Clin Epigenet (2011) 2:17–25
DOI 10.1007/s13148-010-0015-1

CARCINOGENESIS

Interaction of SET domains with histones and nucleic acid structures in active chromatin

Wladyslaw A. Krajewski · Oleg L. Vassiliev

Received: 15 July 2010 / Accepted: 16 November 2010 / Published online: 14 January 2011

Abstract Changes in the normal program of gene expression are the basis for a number of human diseases. Epigenetic control of gene expression is programmed by chromatin modifications—the inheritable "histone code"—the major component of which is histone methylation. This chromatin methylation code of gene activity is created upon cell differentiation and is further controlled by the "SET" (methyltransferase) domain proteins which maintain this histone methylation pattern and preserve it through rounds of cell division. The molecular principles of epigenetic gene maintenance are essential for proper treatment and prevention of disorders and their complications. However, the principles of epigenetic gene programming are not resolved. Here we discuss some evidence of how the SET proteins determine the required states of target genes and maintain the required levels of their activity. We suggest that, along with other recognition pathways, SET domains can directly recognize the nucleosome and nucleic acids intermediates that are specific for active chromatin regions.

Keywords Chromatin · Gene expression · Histone methylation

W. A. Krajewski (✉)
Institute of Developmental Biology,
Russian Academy of Sciences,
ul. Vavilova 26,
Moscow 119334, Russia
e-mail: wkrajewski@hotmail.com

O. L. Vassiliev
Shemyakin Institute of Bioorganic Chemistry,
Russian Academy of Sciences,
Ul. Miklukho-Maklaya, 16/10,
Moscow 437, Russia

The accumulated data suggest that many diseases and metabolic disorders are caused by altered patterns of gene expression (Kaminsky et al. 2006; Perini and Tupler 2006; Maekawa and Watanabe 2007). The chromatin-templated processes are controlled by a complex pattern of posttranslational modifications of the flexible N-terminal tails of histone proteins, including methylation, acetylation, phosphorylation, ubiquitination, etc., which comprise the inheritable "histone code" of gene function (Marmorstein and Trievel 2009), although the effects of the histone code may depend on the particular situation (Lee et al. 2010). Chromatin activity is largely determined by the methylation status of specific lysine and arginine residues in the N termini of histones H3 and H4 (Li et al. 2007; Shilatifard 2008). For example, methylation of K4, 36, and 79 in histone H3 promotes gene activation, while methylation of H3 K9, 27, and H4 lys20 is associated with gene silencing. The knowledge of epigenetic gene regulatory principles is essential for developing targeted therapies.

With the exception of Dot1 which methylates H3 K79 (van Leeuwen et al. 2002), the methylation of histone lysines for chromatin activity and silencing is conferred, respectively, by the Trithorax and Polycomb group histone methyltransferases, containing a conserved 130-amino-acid catalytic "SET" domain [*S*u(var), *E*nhancer of zeste, *T*rithorax] (Qian and Zhou 2006). The SET domain is a paradigm for both positive and negative regulators of chromatin activity. Less well-conserved pre- and post-SET sequences may flank SET domains at N and C boundaries, respectively. The SET domain proteins assume full control on the maintenance of chromatin lysine methylation through rounds of cell divisions, after the gene activity patterns have been established in early embryogenesis by a cascade of maternal and zygotic transcription factors (Breiling et al. 2007). However, it is not fully understood

how SET domain proteins can "decide" on the required states of gene activity in a tissue-specific manner, how SET proteins can initially recognize their target genes, and the way this "recognition" is transmitted to progeny cells.

The molecular mechanisms behind SET-domain recognition of chromatin activity states are essential for understanding the molecular etiology of epigenetic-related disorders. The best studied example is cancer (Albert and Helin 2010), although a number of other diseases also result from aberrant chromatin methylation patterns. For example, it has been shown that correct functioning of islet beta cell depends on the methyltransferase Set7/9 which is implicated in maintaining the active chromatin status of genes required for glucose-stimulated secretion of insulin (Deering et al. 2009). An aberrant histone methylation pattern may be a major underlying mechanism for sustained proinflammatory phenotype of diabetic cells. The SET7/9 protein recognizes and preserves the correct methylation states of K4 in histone H3 in chromatin of NF-kappa B-dependent inflammatory genes, the correct functioning of which is critical for preventing the progression of diabetes and the metabolic syndrome (Li et al. 2008). Trimethylation of histone H3-K9 by Suv39h1 is essential in preventing the pre-activated state of diabetic vascular smooth muscle cells (Villeneuve et al. 2008). Methylation of histone H3 controls the expression of insulin (Cavener 2009). The tissue-specific heritable aberrations in histone methylation pattern could be a reason for the propagation of the trans-generational insulin-resistant phenotype in gestational diabetes (Devaskar and Thamotharan 2007). The Ezh2 methyltransferase is implicated in the maintenance of normal pancreatic beta cell proliferation, likely by preserving histone H3 trimethylation at the Ink4a/Arf locus in islet beta cells, thus preventing beta cell regenerative failure and diabetes (Chen et al. 2009; Villeneuve et al. 2008). In addition, the Ezh2 protein is also likely to be involved in formation of cancerous tissues, including prostate and breast cancers (Simon and Lange 2008). Deregulation by MLL1, the human homolog of *Drosophila* Trithorax, results in lymphoid and myeloid acute leukemias (Cosgrove and Patel 2010). Epigenetic alterations are also implicated in the development of cardiac hypertrophy, ischemia (Maekawa and Watanabe 2007; Granger et al. 2008; Kaneda et al. 2009), rheumatic arthritis (Strietholt et al. 2008), autoimmune disease (Szyf 2010), asthma (Schwartz 2010), and other diseases (Perini and Tupler 2006; Maekawa and Watanabe 2007). There are many of such studies, but their practical implications are still limited by insufficient understanding of the principles of how the SET-domain proteins recognize, maintain, and propagate the states of chromatin activity to descendant cells.

