Cell. Mol. Life Sci. 63 (2006) 2755–2763
1420-682X/06/232755-9
DOI 10.1007/s00018-006-6274-5

Cellular and Molecular Life Sciences

Review

SET domain protein lysine methyltransferases: Structure, specificity and catalysis

C. Qian and M.-M. Zhou*

Department of Molecular Physiology and Biophysics, Mount Sinai School of Medicine,
New York University, One Gustave L. Levy Place, New York, New York 10029 (USA), Fax: +1 212 849 2456,
e-mail: ming-ming.zhou@mssm.edu

Received 10 June 2006; accepted 22 August 2006
Online First 29 September 2006

Abstract. Site- and state-specific lysine methylation of histones is catalyzed by a family of proteins that contain the evolutionarily conserved SET domain and plays a fundamental role in epigenetic regulation of gene activation and silencing in all eukaryotes. The recently determined three-dimensional structures of the SET domains from chromosomal proteins reveal that the core SET domain structure contains a two-domain architecture, consisting of a conserved anti-parallel β-barrel and a structurally variable insert that surround a unusual knot-like structure that comprises the enzyme active site. These structures of the SET domains, either in the free state or when bound to cofactor *S*-adenosyl-L-homocysteine and/or histone peptide, mimicking an enzyme/cofactor/substrate complex, further yield the structural insights into the molecular basis of the substrate specificity, methylation multiplicity and the catalytic mechanism of histone lysine methylation.

Keywords. SET domain, histone lysine methylation, structure-function, chromatin modifications.

Introduction

Site-specific post-translational modifications of histones that include methylation, acetylation, phosphorylation and ubiquitination function individually or combinatorially as the fundamental epigenetic mechanism that governs all nuclear DNA-templated processes in eukaryotic organisms [1–3]. Histone lysine methylation, with the exception of histone H3 K79, has been shown to be catalyzed exclusively by the conserved SET domain family proteins originally identified in *Drosophila* suppressor of variegation [Su(var)3-9] [4], enhancer of zeste [E(z)] [5] and trithorax [6] – hence the name. Many SET domain lysine methyltransferases (Fig. 1) that target site-specific lysines in histones have been identified and characterized [7–12]. Lysine methylation is generally more complex than lysine acetylation, as a lysine can be subject to mono-, di- or tri-methylation. It has been hypothesized that acting in a combinatorial or sequential manner with other histone modifications on one or multiple histones, position- and state-specific lysine methylation of histones that is accomplished by a SET domain lysine methyltransferase in a specific biological context specifies unique functional consequences [13–16].

Histone lysine methylation plays a pivotal role in a wide array of cellular processes including heterochromatin formation, X-chromosome inactivation and transcription regulation [12]. Aberrant histone methylation has been linked to a number of developmental disorders and human disease [9, 17, 18]. For example, histone H3 K4 methylation is associated with transcriptionally active chromatin. H3-K4 di-methylation by Set1 correlates with basal transcription, whereas H3-K4 tri-methylation is observed at fully activated promoters [19]. In *Drosophila,* H3-K9 methylation by the SET domain protein Su(var)3-9 is required for the establishment and maintenance of

* Corresponding author.

