression vector was constructed by digestion of the GST-SET7 with EcoRI and cloned into the EcoRI site of the pCDNA- β -Flag vector. All the GST-SET7 deletion constructs were generated by PCR and cloned into the EcoRI/XhoI sites of the pGEX-KG vector. Constructs generated through PCR were verified by DNA sequencing. The H297G mutant was generated by PCR-based mutation and was verified by DNA sequencing. All the GST-H3 tail fusion constructs used in Figure 3B have been previously described (Tachibana et al., 2001). Antibody against SET7 was generated by injecting the full-length recombinant SET7 protein into rabbits. Methyl-K4-specific antibody was purchased from Upstate Biotechnology and was described (Noma et al., 2001). Transfection and histone isolation from transfected 293T cells were performed as described (Wang et al., 2001).

Acknowledgments

We thank Dr. Takahiro Nagase for providing the KIAA1717 cDNA clones and Dr. Yoichi Shinkai for the GST-histone tail plasmids. Y.Z. is a Kimmel Scholar and is supported by the National Institutes of Health (GM63067-01) and the American Cancer Society (RSG-00-351-01-GMC). Research in the lab of P.T. is supported by a National Cancer Institute grant (P30CA08748).

Received October 26, 2001; revised November 14, 2001.

References

Annunziato, A.T., Eason, M.B., and Perry, C.A. (1995). Relationship between methylation and acetylation of arginine-rich histones in cycling and arrested HeLa cells. *Biochemistry* 34, 2916–2924.

Bannister, A.J., Zegerman, P., Partridge, J.F., Miska, E.A., Thomas, J.O., Allshire, R.C., and Kouzarides, T. (2001). Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* 410, 120–124.

Chen, D., Ma, H., Hong, H., Koh, S.S., Huang, S.M., Schurter, B.T., Aswad, D.W., and Stallcup, M.R. (1999). Regulation of transcription by a protein methyltransferase. *Science* 284, 2174–2177.

Cheung, P., Tanner, K.G., Cheung, W.L., Sassone-Corsi, P., Denu, J.M., and Allis, C.D. (2000). Synergistic coupling of histone H3 phosphorylation and acetylation in response to epidermal growth factor stimulation. *Mol. Cell* 5, 905–915.

Dignam, J.D., Lebovitz, R.M., and Roeder, R.G. (1983). Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res.* 11, 1475–1489.

Erdjument-Bromage, H., Lui, M., Lacomis, L., Grewal, A., Annan, R.S., McNulty, D.E., Carr, S.A., and Tempst, P. (1998). Examination of micro-tip reversed-phase liquid chromatographic extraction of peptide pools for mass spectrometric analysis. *J. Chromatogr. A* 826, 167–181.

Frydman, J., Erdjument-Bromage, H., Tempst, P., and Hartl, F.U. (1999). Co-translational domain folding as the structural basis for

the rapid de novo folding of firefly luciferase. *Nat. Struct. Biol.* 6, 697–705.

Geromanos, S., Freckleton, G., and Tempst, P. (2000). Tuning of an electrospray ionization source for maximum peptide-ion transmission into a mass spectrometer. *Anal. Chem.* 72, 777–790.

Grant, P.A., Duggan, L., Cote, J., Roberts, S.M., Brownel, J.E., Candau, R., Ohba, R., Owen-Hughes, T., Allis, C.D., Winston, F., et al. (1997). Yeast Gcn5 functions in two multisubunit complexes to acetylate nucleosomal histones: characterization of an Ada complex and the SAGA (Spt/Ada) complex. *Genes Dev.* 11, 1640–1650.

Hendzel, M.J., and Davie, J.R. (1989). Distribution of methylated histones and histone methyltransferases in chicken erythrocyte chromatin. *J. Biol. Chem.* 264, 19208–19214.

Jenuwein, T., and Allis, C.D. (2001). Translating the histone code. *Science* 293, 1074–1080.

Jenuwein, T., Laible, G., Dorn, R., and Reuter, G. (1998). SET domain proteins modulate chromatin domains in eu- and heterochromatin. *Cell. Mol. Life Sci.* 54, 80–93.

Jones, R.S., and Gelbart, W.M. (1993). The Drosophila Polycomb-group gene Enhancer of zeste contains a region with sequence similarity to trithorax. *Mol. Cell. Biol.* 13, 6357–6366.

Kingston, R.E., and Narlikar, G.J. (1999). ATP-dependent remodeling and acetylation as regulators of chromatin fluidity. *Genes Dev.* 13, 2339–2352.

Kozak, M. (1980). Evaluation of the “scanning model” for initiation of protein synthesis in eucaryotes. *Cell* 22, 7–8.

Kuo, M.-H., Brownell, J.E., Ranalli, T.A., Cook, R.G., Edmonson, D.G., Roth, S.Y., and Allis, C.D. (1996). Transcription-linked acetylation by Gcn5p of histones H3 and H4 at specific lysines. *Nature* 383, 269–272.