Most of the SET-domain methyltransferases (HKMTs) can mono-, di-, and trimethylate one specific lysine residue in core histones. For example, Set1, Trithorax and MLL1/4 methylate H3-K4, Su(var)3–9 and Suv39h1 methylate H3-K9, E(z) and Ezh2 methylate H3-K27, Set2 and HPBD methylate H3-K36, PR-Set7/8, and Suv4h20 methylate H4-K20. Some proteins methylate multiple sites. For example, Ash1 can methylate H3-K4, -9, and H4-K20 (Gregory et al. 2007), NSD1-3 can methylate H3-K36 and H4-K44 (Li et al. 2009). HKMTs often function within multiprotein complexes. For example, Trithorax functions within the acetylation complex, TAC1 (Petruk et al. 2001) and Polycomb Repressive complexes type 2 (PRC2) are based around the Ezh2 [the E(z) in *Drosophila*] methyltransferases (Schwartz and Pirrotta 2007). Isolated HKMT complexes may include tens of different transcription factors, like the MLL1 complex, reflecting the complexity of the HKMT interactions in vivo (Nakamura et al. 2002), although only a few core components may be critical for HKMT functional activity (like the Wdr5, RbPB5, Ash2, and Menin subunits of MLL1; Dou et al. 2006).

Data in the literature suggest several, not mutually exclusive, possibilities of how SET proteins may recognize their chromatin targets.

(a) SET-domain proteins can recognize their target genes through direct or indirect interaction with site-specific chromatin-binding factors, including factors recognizing specific histone modifications.

Thus, human MLL1/MLL2 methyltransferases can bind activated estrogen receptor-α through the associated tumor suppressor Menin (Dreijerink et al. 2006), TRR, the major *Drosophila* H3-K4 methyltransferase, can be targeted to ecdysone-responsive promoters through direct association with ecdysone nuclei receptor (Sedkov et al. 2003), MLL1 can associate with E2F transcription factor 6 (Dou et al. 2005), Trithorax and MLL methyltransferases may be targeted to chromatin through association with heat shock protein HSP90 (Tariq et al. 2009), PRC2 complexes can be site-specifically anchored to DNA by PHO/PHO-like/YY-1 DNA-binding proteins (Brown et al. 2003), etc. The recruitment of SET-domain proteins may also involve direct interactions of HKMTs with specific DNA sequences; for example, the interaction of MLL1 with DNA through CXXC domain, which binds to non-methylated CpG DNA sites (Cierpicki et al. 2010), could contribute to stable association of MLL1 with HoxA9 genes (Milne et al. 2010). Direct interaction of NSD1, -2, -3 and PR-SET7/8 SET domains with DNA may be essential for methylation specificity and activity of these enzymes (Li et al. 2009).

The recruitment of SET-domain proteins may implicate the recognition of site-specific histone modifications and histone variants. For example, PHD motifs of MLL1 and Trithorax proteins can recognize histone H3 trimethylated at

lysine 4 and thus contribute to the stable chromatin association (Chang et al. 2010; Milne et al. 2010). Suv39h1, -2 HKMTs can be targeted to chromatin through association of their C-terminal chromoshadow domain with HP1 chromodomain protein, which selectively binds di- and trimethylated lysine 9 in histone H3. Similarly, E(z) can associate through its Esc subunit with Polycomb, which recognizes H3-K27 trimethylation (Daniel et al. 2005; Schuettengruber and Cavalli 2009). The bromodomains of Trithorax and MLL methyltransferases and of their associated proteins can recognize histone tails acetylated at specific lysine residues (Yang 2004). Association of SET-domain proteins with chromatin may also involve recognition of histone variants. For example, histone variant H3.3, which is preferentially deposited at gene regulatory elements, is enriched in lysine methylation associated with active gene transcription (Ng and Gurdon 2008), which suggests that it may facilitate recruitment of SET proteins, presumably by promoting more accessible chromatin configuration (ibid).

Many of the HKMT-associated subunits in vitro can selectively bind histones with di- and trimethylated substrate lysine through their histone-recognition motifs. However, in vivo, this recognition of specific histone methylation states most likely confers proper di- and trimethylation of target lysines through control of the catalytic cycle, but not for the recruitment of HKMTs to their chromatin loci per se or for the basic monomethylation of chromatin. HKMT conserved subunits—the WD40 repeat proteins (Smith 2008) such as the human WDR5 (WDS in *Drosophila*, Cps30 in yeast), RbBP5, RbAp48/46 (p55 in *Drosophila*), and Eed (Esc/Escl in *Drosophila*)—can recognize histones through their repeated regions of β-propeller structures and "present" them to methyltransferase catalytic domains (Ruthenburg et al. 2007a; Suganuma et al. 2008). Wdr5, the common subunit of Trithorax-group HKMT complexes, recognizes dimethylated K4 in histone H3 and promotes H3-K4 trimethylation in vitro (Schuetz et al. 2006; Suganuma et al. 2008; Trievel and Shilatifard 2009), although Wdr5 is dispensable for association of MLL1 with its chromatin targets and for chromatin di- (but not tri-) methylation in vivo (Wysocka et al. 2005; Suganuma et al. 2008). RbAp48, the common subunit of Polycomb PRC2 complexes, recognizes histone H3 and H4 termini, although within PRC2 complex its H4 binding is restricted (Song et al. 2008; Suganuma et al. 2008). The Esc and p55 subunits of *Drosophila* PRC2 are both required for association of PRC2 with nucleosomes in vitro (Nekrasov et al. 2005). *Drosophila* Esc and Escl and human Eed have been shown to specifically bind histone H3 in vitro in a H3 tail- and modification-independent manner that was essential for E(z)-dependent trimethylation of H3-K27 in vivo (Tie et al. 2007). However, Esc and Escl were dispensable for E(z) targeting and monomethylation of chromatin in vivo (Kurzhals et al. 2008). Human Ezh2, in association with Suz12 and Eed, specifically binds trimethylated H3-K27 (Hansen et al. 2008), although it also has been reported that Eed alone can recognize trimethylated forms of K9 or 27 in histone H3 and K20 in H4 (gene silencing methylation) that results in allosteric activation of Ezh2 methyltransferase (Margueron et al. 2009; Suganuma and Workman 2010)—it has been proposed that these mechanisms may be implicated in propagating repressive H3 K27 trimethylation over extended genomic domains, as well as in transmitting gene silence to progeny cells (ibid).