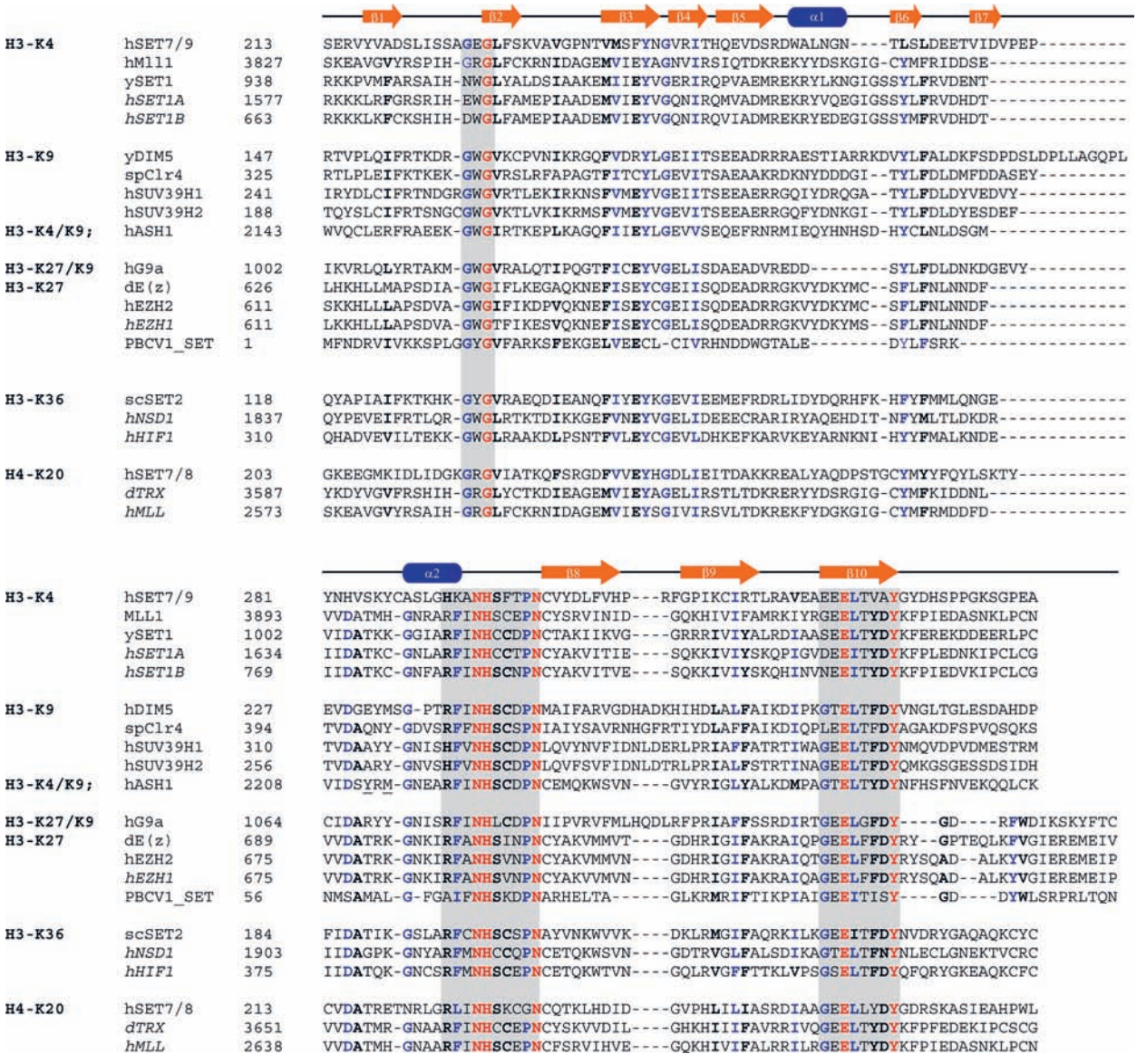

Figure 1. Sequence alignment of select members of the SET domain family. Absolutely conserved residues are shown in red, highly conserved residues black or blue, with blue for a higher degree of conservation. Three highly conserved motifs have been highlighted by light gray background.

heterochromatin [20]. In fission yeast, H3-K9 methylation by the lysine methyltransferase Clr4 has been linked to DNA methylation and to heterochromatin assembly by the RNAi effector complex RITS [21–23]. Moreover, di-methylation of H3-K9 by G9a is known to be associated with transcriptional silencing in euchromatin [24]. During cell differentiation, extended H3-K27 di- and tri-methylation carried out by the *Drosophila* polycomb group E(z)-Esc complex [25, 26] or by its mammalian Ezh-Eed counterpart [27, 28] marks for long-term gene silencing. Regional H3-K9 tri-methylation by Suv39h at transcriptionally inert chromatin is regarded as a hallmark of constitutive heterochromatin [15]. H4-K20 mono-methylation catalyzed by *Drosophila* PreSET7 (or SET8) is important for silencing general gene expression in mitosis. Finally, tri-methylation of H4 K20 by SUV4-20h is associated with transcriptionally inactive pericentric heterochromatin [29, 30].

Different methylation states of multiple histone lysines have been shown to have distinct biological distributions in chromatin. For instance, H3-K27 mono-methylation and H3-K9 tri-methylation are selectively enriched in pericentric heterochromatin, whereas H3-K27 tri-methylation along with H3-K9 di-methylation is considered epigenetic imprints of inactive X chromosome [23, 31]. Ezh2-containing complexes PRC2 and PRC3, selectively associated with distinct isoforms of Eed proteins, exhibit differential targeting of specific histones towards

H3-K27 or H1-K26 methylation, which has been shown to be important for transcriptional repression by Ezh2 [32, 33]. Moreover, studies in yeast have shown that trimethylation of H3-K4 by SET1 together with H3-K36 methylation by SET2 functionally mediate transcriptional elongation [34, 35]. The SET domain of SET1 closely resembles the mixed-lineage leukemia (MLL) SET domain in humans, which also processes methyltransferase activity toward H3-K4 that leads to transcriptional activation at Hox gene promoters [36].

Collectively, the high degree of modification complexity and coding potential of histone lysine methylation in epigenetic control of chromatin biology may explain the existence of an unusually large family of SET domain-containing proteins, which contain more than 700 members (greater than 100 members in humans) [37]. In addition to histones, recent studies show that SET domain proteins can catalyze lysine methylation of cellular proteins including cytochrome c, Rubisco [10], p53 and Taf10 [38, 39]. Thus, the SET domain proteins function generally as protein lysine methyltransferases. In this article, we review the current understanding of the structure, substrate specificity and catalysis of the SET domain lysine methyltransferases.

Overall structure

To date, more than ten SET domain structures, either in the free form or in complex with the methyl donor cofactor product *S*-adenosyl-L-homocysteine (SAH) and/or substrate peptides, have been solved, which include those of SET7/9, Dim-5, Clr4, SET8/PreSET7 and a viral SET domain protein (referred to as vSET) from paramecium bursaria chlorella virus 1 (PBCV-1) [7, 10, 11, 38–47]. These structures reveal that the conserved SET domain has a unique structural fold and is different from other classes of protein methyltransferases that use the cofactor *S*-adenosyl-L-methionine (SAM) as the methyl donor cofactor. The SET domain contains a series of β strands folding into three discrete sheets that surround a knot-like structure [48] (Fig. 2). This unusual pseudo-knot is formed by the C-terminal segment of the SET domain passing through a loop formed by a preceding stretch of the sequence. The C-terminal segment and the loop contain the two most-conserved sequence motifs in the SET domains consisting of ELxF/YDY and NHS/CxxPN (where × is any amino acid), respectively (Fig. 1).

Of the mammalian SET domain proteins that structures have been solved so far, the core SET domain is flanked by preSET (or nSET) and postSET (or cSET) sequences (Fig. 2). The preSET helps keep the structural stability by interacting with different surfaces of the core SET domain [7, 10]. The postSET domain forms part of the active site by providing an aromatic residue to pack against the core SET domain and thus construct a hydrophobic channel [42]. However, neither the preSET nor postSET sequence is conserved in SET domain lysine methyltransferases. For example, vSET from PBCV-1 does not contain a preSET domain and exists as a dimeric form in solution; The post-SET region of Dim-5 contains three cysteine residues that are essential for methylated activity by forming tetrahedral coordination with a zinc ion near the active site [42]. However, SET7/9, SET8 and Rubisco methyltransferase do not have this cysteine-rich region in the postSET domain. Another unusual feature of the SET