Lachner, M., O’Carroll, D., Rea, S., Mechtler, K., and Jenuwein, T. (2001). Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* 410, 116–120.

LeRoy, G., Orphanides, G., Lane, W.S., and Reinberg, D. (1998). Requirement of RSF and FACT for transcription of chromatin templates in vitro. *Science* 282, 1900–1904.

Litt, M.D., Simpson, M., Gaszner, M., Allis, C.D., and Felsenfeld, G. (2001). Correlation between histone lysine methylation and developmental changes at the chicken beta-globin locus. *Science* 293, 2453–2455.

Lo, W.S., Trievel, R.C., Rojas, J.R., Duggan, L., Hsu, J.Y., Allis, C.D., Marmorstein, R., and Berger, S.L. (2000). Phosphorylation of serine 10 in histone H3 is functionally linked in vitro and in vivo to Gcn5-mediated acetylation at lysine 14. *Mol. Cell* 5, 917–926.

Murray, K. (1964). The occurrence of e-N-methyl lysine in histones. *Biochemistry* 3, 10–15.

Nagase, T., Kikuno, R., Hattori, A., Kondo, Y., Okumura, K., and Ohara, O. (2000). Prediction of the coding sequences of unidentified human genes. XIX. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro. *DNA Res.* 7, 347–355.

Nakayama, J., Rice, J.C., Strahl, B.D., Allis, C.D., and Grewal, S.I. (2001). Role of histone H3 lysine 9 methylation in epigenetic control of heterochromatin assembly. *Science* 292, 110–113.

Noma, K., Allis, C.D., and Grewal, S.I. (2001). Transitions in distinct histone H3 methylation patterns at the heterochromatin domain boundaries. *Science* 293, 1150–1155.

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

Roth, S.Y., Denu, J.M., and Allis, C.D. (2001). Histone Acetyltransferases. *Annu. Rev. Biochem.* 70, 81–120.

Spencer, V.A., and Davie, J.R. (1999). Role of covalent modifications of histones in regulating gene expression. *Gene* 240, 1–12.

Stassen, M.J., Bailey, D., Nelson, S., Chinwalla, V., and Harte, P.J. (1995). The Drosophila trithorax proteins contain a novel variant of the nuclear receptor type DNA binding domain and an ancient

conserved motif found in other chromosomal proteins. *Mech. Dev.* 52, 209–223.

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

Strahl, B.D., Ohba, R., Cook, R.G., and Allis, C.D. (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.

Strahl, B.D., Briggs, S.D., Brame, C.J., Caldwell, J.A., Koh, S.S., Ma, H., Cook, R.G., Shabanowitz, J., Hunt, D.F., Stallcup, M.R., and Allis, C.D. (2001). Methylation of histone H4 at arginine 3 occurs in vivo and is mediated by the nuclear receptor coactivator PRMT1. *Curr. Biol.* 11, 996–1000.

Tachibana, M., Sugimoto, K., Fukushima, T., and Shinkai, Y. (2001). SET-domain containing protein, G9a, is a novel lysine-preferring mammalian histone methyltransferase with hyperactivity and specific selectivity to lysines 9 and 27 of histone H3. *J. Biol. Chem.* 276, 25309–25317.

Tschiersch, B., Hofmann, A., Krauss, V., Dorn, R., Korge, G., and Reuter, G. (1994). The protein encoded by the *Drosophila* position-effect variegation suppressor gene *Su(var)3-9* combines domains of antagonistic regulators of homeotic gene complexes. *EMBO J.* 13, 3822–3831.

van Holde, K.E. (1988). Histone modifications. In *Chromatin*, Springer series in molecular biology, A. Rich, ed. (New York: Springer), pp. 111–148.

Wang, H., Huang, Z.Q., Xia, L., Feng, Q., Erdjument-Bromage, H., Strahl, B.D., Briggs, S.D., Allis, C.D., Wong, J., Tempst, P., and Zhang, Y. (2001). Methylation of histone H4 at arginine 3 facilitates transcriptional activation by nuclear hormone receptor. *Science* 293, 853–857.

Yang, L., Xia, L., Wu, D.Y., Wang, H., Chansky, H.A., Schubach, W.H., Hickstein, D.D., and Zhang, Y. (2002). Molecular cloning of ESET, a novel histone H3-specific methyltransferase that interacts with ERG transcription factor. *Oncogene* 21, 148–152.

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

Zhang, Y., LeRoy, G., Seelig, H.-P., Lane, W.S., and Reinberg, D. (1998). The dermatomyositis-specific autoantigen Mi2 is a component of a complex containing histone deacetylase and nucleosome remodeling activities. *Cell* 95, 279–289.

Accession Numbers

The GenBank accession numbers of the SET domain-containing H3-specific HMTase from human and mouse are AF448509 and AF448510, respectively.