Proper di- and trimethylation of lysine 4 in histone H3 depend on monoubiquitination of K123 in histone H2B (K120 in mammals) (Shilatifard 2008; Weake and Workman 2008; Shukla et al. 2009). However, lack of H2B ubiquitination does not affect the recruitment of SET-domain methyltransferases to gene-specific loci and monomethylation of histone lysines. Instead, it likely affects the correct compositional assembly of recruited HKMTs with WD-repeated histone-binding subunits (Dehe and Geli 2006; Shilatifard 2008), such as the Cps35 of yeast Set1 and its homolog Wdr82 from human hSet1, which associate with chromatin in H2B ubiquitination-dependent but Set1/hSet1-independent manner (Lee et al. 2007; Wu et al. 2008). Thus, complementation of Set1 from a ubiquitination-deficient background with wild-type Cps35 confers Set1 trimethylation activity (Lee et al. 2007). However, some reports contest the absolute need of H2B ubiquitylation for Set1 trimethylation activity (Foster and Downs 2009; Wang et al. 2009) or recruitment of Cps35 (Vitaliano-Prunier et al. 2008).

In addition, proper methylation of a specific lysine residue in a particular histone often depends on histone termini of the same or another histone (Fingerman et al. 2008; Lee et al. 2010). For example, Set2 H3-K36 methyltransferase, after being recruited to elongating RNA polymerase II, requires interaction of the Set2 amino-terminus with histone H4 at K44 (Du et al. 2008), histone H2A at L116,117 (Du and Briggs 2010), and with additional regions in histone H2A (ibid) and histone H3 N termini (Psathas et al. 2009), although the latter affects Set2 activity at a post-recruitment step (ibid). Phosphorylation of serine 10 in histone H3 prevents H3-K9 methylation (Schultz et al. 2002), H3-K9 acetylation may facilitate H3-K4 methylation in vitro (Southall et al. 2009), H3-K9, -R2 methylation and H3-K4 methylation may inhibit each other (Wang et al. 2001; Guccione et al. 2007), etc.

(b) SET-domain methyltransferases can be recruited to chromatin at 5′ of active genes by the components of transcriptional machinery through association with the Ser2/Ser5-phosphorylated carboxy-terminal heptapeptide (CTD) of newly elongated RNA polymerase II

(Dehe and Geli 2006; Ruthenburg et al. 2007a; Shilatifard 2008), as was demonstrated for *Saccharomyces cerevisiae* Set1 (Ng et al. 2003) and Set2 (Schaft et al. 2003), *Drosophila* Trithorax (Smith et al. 2004), human MLL1, MLL2 (Milne et al. 2005), and hSet1 (Wysocka et al. 2005) and Ash1 methyltransferases (Gregory et al. 2007), etc.

Although SET proteins are primarily associated with the RNA polymerase II holoenzyme, the components of the Paf elongation complex are essential for recruitment of SET proteins complexes to transcribed genes (Krogan et al. 2003; Milne et al. 2005). However, the proper enzymatic activity of RNA pol II-recruited methyltransferase complexes depends on the presence of WD-repeated proteins like WDR5 for MLL/Trithorax or Cps35 for Set1 and on the monoubiquitination of histone H2B (Ruthenburg et al. 2007b; Shilatifard 2008).

(c) the above SET-domain recruitment principles are comprehensively overviewed in a number of recent reviews cited above. Here, we would like to discuss in more detail one more possible mechanism of chromatin recognition by SET-domain proteins, which may be implemented through direct interaction of SET domain with nucleic acid or histone components of transcriptionally active chromatin.

Evidence for this mechanism has come from the discovery of a high-affinity histone-binding motif in the SET domain of *Drosophila* Trithorax, which selectively and efficiently binds the N-terminal region of histone H3 (Katsani et al. 2001). A point mutation trxZ11 in Trithorax SET (G3601S), which severely impairs SET-domain functions and results in homeotic transformations and lethality in flies, also incapacitates the histone-binding ability of SET domain, suggestive that histone–SET interactions are involved the maintenance of the pattern of gene activity in vivo (ibid). As illustrated in Fig. 1, the SET-domain polypeptide of Set7/9 methyltransferase preferentially methylates SET-associated histones H3, suggesting that tight binding of histone H3 is essential for histone methylation. However, while SET domains of Trithorax and some other methyltransferases possess high affinity to single histone H3, SET domains inefficiently bind nucleosomes (Krajewski and Reese 2010), consistent with the observations that various SET domains exhibit high methyltransferase activity on histones but poorly methylate histones inside nucleosomes (Wang et al. 2001; Martin et al. 2006; Krajewski and Reese 2010). This suggests that histone–SET interactions are restricted in compact "canonical" nucleosomes and that the nucleosome particle must be structurally altered for association with SET domain.