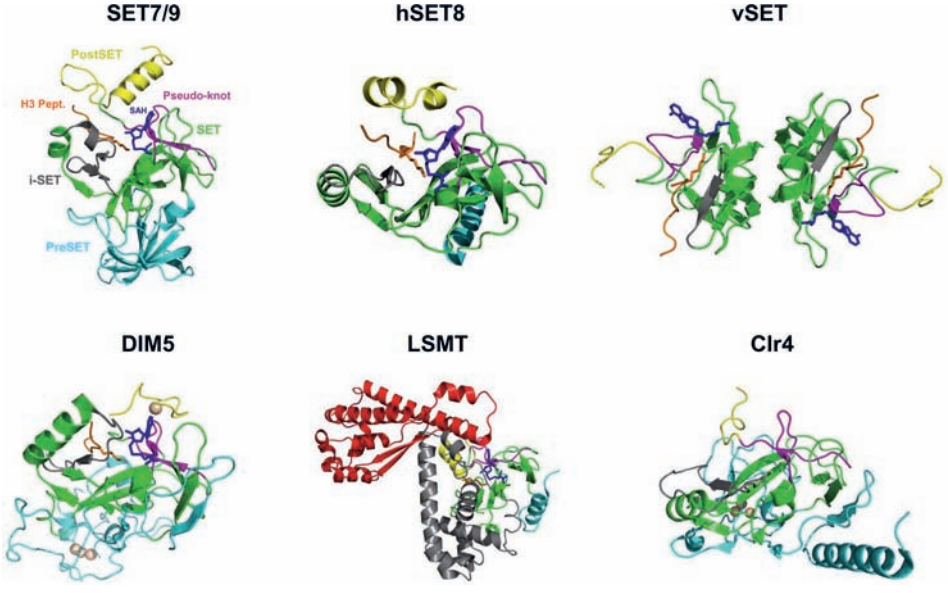

Figure 2. Three-dimensional structures of SET domains. In each proteins, PreSET, i-SET, SET and PostSET regions are depicted in cyan, light gray, green and yellow; the pseudo-knot, cofactor product SAH and substrate peptides are shown in magenta, blue and orange.

domain is an inserted region called i-SET, the sequence of which varies considerably in length, but the amino acid residues have been observed to interact with the substrate peptide in the three-dimensional structures of the different SET domains in ternary complex with substrate and cofactor.

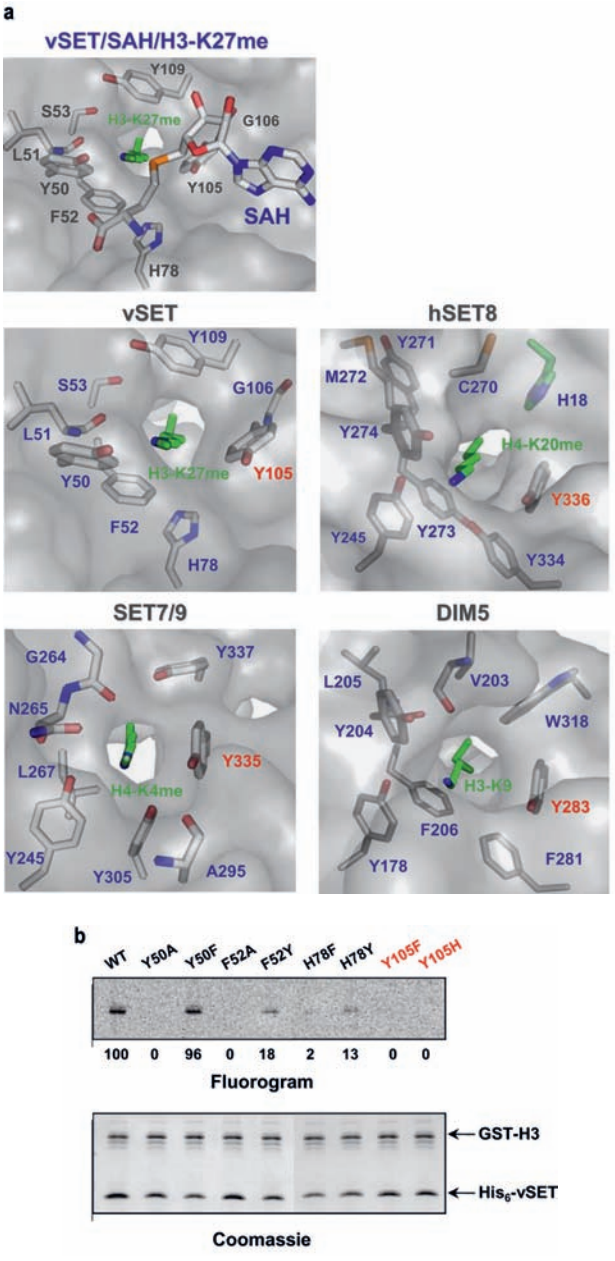

Figure 3. The molecular determinants of histone lysine methylation by SET domains. (*a*) The geometry and molecular environment of the lysine access channel in SET domains is shown for vSET, hSET8, SET7/9, and DIM-5. The SAH molecule is omitted in the latter three structures for clarity. (*b*) *In vitro* HKMT activity of vSET and its mutants measured using GST-fusion histone H3 N terminus (residues 1-57). Relative equal amounts of GST-H3 and vSET proteins used in the assay are shown in an SDS-PAGE gel (lower panel). Signals in the fluorogram represent HKMT activity of vSET for GST-H3 (upper panel). Phosphorescence values are normalized against wild-type vSET signal (lane 1).

Enzyme active site

A narrow hydrophobic channel that links the substrate lysine or methylated lysine and cofactor product SAH has been observed in all SET domain complex structures. The cofactor and substrate bind in distinct clefts on the opposite surfaces of the SET domain, while methylated lysine is buried in the channel at the enzyme active site formed by several conserved hydrophobic residues (Fig. 3a). The cofactor SAH is bound at the other end of this lysine access channel with the sulfur atom positioned at the tip of the methyl group of the mono-methylated lysine. In the ternary complex of vSET bound to SAH and a mono-methylated K27 histone H3 peptide, Tyr 50, Leu 51, Phe 52 and Ser 53 of β6, His 78 of β8, and Tyr 105 and Tyr 109 of the C-terminal sequence form the wall of the channel in the binding pocket [47]. Direct participation of the C-terminal residue Tyr 109 in the formation of the lysine access channel and its interactions with SAH may explain how the C-terminal sequence becomes more structurally ordered upon the ternary complex formation [47]. Similar dynamics have also been seen in SET7/9 [45] by the observation that its C-flanking domain becomes more structured and completes the lysine access channel upon the binding of the cofactor and substrate. It is also worth noting that Tyr 109 in vSET is reminiscent of Trp 318 in Dim-5 or Tyr 337 in SET7/9 [8, 43]. The latter two residues are directly involved in completeness of the lysine access channel through structural arrangement upon substrate binding. Interesting, in the SET8 SET domain complex structure, a substrate residue His 18 (-2) in the histone H4 Lys 20-containing peptide appears to fulfill such a role in the formation of the substrate lysine access channel (Fig. 3a, upper left panel) [46].

The cofactor SAH bound to the SET domain adapts a U-shaped conformation by forming hydrogen bonds with main-chain atoms of protein residues in the conserved NHS/CxxPN motif in the pseudo-knot and GxG motif in the N-terminal region. Notably, this compact conformation of the bound SAH and the geometry of its binding cleft are highly conserved in SET domain lysine methyltransferase, but distinct from those of protein arginine methyltransferases, which also use SAM as methyl donor. The C-terminal region of the SET domain also contributes to the binding cleft of the bound SAH. For instance, in vSET, the indole ring of the C-terminal Trp 110 is positioned in close proximity to SAH, whereas in SET8 the indole ring of Trp 349 (corresponding to Trp 352 in SET7/9) packs against the adenine base of SAH [8, 45–47].

Substrate specificity

Histone lysine methylation is a functionally complex process, as it can either activate or repress transcription, depending on sequence-specific lysine methylation site in histones and the methylation state of the ε-amino group of the target lysine. The high degree of modification complexity and coding potential of histone lysine methylation in epigenetic control of chromatin biology may explain the existence of an unusually large family of SET domain-containing proteins [37]. Consistent with their high substrate specificity, SET domain lysine methyltransferases show overall low sequence similarity and the residues at the active site are not all conserved. Structural details of the available SET domains in complex with cofactor and histone peptide substrate provide insight into the substrate specificity and the general mechanism of catalysis.

The network of polar interactions between the peptide substrate and a SET domain lysine methyltransferase makes a major contribution to the substrate specificity. The residues in the i-SET region are directly involved in interactions with the peptide substrate, as described above. Because of its low sequence similarity in SET domain lysine methyltransferases that exhibit different substrate specificities, this i-SET region is likely a determinant for target sequence discrimination. However, it is not fully understood how some SET domain proteins such as SET7/9 and MLL1 have identical substrate specificities but very different i-SET sequences.

For a specific protein lysine methyltransferase, the site specificity of lysine methylation is probably determined by amino acid residues flanking the target lysine in the substrate. This suggests that a particular SET domain lysine methyltransferase likely achieves its biological substrates by recognizing a consensus motif in the biological targets. For instance, human SET7/9 methylates histone H3 at lysine 4, as well as Taf10 and p53 peptides, and all three substrates share the conserved K/R-S/T-K motif preceding the substrate lysine [38, 49]. Histone K20-specific methyltransferase hSET8 recognizes the motif R-Ω-ξ-K-x-φ (where Ω is an aromatic, ξ is a non-acidic, x is any, and φ is a bulky hydrophobic amino acid) [46]. Interestingly, the sequence of histone H3 at Lys 27 and Lys 9 sites has similar residues flanking the corresponding lysine. vSET domain selectively targets at histone H3 Lys 27 but not Lys 9, which, as revealed by biochemical and structural studies, is due to its recognition of the motif APA that is present at the H3-K27 but not the H3-Lys 9 site [11, 47].

Methylation multiplicity

The geometry, shape, and type of amino acids that comprise the substrate lysine access channel in the active site have been suggested to confer methylation multiplicity in the SET domain. Figure 3a shows the lysine access channels of the four ternary complex structures. The walls of lysine access channel are mainly formed by hydrophobic residues. These residues are not so conserved among the SET domains and may be responsible for methylation multiplicity because the variability of amino acids will alter the diameter and the shape of the channel. Mutation of Tyr 50 or Phe 52 to an Ala caused a nearly complete loss of methylation activity of vSET domain, whereas their more conservative amino acid mutations showed that Y50F became largely a mono-methylase for H3-K27, and that F52Y can only mono- and di- but not tri-methylate H3-K27 [47]. Mutation of His 78 to a Phe or Tyr resulted in a nearly complete loss of the H3-K27 methylation activity except for a little mono-methylation activity (Fig. 3b) [47]. The mutation effects of His 78 are consistent with the reported mutational results of the corresponding Phe 281 in Dim-5, which showed that F281Y exhibited only mono- and di- but not tri-methylation of the target H3-K27. Similarly, mutations of Tyr 305 to Phe in SET7/9 or Tyr 334 to Phe in SET8, which are positioned similarly to that of His 78 in vSET or Phe 281 in Dim-5, have been shown to alter methylation multiplicity from sole mono-methylation to mono- and di-methylation [8, 43]. Taken together, these observations suggest that several residues forming the substrate-specific channel play an important role in determining the methylation multiplicity of the SET domain by defining the geometry and size of the narrow lysine access channel and by maintaining a network of molecular interactions between the enzyme, cofactor and the substrate lysine.