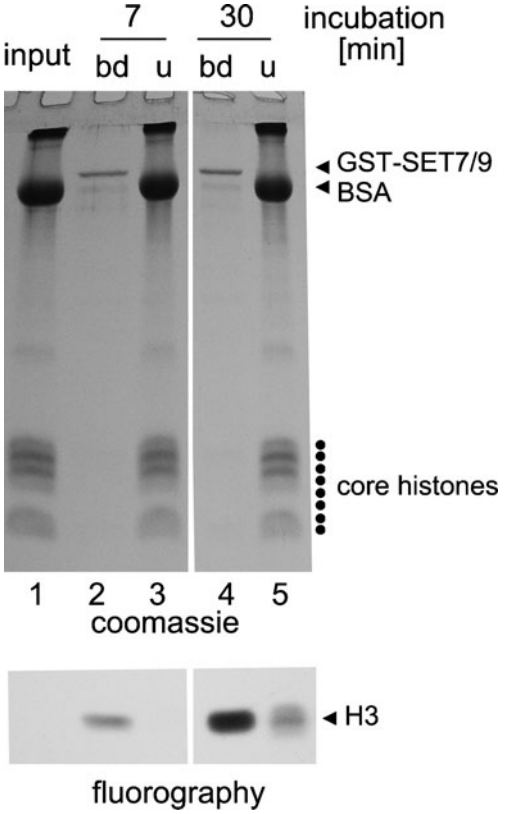

Fig. 1 An illustration of the relationship between SET-domain histone binding and histone methylation. Set7/9 methyltransferase preferentially methylates SET-domain-associated histones H3. Bacterially expressed GST-tagged full-size SET7/9 protein, immobilized on glutathione–sepharose, was incubated for the indicated time with excess of HeLa cell core histones and 100 μg/ml BSA in the presence of ^{3}H-*S*-adenosyl-methionine. The GST-SET beads were sequentially washed in buffers containing 0.2% of NP-40 and 0.2, 0.4, and 0.6 M NaCl. Bead-associated proteins ("bound", *bd*—lanes *2* and *4*) and TCA-precipitated pooled wash fractions ("unbound", *u*—lanes *3* and *5*) were resolved on an SDS gel and stained with Coomassie (*top panel*). The H3-containing slice was excised from gel, treated with EN3HANCE reagent (Perkin-Elmer), dried, and exposed to film. Lane *1* shows input histones in the reaction buffer

The studies of most SET domains complexed with histone polypeptides suggest that the target lysine residue accesses the cofactor *S*-adenosylmethionine, located on the opposite face of catalytic unit, through a narrow "trans-enzyme" lysine-access channel, composed of hydrophobic residues (Bottomley 2004; Southall et al. 2009; Wu et al. 2010 and references therein). The steric features of the lysine-access channel determine the SET-domain specificity and levels of methylation. By changing the size of the lysine-binding groove to match mono- or di- and trimethylated lysines, it was possible to convert SET7/9 monomethylase to trimethylase (Xiao et al. 2003) and Dim5 trimethylase to monomethylase (Zhang et al. 2003). The proper levels of histone methylation in vivo are likely promoted by the WD40 repeated HKMT's subunits, which control the interactions of the target lysine with the methyl-donating cofactor in the SET-domain lysine-

access channel (Song and Kingston 2008). The strict spatial disposition of the histone terminus within the catalytic domain may explain the requirement for a steric accessibility of histone termini for their unhindered binding to the SET-domain lysine-access pocket.

The nucleosome consists of a 147-bp DNA fragment wrapped around an octamer of histone proteins H2A, H2B, H3, and H4, which compose a H3-H4 tetramer flanked on either side with a H2A-H2B dimer (Luger et al. 1997). The flexible N termini of histones are engaged by interactions with extranucleosome DNA and between themselves in an intra- and internucleosomal manner and, thus, are not readily accessible to external proteins (Davey et al. 2002; Zheng and Hayes 2003). The dense nucleosome packaging can be altered by a family of ATP-dependent chromatin remodelers which could also increase the accessibility of histone octamers through multiple mechanisms, including partial or complete eviction of histone H2A–H2B dimers, nucleosome dimerization, formation of intra- and internucleosome DNA loops, etc. (Clapier and Cairns 2009). Nucleosomes can be unfolded during transcription, resulting in a transient loss of H2A–H2B dimers (Kulaeva et al. 2007) and formation of stably altered nucleosome transcriptional intermediates (Nacheva et al. 1989; Bazett-Jones et al. 1996). The accessibility of histone termini can be affected by nucleosome–nucleosome interactions and by binding of linker histone H1 (Arya and Schlick 2009; Kan et al. 2007, 2009), or by posttranslational modification of histones (Campos and Reinberg 2009). For example, the acetylation of histone N termini—a hallmark for active chromatin regions—causes weakening of electrostatic interaction of histone tails with DNA (Choi and Howe 2009) that promotes opening of the nucleosome structure and increases accessibility of histone termini.

Internucleosome interactions within di- and oligonucleosomes and/or remodeling of di- and oligonucleosome by incorporations of histone H1 significantly stimulated the methyltransferase activity of the Ezh2 SET domain, which exhibited weak methyltransferase activity towards intact mononucleosomes (Martin et al. 2006). Remodeling of dinucleosome, but not mononucleosome, templates with purified Isw1 and Isw2 complexes facilitated association of nucleosomes with GST-tagged SET-domain polypeptides of MLL1 and SET7/9 (Krajewski and Reese 2010). It also has been shown that nucleosomes reconstituted from hyperacetylated histones possess an increased affinity to GST-SET polypeptide of MLL1 (Krajewski and Vassiliev 2010). It may be presumed that any process that changes the chromatin conformation with a concomitant release of histone N termini could provide a structural basis for binding of SET domains to nucleosomes. Liberation of histone termini could alleviate their proper positioning in the SET-domain lysine-binding pocket, thereby partially eliminating the requirement for auxiliary WD-repeated factors.

In contrast, the yeast Set2 H3-K36- and human PR-Set7/8 H4-K20 methyltransferases exhibit a preference for nucleosomes over histones (Nishioka et al. 2002; Strahl et al. 2002). Set2, unlike many other SET-domain HKMTs, binds histone H3 peptide inefficiently (only H3-K36 non-methylated form), but can tightly bind nucleosomes in vitro through interaction of its N terminus H3K36-like motif in histone H4 ($G_{41}GVKR_{45}$ vs. $G_{33}GVKK_{37}$ in H3) (Du et al. 2008), which is relatively exposed (Davey et al. 2002), and with other histone regions (Psathas et al. 2009; Du and Briggs 2010) (see text above). However, in vivo, multiple Set2-nucleosome contacts are essential for H3-K36 di- and trimethylation, but are dispensable for recruitment of Set2 to chromatin and H3-K36 monomethylation (ibid). The nucleosome preference of PR-Set7/8 may be explained by an unusually high flexibility of its lysine-binding channel so the substrate itself contributes to the structure of the channel (Xiao et al. 2005). This facilitates docking of the H4-K20 lysine to the methyl-donating cofactor. Consistent with this observation, at low NaCl concentration, the PR-Set7/8 is equally efficient on nucleosome and bare histones substrates (ibid). In addition, the nucleosome preference of PR-Set7/8 is lost upon deletion of the first 14 amino acids from the N terminus of the protein (Nishioka et al. 2002) or in the presence of short nucleic acid fragments (Li et al. 2009), which likely facilitates the conformational changes within SET domain (ibid).