Catalytic mechanism

Several possible mechanisms have been discussed in the literature, but there has been no consensus in the understanding of the catalytic mechanism. The network of molecular interactions revealed in the structures of the ternary complexes of several SET domains supports a notion that methyl transfer from SAM to the ε-amino of the target lysine likely proceeds by a direct in-line S_N2 nucleophilic attack, and this mechanism has been supported by a recently theoretical study on the SET7/9 [8, 50]. While it is believed that only a deprotonated substrate lysine bearing a free lone pair of electrons is capable of the nucleophilic attack on the SAM methyl group, the first mechanism is related to the hydrogen bonds formed in the SET domain active site [51]. The hydrogen bond between the δ-methyl group carbon and several main chain carbonyl groups and hydroxyl of the invariant tyrosine residue can increase the electrophilic character of δ-methyl group of SAM and facilitate its

transfer. However, the carbon-oxygen hydrogen bond is not present in some SET domains because the residues around the active site are not absolutely conserved. The second mechanism is of a general base catalysis, which predicts an active site residue deprotonates the substrate lysine before methyl transfer; the supporting evidence for this mechanism is not conclusive. For instance, the only invariant residue Tyr 335 in SET7/9 or Tyr 283 in Dim-5 SET domain that could act as a general base for lysine deprotonation is separated from the ε-aminonium group of the target lysine by greater than 3–4 Å in the crystal structures, indicating a lacking of hydrogen bonding between them [43, 52]. In vSET, the pH dependence of the H3-K27 methylation activity of the wild-type vSET exhibited an apparent pKa of pH 9.0; mutation of the invariant tyrosine residue Tyr 105 into Phe or His exhibited little activity at pH 8.0 or above pH 9.0, arguing that this invariant Tyr 105 likely does not act as a general base that deprotonates substrate lysine [47]. Thus, the general base mechanism is less likely, and Tyr 105 may instead function to provide a charge-dipole interaction between its phenoxyl oxygen and the positively charged sulfur of SAM, or may interact with and align the amino nitrogen of the substrate lysine to the methyl carbon of SAM to facility the S_N2 methyl transfer. Such a mechanism, which needs to be further validated experimentally, is reminiscent of what has been shown for glycine *N*-methyltransferase [53]. One feature of the SET domain structure that is consistent with this proposed catalytic mechanism is that the target lysine can lose a proton to the solvent as the C-flanking sequence of the SET domain is flexible and can expose the target lysine to the bulky solvent [45].

Prospective

Since the identification of the first histone lysine methyltransferase Su(var)3-9 [11], tremendous progress has been made on our understanding of the three-dimensional structure, substrate specificity and catalytic mechanism of SET domain lysine methyltransferases, which has facilitated our characterization of the role of histone lysine methylation in histone-directed chromatin biology (Fig. 4). However, because of the modification complexity and its extraordinarily high degree of coding potential (in combination with other histone modifications), major challenges lie ahead to fully understand the chromatin regulatory capability of histone lysine methylation in directing targeted gene silencing and 'on demand' expression in response to physiological and environmental stimuli. In the near future, we expect that the chromatin biology research will address several key outstanding questions in the epigenetic regulation as described below.

First, more novel histone lysine methyltransferases with distinct cellular functions may be identified. Several proteins that contain the conserved PR domain, such as Meisetz and RIZ, have been identified to show histone lysine methyltransferase activity [54]. Note that PR domain shares only about 20% sequence identity with the SET domain. Detailed structural analysis of PR domains will be needed to understand their molecular function and catalytic mechanism in histone lysine methylation.

Second, several groups of conserved protein modules that belong to the super royal family [55] of protein modules have been characterized as methyl lysine recognition domains. These include the chromo [56–58], Tudor [59, 60], MBT and WD repeat [61] domains, which have been

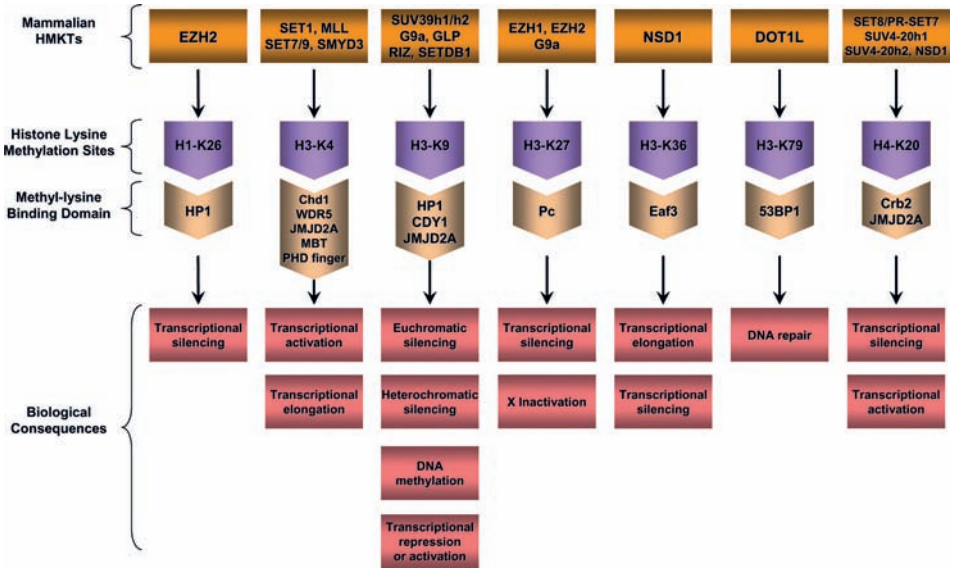

Figure 4. Histone lysine methyltransferases, target lysine methylation sites in histones, methyl-lysine binding protein modular domains and the related biological functions.

shown to read distinct methyl signals and subsequently direct different downstream effector proteins in histone signaling. Because different histone modifications can in principle work in a combinational fashion, it will not be surprising to see that some of these proteins or protein domains may function in combinational recognition of distinct patterns of histone modifications.

Third, as SET domain lysine methyltransferases are often functionally associated with other chromosomal proteins in histone methylation. It is conceivable that some of these cellular proteins will be identified as substrates of SET domain lysine methyltransferases. As such, we predict that lysine methylation serves a general mechanism that regulates gene transcription through site- and state-specific lysine methylation of not only histones but also transcriptional proteins.

Fourth, in contrast to what was thought previously, recent identification of histone lysine demethylases confirms that lysine methylation is a reversible modification. Similar to SET domain lysine methyltransferases, substrate specificity of the first identified histone lysine demethylase LSD1, which can demethylate mono- or di-methylated but not tri-methylated lysine K4 of H3 [62], is tightly regulated by its associated protein cofactors androgen receptor or CoREST [63–65]. JmjC-containing proteins represent another class of histone lysine demethylases. JHDM1 can specifically demethylate mono- or di-methylated H3 K36 [66], whereas JHDM2 targets mono- or di-methylated H3 K9 [67], and functions as a tri-methylation demethylase [68]. If methylation at every known lysine sites on histone H3 is reversible, we would expect more histone demethylases to be discovered. At the same time, we also anticipate that these lysine demethylases may target cellular proteins, as shown for the SET domain lysine methyltransferases.

Lastly, growing evidence suggests that histone lysine methylation plays an important role in the development of various forms of human disease including cancer [9, 17, 18]. For example, suppression of SMYD3 protein expression significantly inhibits the growth of colorectal and hepatocellular carcinoma cells [9]. Future development of small-molecule chemicals that are capable of inhibiting or modulating the SET domain enzymatic activity could lead to novel anticancer therapies.

Acknowledgement. This work was supported by grants from the National Institutes of Health to M.-M.Z.

1 Jaskelioff, M. and Peterson, C. L. (2003) Chromatin and transcription: histones continue to make their marks. Nat. Cell Biol. 5, 395–399.
2 Fischle, W., Wang, Y. and Allis, C. D. (2003) Histone and chromatin cross-talk. Curr. Opin. Cell Biol. 15, 172–183.
3 Margueron, R., Trojer, P. and Reinberg, D. (2005) The key to development: interpreting the histone code? Curr. Opin. Genet. Dev. 15, 163–176.
4 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.
5 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.
6 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.
7 Wilson, J., Jing, C., Walker, P., Martin, S., Howell, S., Blackburn, G., Gamblin, S. and Xiao, B. (2002) Crystal structure and functional analysis of the histone methyltransferase SET7/9. Cell 111, 105–115.
8 Xiao, B., Jing, C., Wilson, J. R., Walker, P. A., Vasisht, N., Kelly, G., Howell, S., Taylor, I. A., Blackburn, G. M. and Gamblin, S. J. (2003) Structure and catalytic mechanism of the human histone methyltransferase SET7/9. Nature 421, 652–656.
9 Hamamoto, R., Furukawa, Y., Morita, M., Iimura, Y., Silva, F. P., Li, M., Yagyu, R. and Nakamura, Y. (2004) SMYD3 encodes a histone methyltransferase involved in the proliferation of cancer cells. Nat. Cell Biol. 6, 731–740.
10 Trievel, R. C., Beach, B. M., Dirk, L. M., Houtz, R. L. and Hurley, J. H. (2002) Structure and catalytic mechanism of a SET domain protein methyltransferase. Cell 111, 91–103.
11 Manzur, K. L., Farooq, A., Zeng, L., Plotnikova, O., Koch, A. W., Sachchidanand and Zhou, M. M. (2003) A dimeric viral SET domain methyltransferase specific to Lys27 of histone H3. Nat. Struct. Biol. 10, 187–196.
12 Martin, C. and Zhang, Y. (2005) The diverse functions of histone lysine methylation. Nat. Rev. Mol. Cell. Biol. 6, 838–849.
13 Strahl, B. D. and Allis, C. D. (2000) The language of covalent histone modifications. Nature 403, 41–45.
14 Turner, B. M. (2002) Cellular memory and the histone code. Cell 111, 285–291.
15 Lachner, M., O'Sullivan, R. J. and Jenuwein, T. (2003) An epigenetic road map for histone lysine methylation. J. Cell Sci. 116, 2117–2124.
16 Bannister, A. J., Schneider, R. and Kouzarides, T. (2002) Histone methylation: dynamic or static? Cell 109, 801–806.
17 Fraga, M. F., Ballestar, E., Villar-Garea, A., Boix-Chornet, M., Espada, J., Schotta, G., Bonaldi, T., Haydon, C., Ropero, S., Petrie, K., Iyer, N. G., Perez-Rosado, A., Calvo, E., Lopez, J. A., Cano, A., Calasanz, M. J., Colomer, D., Piris, M. A., Ahn, N., Imhof, A., Caldas, C., Jenuwein, T. and Esteller, M. (2005) Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer. Nat. Genet. 37, 391–400.
18 Hake, S. B., Xiao, A. and Allis, C. D. (2004) Linking the epigenetic 'language' of covalent histone modifications to cancer. Br. J. Cancer 90, 761–769.
19 Bernstein, B. E., Humphrey, E. L., Erlich, R. L., Schneider, R., Bouman, P., Liu, J. S., Kouzarides, T. and Schreiber, S. L. (2002) Methylation of histone H3 Lys 4 in coding regions of active genes. Proc. Natl. Acad. Sci. USA 99, 8695–8700.
20 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., Jenuwein, T. (2000) Regulation of chromatin structure by site-specific histone H3 methyltransferases. Nature 406, 593–599.
21 Volpe, T. A., Kidner, C., Hall, I. M., Teng, G., Grewal, S. I. and Martienssen, R. A. (2002) Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. Science 297, 1833–1837.
22 Freitag, M. and Selker, E. U. (2005) Controlling DNA methylation: many roads to one modification. Curr. Opin. Genet. Dev. 15, 191–199.