Additional evidence for the ability of SET domains to specifically target "open" chromatin is based on the ability of SET to bind single-stranded nucleic acids. The SET-domain region of diverse HKMTs contains a polypeptide motif named SSBL (SSB-like), located in the pre-SET region or the boundary of SET and pre-SET regions (Fig. 2), which can tightly bind single-stranded DNA

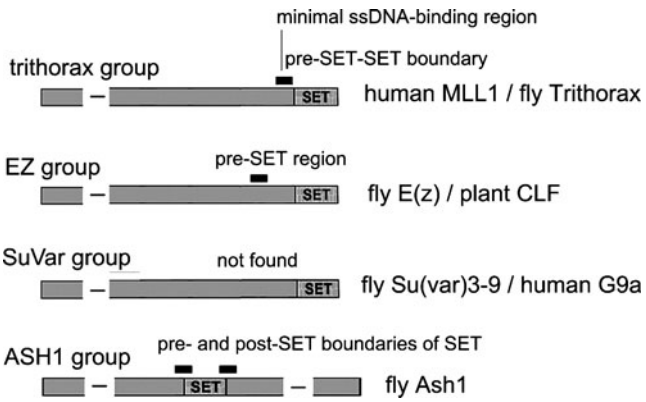

Fig. 2 Schematic diagram showing positions (*black rectangle*) of the ssDNA-binding motif in the SET-domain regions of representatives from the Trithorax, E(z), Su(var)3–9, and ASH1 SET-domain protein families (Krajewski et al. 2005). In addition to the indicated proteins, the ssDNA-binding motif was found (although not mapped precisely) in the SET-domain region of yeast Set1, Set2, and human ALR1 proteins. This motif was not found in the Set7/9 methyltransferase (unpublished observations)

(ssDNA) and RNA and can recognize ssDNA stretches in supercoiled and in vitro-transcribed DNA (Krajewski et al. 2005). The SSBL motif was found a SET domains belonging to E(z), Ash1, and Trithorax, but not Su(var)3–9, SET families (SET classification as suggested by Alvarez-Venegas and Avramova 2002). Therefore, the existence and location (Fig. 2) of the SSBL motif correlates with SET-domain functional activity. The SSBL binds single-stranded DNA substrates with similar selectivity and efficiency as the major *Escherichia coli* SSB protein (ibid), suggestive that SSBL motif in vivo may participate in the processes which involve the regular SSB proteins, i.e., almost any time that single-stranded DNA is present or requires manipulation.

It has been shown that the SSBL motif in the SET domain of Ash1 can recruit Ash1 protein to its target sites in the *Drosophila* homeotic gene Ultrabithorax (UBX) through association with UBX-associated non-coding RNAs (Sanchez-Elsner et al. 2006). The non-coding RNA transcripts of three Trithorax response elements (TREs) in UBX can mediate transcription activation by recruiting Ash1 to the template TREs. The SET domain of Ash1 binds all three TRE transcripts with each TRE transcript recruiting Ash1 only to the corresponding TRE in chromatin (ibid). The authors suggest that Ash1 is recruited to TREs through association with the stretches of ssDNA exposed during ongoing transcription and the non-coding RNAs generated by transcription (ibid). Schmitt and Paro (2006) suggested that non-coding RNAs that survived mitosis might serve for re-targeting of Ash1 to a TRE to ensure the re-establishment of epigenetic activation after DNA replication. In addition, direct interactions of SET domain with nucleic acid (RNA or DNA) may regulate their HKMT's activity and specificity, facilitating proper positioning the substrate lysine side chain in the SET-domain lysine-access channel, similar to what has been shown for NSD2 and PR-SET7/8 in vitro (Li et al. 2009).

Various types of histone methylation are hallmarks of not only gene-regulatory and transcription initiation regions but also of the entire transcribed genes (Li et al. 2007). Functionally active chromatin assume a more "open" configuration with differential packaging of DNA into nucleosomes (Gilbert and Ramsahoye 2005). Therefore, it can be presumed that H3 and H4 histone termini in functioning chromatin are more readily accessible to SET domains (see text above). Transcriptionally active DNA also contains extended regions of ssDNA, RNA transcripts, and multiple DNA regions with distorted base pairing, which represent potential targets for SET-domain proteins. This, in principle, may play a role in the HKMT discrimination of active and repressed chromatin. The recruitment of SET-domain-containing proteins to their chromatin loci may involve their association with "exposed" histone tails and ssDNA and RNA structures, which form during transcription and other nucleosome-templated processes. The ability of SET-domain proteins to bind single-stranded nucleic acids may be also involved in the heritable propagation of chromatin methylation patterns. For example, upon progression of the replication fork, the SET-domain proteins may be distributed between the parental and newly synthesized DNA strands due to the high affinity between SET and ssDNA. In this way, SET-domain proteins may be implicated not only in methylating chromatin for transcription and repression but also in transmitting these states to descendant cells.

In view of the above, it is of interest that the Trithorax SET domain is indispensable for targeting Trithorax protein to transcribed heat shock genes in vivo (Smith et al. 2004). The expressed Trithorax SET-domain polypeptide is recruited in vivo to the heat shock genes only during their active transcription (ibid). In addition, maintaining the active state of the *Drosophila* Ultrabithorax gene requires the initial transcription of the Ultrabithorax regulatory region (Bender and Fitzgerald 2002; Hogga and Karch 2002; Rank et al. 2002). Using the ChIP-on-chip technique, it has been shown that *Drosophila* Trithorax binds to the promoters of active genes and non-coding transcripts (Beisel et al. 2007). The SET domain of Ezh proteins is required for their recruitment to target genes. A point mutation in the SET domain that prevents association of PRC2 complexes with chromatin does not affect the integrity of the complex (Margueron et al. 2008), suggesting that the SET domains of Ezh1,-2 are directly implicated in targeting of Ezh to chromatin.

Obviously, the aforementioned mechanism addresses only one of the many aspects of the interaction of SET-domain-containing proteins with chromatin. However, considering the high affinity of SET to histones and single-stranded nucleic acids, these interactions may be important in the transcription regulatory mechanisms involved in cell determination.

Acknowledgments The authors are very grateful to Prof. David Beebe (Washington University School of Medicine, St. Louis, MO) for the critical reading and correction of the manuscript. The research was funded by the grants from the Russian Foundation for Basic Research (08-04-01010).