23 Rice, J. C., Briggs, S. D., Ueberheide, B., Barber, C. M., Shabanowitz, J., Hunt, D. F., Shinkai, Y. and Allis, C. D. (2003) Histone methyltransferases direct different degrees of methylation to define distinct chromatin domains. Mol. Cell 12, 1591–1598.
24 Tachibana, M., Sugimoto, K., Nozaki, M., Ueda, J., Ohta, T., Ohki, M., Fukuda, M., Takeda, N., Niida, H., Kato, H. and Shinkai, Y. (2002) G9a histone methyltransferase plays a dominant role in euchromatic histone H3 lysine 9 methylation and is essential for early embryogenesis. Genes Dev. 16, 1779–1791.
25 Czermin, B., Melfi, R., McCabe, D., Seitz, V., Imhof, A. and Pirrotta, V. (2002) Drosophila enhancer of Zeste/ESC complexes have a histone H3 methyltransferase activity that marks chromosomal polycomb sites. Cell 111, 185–196.
26 Muller, J., Hart, C. M., Francis, N. J., Vargas, M. L., Sengupta, A., Wild, B., Miller, E. L., O'Connor, M. B., Kingston, R. E. and Simon, J. A. (2002) Histone methyltransferase activity of a Drosophila polycomb group repressor complex. Cell 111, 197–208.
27 Cao, R., Wang, L., Wang, H., Xia, L., Erdjument-Bromage, H., Tempst, P., Jones, R. S. and Zhang, Y. (2002) Role of histone H3 lysine 27 methylation in polycomb-group silencing. Science 298, 1039–1043.
28 Kuzmichev, A., Nishioka, K., Erdjument-Bromage, H., Tempst, P. and Reinberg, D. (2002) Histone methyltransferase activity associated with a human multiprotein complex containing the enhancer of zeste protein. Genes Dev. 16, 2893–2905.
29 Rice, J. C., Nishioka, K., Sarma, K., Steward, R., Reinberg, D. and Allis, C. D. (2002) Mitotic-specific methylation of histone H4 Lys 20 follows increased PR-Set7 expression and its localization to mitotic chromosomes. Genes Dev. 16, 2225–2230.
30 Schotta, G., Lachner, M., Sarma, K., Ebert, A., Sengupta, R., Reuter, G., Reinberg, D. and Jenuwein, T. (2004) A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin. Genes Dev. 18, 1251–1262.
31 Peters, A. H., Kubicek, S., Mechtler, K., O'Sullivan, R. J., Derijck, A. A., Perez-Burgos, L., Kohlmaier, A., Opravil, S., Tachibana, M., Shinkai, Y., Martens, J. H. and Jenuwein, T. (2003) Partitioning and plasticity of repressive histone methylation states in mammalian chromatin. Mol. Cell 12, 1577–1589.
32 Kuzmichev, A., Jenuwein, T., Tempst, P. and Reinberg, D. (2004) Different EZH2-containing complexes target methylation of histone H1 or nucleosomal histone H3. Mol. Cell 14, 183–193.
33 Daujat, S., Zeissler, U., Waldmann, T., Happel, N. and Schneider, R. (2005) HP1 binds specifically to Lys26-methylated histone H1. 4, whereas simultaneous Ser27 phosphorylation blocks HP1 binding. J. Biol. Chem. 280, 38090–38095.
34 Ng, H. H., Robert, F., Young, R. A. and Struhl, K. (2003) Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. Mol. Cell 11, 709–719.
35 Krogan, N. J., Kim, M., Tong, A., Golshani, A., Cagney, G., Canadien, V., Richards, D. P., Beattie, B. K., Emili, A., Boone, C., Shilatifard, A., Buratowski, S. and Greenblatt, J. (2003) Methylation of histone H3 by Set2 in *Saccharomyces cerevisiae* is linked to transcriptional elongation by RNA polymerase II. Mol. Cell. Biol. 23, 4207–4218.
36 Milne, T. A., Briggs, S. D., Brock, H. W., Martin, M. E., Gibbs, D., Allis, C. D. and Hess, J. L. (2002) MLL targets SET domain methyltransferase activity to Hox gene promoters. Mol. Cell 10, 1107–1117.
37 Schultz, J., Milpetz, F., Bork, P. and Ponting, C. P. (1998) SMART, a simple modular architecture research tool: identification of signaling domains. Proc. Natl. Acad. Sci. USA 95, 5857–5864.
38 Chuikov, S., Kurash, J. K., Wilson, J. R., Xiao, B., Justin, N., Ivanov, G. S., McKinney, K., Tempst, P., Prives, C., Gamblin, S. J., Barlev, N. A. and Reinberg, D. (2004) Regulation of p53 activity through lysine methylation. Nature 432, 353–360.
39 Kouskouti, A., Scheer, E., Staub, A., Tora, L. and Talianidis, I. (2004) Gene-specific modulation of TAF10 function by SET9-mediated methylation. Mol. Cell 14, 175–182.
40 Kwon, T., Chang, J. H., Kwak, E., Lee, C. W., Joachimiak, A., Kim, Y. C., Lee, J. and Cho, Y. (2003) Mechanism of histone lysine methyl transfer revealed by the structure of SET7/9-AdoMet. EMBO J. 22, 292–303.
41 Min, J., Zhang, X., Cheng, X., Grewal, S. I. and Xu, R. M. (2002) Structure of the SET domain histone lysine methyltransferase Clr4. Nat. Struct. Biol. 9, 828–832.
42 Zhang, X., Tamaru, H., Khan, S., Horton, J., Keefe, L., Selker, E. and Cheng, X. (2002) Structure of the neurospora SET domain protein DIM-5, a histone H3 lysine methyltransferase. Cell 111, 117–127.
43 Zhang, X., Yang, Z., Khan, S. I., Horton, J. R., Tamaru, H., Selker, E. U. and Cheng, X. (2003) Structural basis for the product specificity of histone lysine methyltransferases. Mol. Cell 12, 177–185.
44 Jacobs, S. A., Harp, J. M., Devarakonda, S., Kim, Y., Rastinejad, F. and Khorasanizadeh, S. (2002) The active site of the SET domain is constructed on a knot. Nat. Struct. Biol. 9, 833–838.
45 Xiao, B., Jing, C., Kelly, G., Walker, P. A., Muskett, F. W., Frenkiel, T. A., Martin, S. R., Sarma, K., Reinberg, D., Gamblin, S. J. and Wilson, J. R. (2005) Specificity and mechanism of the histone methyltransferase Pr-Set7. Genes Dev. 19, 1444–1454.
46 Couture, J. F., Collazo, E., Brunzelle, J. S. and Trievel, R. C. (2005) Structural and functional analysis of SET8, a histone H4 Lys-20 methyltransferase. Genes Dev. 19, 1455–1465.
47 Qian, C., Wang, X., Manzur, K., Sachchidanand, Farooq, A., Zeng, L., Wang, R. and Zhou, M. M. (2006) Structural insights of the specificity and catalysis of a viral histone H3 lysine 27 methyltransferase. J. Mol. Biol. 359, 86–96.
48 Taylor, W. R., Xiao, B., Gamblin, S. J. and Lin, K. (2003) A knot or not a knot? SETting the record 'straight' on proteins. Comput. Biol. Chem. 27, 11–15.
49 Couture, J. F., Collazo, E., Hauk, G. and Trievel, R. C. (2006) Structural basis for the methylation site specificity of SET7/9. Nat. Struct. Mol. Biol. 13, 140–146.
50 Hu, P. and Zhang, Y. (2006) Catalytic mechanism and product specificity of the histone lysine methyltransferase SET7/9: an *ab initio* QM/MM-FE study with multiple initial structures. J. Am. Chem. Soc. 128, 1272–1278.
51 Trievel, R. C., Flynn, E. M., Houtz, R. L. and Hurley, J. H. (2003) Mechanism of multiple lysine methylation by the SET domain enzyme Rubisco LSMT. Nat. Struct. Biol. 10, 545–552.
52 Xiao, B., Wilson, J. R. and Gamblin, S. J. (2003) SET domains and histone methylation. Curr. Opin. Struct. Biol. 13, 699–705.
53 Takata, Y., Huang, Y., Komoto, J., Yamada, T., Konishi, K., Ogawa, H., Gomi, T., Fujioka, M. and Takusagawa, F. (2003) Catalytic mechanism of glycine N-methyltransferase. Biochemistry 42, 8394–8402.
54 Hayashi, K., Yoshida, K. and Matsui, Y. (2005) A histone H3 methyltransferase controls epigenetic events required for meiotic prophase. Nature 438, 374–378.
55 Maurer-Stroh, S., Dickens, N. J., Hughes-Davies, L., Kouzarides, T., Eisenhaber, F. and Ponting, C. P. (2003) The Tudor domain 'Royal Family': Tudor, plant Agenet, Chromo, PWWP and MBT domains. Trends Biochem. Sci. 28, 69–74.
56 Cao, R., Wang, L., Wang, H., Xia, L., Erdjument-Bromage, H., Tempst, P., Jones, R. S. and Zhang, Y. (2002) Role of histone H3 lysine 27 methylation in polycomb-group silencing. Science 298, 1039–1043.
57 Min, J., Zhang, Y. and Xu, R. M. (2003) Structural basis for specific binding of polycomb chromodomain to histone H3 methylated at Lys 27. Genes Dev. 17, 1823–1828.