Conflict of interest None.

References

Albert M, Helin K (2010) Histone methyltransferases in cancer. Semin Cell Dev Biol 21:209–220

Alvarez-Venegas R, Avramova Z (2002) SET-domain proteins of the Su(var)3–9, E(z) and trithorax families. Gene 285:25–37

Arya G, Schlick T (2009) A tale of tails: how histone tails mediate chromatin compaction in different salt and linker histone environments. J Phys Chem A 113:4045–4059

Bazett-Jones DP, Mendez E, Czarnota GJ et al (1996) Visualization and analysis of unfolded nucleosomes associated with transcribing chromatin. Nucleic Acids Res 24:321–329

Beisel C, Buness A, Roustan-Espinosa IM et al (2007) Comparing active and repressed expression states of genes controlled by the Polycomb/Trithorax group proteins. Proc Natl Acad Sci USA 104:16615–16620

Bender W, Fitzgerald DP (2002) Transcription activates repressed domains in the Drosophila bithorax complex. Development 129:4923–4930

Bottomley MJ (2004) Structures of protein domains that create or recognize histone modifications. EMBO Rep 5:464–469

Breiling A, Sessa L, Orlando V (2007) Biology of polycomb and trithorax group proteins. Int Rev Cytol 258:83–136

Brown JL, Fritsch C, Mueller J et al (2003) The Drosophila pho-like gene encodes a YY1-related DNA binding protein that is redundant with pleiohomeotic in homeotic gene silencing. Development 130:285–294

Campos EI, Reinberg D (2009) Histones: annotating chromatin. Annu Rev Genet 43:559–599

Cavener DR (2009) Sleeping Beauty, awake! Regulation of insulin gene expression by methylation of histone H3. Diabetes 58:28–29

Chang PY, Hom RA, Musselman CA et al (2010) Binding of the MLL PHD3 finger to histone H3K4me3 is required for MLL-dependent gene transcription. J Mol Biol 400:137–144

Chen H, Gu X, Su IH et al (2009) Polycomb protein Ezh2 regulates pancreatic beta-cell Ink4a/Arf expression and regeneration in diabetes mellitus. Genes Dev 23:975–985

Choi JK, Howe LJ (2009) Histone acetylation: truth of consequences? Biochem Cell Biol 87:139–150

Cierpicki T, Risner LE, Grembecka J et al (2010) Structure of the MLL CXXC domain–DNA complex and its functional role in MLL-AF9 leukemia. Nat Struct Mol Biol 17:62–68

Clapier CR, Cairns BR (2009) The biology of chromatin remodeling complexes. Annu Rev Biochem 78:273–304

Cosgrove MS, Patel A (2010) Mixed lineage leukemia: a structure–function perspective of the MLL1 protein. FEBS J 277:1832–1842

Daniel JA, Pray-Grant MG, Grant PA (2005) Effector proteins for methylated histones: an expanding family. Cell Cycle 4:919–926

Davey CA, Sargent DF, Luger K et al (2002) Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 a resolution. J Mol Biol 319:1097–1113

Deering TG, Ogihara T, Trace AP et al (2009) Methyltransferase Set7/9 maintains transcription and euchromatin structure at islet-enriched genes. Diabetes 58:185–193

Dehe PM, Geli V (2006) The multiple faces of Set1. Biochem Cell Biol 84:536–548

Devaskar SU, Thamotharan M (2007) Metabolic programming in the pathogenesis of insulin resistance. Rev Endocr Metab Disord 8:105–113

Dou Y, Milne TA, Tackett AJ et al (2005) Physical association and coordinate function of the H3 K4 methyltransferase MLL1 and the H4 K16 acetyltransferase MOF. Cell 121:873–885

Dou Y, Milne TA, Ruthenburg AJ et al (2006) Regulation of MLL1 H3K4 methyltransferase activity by its core components. Nat Struct Mol Biol 13:713–719

Dreijerink KM, Mulder KW, Winkler GS et al (2006) Menin links estrogen receptor activation to histone H3K4 trimethylation. Cancer Res 66:4929–4935

Du HN, Briggs SD (2010) A nucleosome surface formed by histone H4, H2A, and H3 residues is needed for proper histone H3 Lys36 methylation, histone acetylation, and repression of cryptic transcription. J Biol Chem 285:11704–11713

Du HN, Fingerman IM, Briggs SD (2008) Histone H3 K36 methylation is mediated by a trans-histone methylation pathway involving an interaction between Set2 and histone H4. Genes Dev 22:2786–2798

Fingerman IM, Du HN, Briggs SD (2008) Controlling histone methylation via trans-histone pathways. Epigenetics 3:237–242

Foster ER, Downs JA (2009) Methylation of H3 K4 and K79 is not strictly dependent on H2B K123 ubiquitylation. J Cell Biol 184:631–638

Gilbert N, Ramsahoye B (2005) The relationship between chromatin structure and transcriptional activity in mammalian genomes. Brief Funct Genomic Proteomic 4:129–142

Granger A, Abdullah I, Huebner F et al (2008) Histone deacetylase inhibition reduces myocardial ischemia–reperfusion injury in mice. FASEB J 22:3549–3560

Gregory GD, Vakoc CR, Rozovskaia T et al (2007) Mammalian ASH1L is a histone methyltransferase that occupies the transcribed region of active genes. Mol Cell Biol 27:8466–8479

Guccione E, Bassi C, Casadio F et al (2007) Methylation of histone H3R2 by PRMT6 and H3K4 by an MLL complex are mutually exclusive. Nature 449:933–937

Hansen KH, Bracken AP, Pasini D et al (2008) A model for transmission of the H3K27me3 epigenetic mark. Nat Cell Biol 10:1291–1300

Hogga I, Karch F (2002) Transcription through the iab-7 cis-regulatory domain of the bithorax complex interferes with maintenance of Polycomb-mediated silencing. Development 129:4915–4922

Kaminsky Z, Wang SC, Petronis A (2006) Complex disease, gender and epigenetics. Ann Med 38:530–544

Kan PY, Lu X, Hansen JC et al (2007) The H3 tail domain participates in multiple interactions during folding and self-association of nucleosome arrays. Mol Cell Biol 27:2084–2091

Kan PY, Caterino TL, Hayes JJ (2009) The H4 tail domain participates in intra- and internucleosome interactions with protein and DNA during folding and oligomerization of nucleosome arrays. Mol Cell Biol 29:538–546

Kaneda R, Takada S, Yamashita Y et al (2009) Genome-wide histone methylation profile for heart failure. Genes Cells 14:69–77

Katsani KR, Arredondo JJ, Kal AJ et al (2001) A homeotic mutation in the trithorax SET domain impedes histone binding. Genes Dev 15:2197–2202

Krajewski WA, Reese JC (2010) SET domains of histone methyltransferases recognize ISWI-remodeled nucleosomal species. Mol Cell Biol 30:552–564

Krajewski WA, Vassiliev OL (2010) Histone acetylation facilitates association of nucleosomes with SET domain of ALL-1 methyltransferase in vitro. Biochem Biophys Res Commun 397:112–116

Krajewski WA, Nakamura T, Mazo A et al (2005) A motif within SET-domain proteins binds single-stranded nucleic acids and transcribed and supercoiled DNAs and can interfere with assembly of nucleosomes. Mol Cell Biol 25:1891–1899

Krogan NJ, Dover J, Wood A et al (2003) The Paf1 complex is required for histone H3 methylation by COMPASS and Dot1p: linking transcriptional elongation to histone methylation. Mol Cell 11:721–729

Kulaeva OI, Gaykalova DA, Studitsky VM (2007) Transcription through chromatin by RNA polymerase II: histone displacement and exchange. Mutat Res 618:116–129

Kurzhals RL, Tie F, Stratton CA et al (2008) Drosophila ESC-like can substitute for ESC and becomes required for Polycomb silencing if ESC is absent. Dev Biol 313:293–306

Lee JS, Shukla A, Schneider J et al (2007) Histone crosstalk between H2B monoubiquitination and H3 methylation mediated by COMPASS. Cell 131:1084–1096

Lee JS, Smith E, Shilatifard A (2010) The language of histone crosstalk. Cell 142:682–685
Li B, Carey M, Workman JL (2007) The role of chromatin during transcription. Cell 128:707–719
Li Y, Reddy MA, Miao F et al (2008) Role of the histone H3 lysine 4 methyltransferase, SET7/9, in the regulation of NF-kappaB-dependent inflammatory genes. Relevance to diabetes and inflammation. J Biol Chem 283:26771–26781
Li Y, Trojer P, Xu CF et al (2009) The target of the NSD family of histone lysine methyltransferases depends on the nature of the substrate. J Biol Chem 284:34283–34295
Luger K, Mader AW, Richmond RK et al (1997) Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature 389:251–260
Maekawa M, Watanabe Y (2007) Epigenetics: relations to disease and laboratory findings. Curr Med Chem 14:2642–2653
Margueron R, Li G, Sarma K et al (2008) Ezh1 and Ezh2 maintain repressive chromatin through different mechanisms. Mol Cell 32:503–518
Margueron R, Justin N, Ohno K et al (2009) Role of the polycomb protein EED in the propagation of repressive histone marks. Nature 461:762–767
Marmorstein R, Trievel RC (2009) Histone modifying enzymes: structures, mechanisms, and specificities. Biochim Biophys Acta 1789:58–68
Martin C, Cao R, Zhang Y (2006) Substrate preferences of the EZH2 histone methyltransferase complex. J Biol Chem 281:8365–8370
Milne TA, Dou Y, Martin ME et al (2005) MLL associates specifically with a subset of transcriptionally active target genes. Proc Natl Acad Sci USA 102:14765–14770
Milne TA, Kim J, Wang GG et al (2010) Multiple interactions recruit MLL1 and MLL1 fusion proteins to the HOXA9 locus in leukemogenesis. Mol Cell 38:853–863
Nacheva GA, Guschin DY, Preobrazhenskaya OV et al (1989) Change in the pattern of histone binding to DNA upon transcriptional activation. Cell 58:27–36
Nakamura T, Mori T, Tada S et al (2002) ALL-1 is a histone methyltransferase that assembles a supercomplex of proteins involved in transcriptional regulation. Mol Cell 10:1119–1128
Nekrasov M, Wild B, Muller J (2005) Nucleosome binding and histone methyltransferase activity of Drosophila PRC2. EMBO Rep 6:348–353
Ng RK, Gurdon JB (2008) Epigenetic inheritance of cell differentiation status. Cell Cycle 7:1173–1177
Ng HH, Robert F, Young RA et al (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
Nishioka K, Rice JC, Sarma K et al (2002) PR-Set7 is a nucleosome-specific methyltransferase that modifies lysine 20 of histone H4 and is associated with silent chromatin. Mol Cell 9:1201–1213
Perini G, Tupler R (2006) Altered gene silencing and human diseases. Clin Genet 69:1–7
Petruk S, Sedkov Y, Smith S et al (2001) Trithorax and dCBP acting in a complex to maintain expression of a homeotic gene. Science 294:1331–1334
Psathas JN, Zheng S, Tan S et al (2009) Set2-dependent K36 methylation is regulated by novel intratail interactions within H3. Mol Cell Biol 29:6413–6426
Qian C, Zhou MM (2006) SET domain protein lysine methyltransferases: structure, specificity and catalysis. Cell Mol Life Sci 63:2755–2763
Rank G, Prestel M, Paro R (2002) Transcription through intergenic chromosomal memory elements of the Drosophila bithorax complex correlates with an epigenetic switch. Mol Cell Biol 22:8026–8034
Ruthenburg AJ, Allis CD, Wysocka J (2007a) Methylation of lysine 4 on histone H3: intricacy of writing and reading a single epigenetic mark. Mol Cell 25:15–30
Ruthenburg AJ, Li H, Patel DJ et al (2007b) Multivalent engagement of chromatin modifications by linked binding modules. Nat Rev Mol Cell Biol 8:983–994
Sanchez-Elsner T, Gou D, Kremmer E et al (2006) Noncoding RNAs of trithorax response elements recruit Drosophila Ash1 to Ultrabithorax. Science 311:1118–1123
Schaft D, Roguev A, Kotovic KM et al (2003) The histone 3 lysine 36 methyltransferase, SET2, is involved in transcriptional elongation. Nucleic Acids Res 31:2475–2482
Schmitt S, Paro R (2006) RNA at the steering wheel. Genome Biol 7:218
Schuettengruber B, Cavalli G (2009) Recruitment of polycomb group complexes and their role in the dynamic regulation of cell fate choice. Development 136:3531–3542
Schuetz A, Allali-Hassani A, Martin F et al (2006) Structural basis for molecular recognition and presentation of histone H3 by WDR5. EMBO J 25:4245–4252
Schultz DC, Ayyanathan K, Negorev D et al (2002) SETDB1: a novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins. Genes Dev 16:919–932
Schwartz DA (2010) Epigenetics and environmental lung disease. Proc Am Thorac Soc 7:123–125
Schwartz YB, Pirrotta V (2007) Polycomb silencing mechanisms and the management of genomic programmes. Nat Rev Genet 8:9–22
Sedkov Y, Cho E, Petruk S et al (2003) Methylation at lysine 4 of histone H3 in ecdysone-dependent development of Drosophila. Nature 426:78–83
Shilatifard A (2008) Molecular implementation and physiological roles for histone H3 lysine 4 (H3K4) methylation. Curr Opin Cell Biol 20:341–348
Shukla A, Chaurasia P, Bhaumik SR (2009) Histone methylation and ubiquitination with their cross-talk and roles in gene expression and stability. Cell Mol Life Sci 66:1419–1433
Simon JA, Lange CA (2008) Roles of the EZH2 histone methyltransferase in cancer epigenetics. Mutat Res 647:21–29
Smith TF (2008) Diversity of WD-repeat proteins. Subcell Biochem 48:20–30
Smith ST, Petruk S, Sedkov Y et al (2004) Modulation of heat shock gene expression by the TAC1 chromatin-modifying complex. Nat Cell Biol 6:162–167
Song JJ, Kingston RE (2008) WDR5 interacts with mixed lineage leukemia (MLL) protein via the histone H3-binding pocket. J Biol Chem 283:35258–35264
Song JJ, Garlick JD, Kingston RE (2008) Structural basis of histone H4 recognition by p55. Genes Dev 22:1313–1318
Southall SM, Wong PS, Odho Z et al (2009) Structural basis for the requirement of additional factors for MLL1 SET domain activity and recognition of epigenetic marks. Mol Cell 33:181–191
Strahl BD, Grant PA, Briggs SD et al (2002) Set2 is a nucleosomal histone H3-selective methyltransferase that mediates transcriptional repression. Mol Cell Biol 22:1298–1306
Strietholt S, Maurer B, Peters MA et al (2008) Epigenetic modifications in rheumatoid arthritis. Arthritis Res Ther 10:219
Suganuma T, Workman JL (2010) WD40 repeats arrange histone tails for spreading of silencing. J Mol Cell Biol 2:81–83
Suganuma T, Pattenden SG, Workman JL (2008) Diverse functions of WD40 repeat proteins in histone recognition. Genes Dev 22:1265–1268
Szyf M (2010) Epigenetic therapeutics in autoimmune disease. Clin Rev Allergy Immunol 39:62–77
Tariq M, Nussbaumer U, Chen Y et al (2009) Trithorax requires Hsp90 for maintenance of active chromatin at sites of gene expression. Proc Natl Acad Sci USA 106:1157–1162