58 Pray-Grant, M. G., Daniel, J. A., Schieltz, D., Yates, J. R., 3rd and Grant, P. A. (2005) Chd1 chromodomain links histone H3 methylation with SAGA- and SLIK-dependent acetylation. Nature 433, 434–438.
59 Brahms, H., Meheus, L., de Brabandere, V., Fischer, U. and Luhrmann, R. (2001) Symmetrical dimethylation of arginine residues in spliceosomal Sm protein B/B′ and the Sm-like protein LSm4, and their interaction with the SMN protein. RNA 7, 1531–1542.
60 Friesen, W. J., Massenet, S., Paushkin, S., Wyce, A. and Dreyfuss, G. (2001) SMN, the product of the spinal muscular atrophy gene, binds preferentially to dimethylarginine-containing protein targets. Mol. Cell 7, 1111–1117.
61 Hennig, L., Bouveret, R. and Gruissem, W. (2005) MSI1-like proteins: an escort service for chromatin assembly and remodeling complexes. Trends Cell Biol. 15, 295–302.
62 Shi, Y., Lan, F., Matson, C., Mulligan, P., Whetstine, J. R., Cole, P. A., Casero, R. A. and Shi, Y. (2004) Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. Cell 119, 941–953.
63 Metzger, E., Wissmann, M., Yin, N., Muller, J. M., Schneider, R., Peters, A. H., Gunther, T., Buettner, R. and Schule, R. (2005) LSD1 demethylates repressive histone marks to promote androgen-receptor-dependent transcription. Nature 437, 436–439.
64 Lee, M. G., Wynder, C., Cooch, N. and Shiekhattar, R. (2005) An essential role for CoREST in nucleosomal histone 3 lysine 4 demethylation. Nature 437, 432–435.
65 Shi, Y. J., Matson, C., Lan, F., Iwase, S., Baba, T. and Shi, Y. (2005) Regulation of LSD1 histone demethylase activity by its associated factors. Mol. Cell 19, 857–864.
66 Tsukada, Y., Fang, J., Erdjument-Bromage, H., Warren, M. E., Borchers, C. H., Tempst, P. and Zhang, Y. (2006) Histone demethylation by a family of JmjC domain-containing proteins. Nature 439, 811–816.
67 Yamane, K., Toumazou, C., Tsukada, Y. I., Erdjument-Bromage, H., Tempst, P., Wong, J. and Zhang, Y. (2006) JHDM2A, a JmjC-containing H3K9 demethylase, facilitates transcription activation by androgen receptor. Cell 125, 483–495.
68 Whetstine, J. R., Nottke, A., Lan, F., Huarte, M., Smolikov, S., Chen, Z., Spooner, E., Li, E., Zhang, G., Colaiacovo, M. and Shi, Y. (2006) Reversal of histone lysine trimethylation by the JMJD2 family of histone demethylases. Cell 125, 467–481.