Tie F, Stratton CA, Kurzhals RL et al (2007) The N terminus of Drosophila ESC binds directly to histone H3 and is required for E(Z)-dependent trimethylation of H3 lysine 27. Mol Cell Biol 27:2014–2026

Trievel RC, Shilatifard A (2009) WDR5, a complexed protein. Nat Struct Mol Biol 16:678–680

van Leeuwen F, Gafken PR, Gottschling DE (2002) Dot1p modulates silencing in yeast by methylation of the nucleosome core. Cell 109:745–756

Villeneuve LM, Reddy MA, Lanting LL et al (2008) Epigenetic histone H3 lysine 9 methylation in metabolic memory and inflammatory phenotype of vascular smooth muscle cells in diabetes. Proc Natl Acad Sci USA 105:9047–9052

Vitaliano-Prunier A, Menant A, Hobeika M et al (2008) Ubiquitylation of the COMPASS component Swd2 links H2B ubiquitylation to H3K4 trimethylation. Nat Cell Biol 10:1365–1371

Wang H, Cao R, Xia L et al (2001) Purification and functional characterization of a histone H3-lysine 4-specific methyltransferase. Mol Cell 8:1207–1217

Wang Z, Cui B, Gorovsky MA (2009) Histone H2B ubiquitylation is not required for histone H3 methylation at lysine 4 in tetrahymena. J Biol Chem 284:34870–34879

Weake VM, Workman JL (2008) Histone ubiquitination: triggering gene activity. Mol Cell 29:653–663

Wu M, Wang PF, Lee JS et al (2008) Molecular regulation of H3K4 trimethylation by Wdr82, a component of human Set1/COMPASS. Mol Cell Biol 28:7337–7344

Wu H, Min J, Lunin VV et al (2010) Structural biology of human H3K9 methyltransferases. PLoS ONE 5:e8570

Wysocka J, Swigut T, Milne TA et al (2005) WDR5 associates with histone H3 methylated at K4 and is essential for H3 K4 methylation and vertebrate development. Cell 121:859–872

Xiao B, Jing C, Wilson JR et al (2003) Structure and catalytic mechanism of the human histone methyltransferase SET7/9. Nature 421:652–656

Xiao B, Jing C, Kelly G et al (2005) Specificity and mechanism of the histone methyltransferase Pr-Set7. Genes Dev 19:1444–1454

Yang XJ (2004) Lysine acetylation and the bromodomain: a new partnership for signaling. Bioessays 26:1076–1087

Zhang X, Yang Z, Khan SI et al (2003) Structural basis for the product specificity of histone lysine methyltransferases. Mol Cell 12:177–185

Zheng C, Hayes JJ (2003) Structures and interactions of the core histone tail domains. Biopolymers 68:539–546