The role of a newly identified SET domain-containing protein, SETD3, in oncogenesis

Zhangguo Chen,[1] Catherine T. Yan,[2] Yali Dou,[3] Sawanee S. Viboolsittiseri,[1] and Jing H. Wang[1]

[1]Integrated Department of Immunology, University of Colorado School of Medicine and National Jewish Health, Denver, CO, USA; [2]Department of Pathology, Beth Israel Deaconess Medical Center, Boston, MA, USA; and [3]Department of Pathology, University of Michigan, Ann Arbor, MI, USA

ABSTRACT

The SET domain is found in histone methyltransferases and other lysine methyltransferases. SET domain-containing proteins such as MLL1 play a critical role in leukemogenesis, while others such as SETD2 may function as a tumor suppressor in breast cancer and renal cell carcinoma. We recently discovered that SETD3, a well-conserved SET domain-containing protein, was involved in a translocation to the immunoglobulin lambda light chain locus in one of the non-homologous end-joining/p53-deficient peripheral B-cell lymphomas. We showed that a truncated mRNA lacking the SET domain sequences in *Setd3* gene was highly expressed in the lymphoma. Furthermore, we found that the truncated SET-less protein displayed oncogenic potential while the full length SETD3 protein did not. Finally, SETD3 exhibits histone methyltransferases activity on nucleosomal histone 3 in a SET-domain dependent manner. We propose that this newly identified *Setd3* gene may play an important role in carcinogenesis.

Introduction

Genomic DNA in the nucleus is packaged into the nucleosome, which consists of an octamer of histones (two of each H2A, H2B, H3, and H4) wrapped around with 146bp genomic DNA.[1] The amino-terminal tails of the core histones can be modified by posttranslational modifications such as lysine (K) methylation (me), which is mediated by histone methyltransferases (HMTs).[2] The suppressor of variegation, enhancer of zeste, trithorax (SET) domain was first recognized as a conserved feature in chromatin-associated proteins[3,4] and a number of SET domain-containing proteins have since been characterized as HMTs.[5,6] For example, SUV39H1, SUV39H2, and SET9 methylate H3 or H4 and function as important epigenetic regulators of mammalian development, heterochromatin, and genomic instability.[7,8] Deregulation of SET-domain function has an important role in carcinogenesis.[9,10] Mixed-lineage leukemia 1 (MLL1) is one of the best studied SET domain-containing proteins in human cancer, which is disrupted by chromosomal translocations resulting in the aberrant expression of chimeric proteins.[11] Notably, all the fusion products contain the N-terminal 1400 amino acids of MLL1 and lack the SET domain.[11] In addition, a common feature of SET domain-containing proteins is that their full length proteins such as RIZ1, MDS1-EVI1 and MMSET-II appear to act as tumor suppressors, while the corresponding SET domain-lacking proteins, RIZ2, EVI1, and MMSET-I, function as oncogenes.[9,12] Truncation of the full length protein can be caused by either translocations or alternative splicing.[12]

V(D)J recombination is a site-specific DNA recombination process[13] which breaks and rejoins DNA segments of antigen receptor loci in lymphocyte progenitors.[13] DNA double strand breaks (DSBs) generated during V(D)J recombination are repaired by non-homologous end-joining (NHEJ). XRCC4 is an essential co-factor of DNA ligase IV (Lig4) and co-operates with Lig4 to catalyze the ligation step of NHEJ.[14] We previously inactivated Xrcc4 specifically in peripheral B cells.[15] Xrcc4 conditional knockout mice are not cancer-prone; however, once bred into a p53 deficient background, they often succumb to B-lineage lymphomas (termed CXP lymphomas) that frequently harbor clonal translocations.[16] In this study, we present a SET domain-containing protein, SETD3, which is involved in a translocation to the immunoglobulin lambda (Igλ) light chain locus in one of the CXP lymphomas.

Design and Methods

Cytogenetic assays, SKY, FISH and chromosomal painting

Preparation of metaphase chromosomes, spectral karyotyping (SKY), fluorescence *in situ* hybridization (FISH), and whole-chromosome painting using single-chromosome-specific paints were performed as previously described.[16] FISH probes were as follows: BACs for *Setd3* on chr12 are RP23-360E20, RP23-42A17, and RP23-32G7 obtained from the BACPAC CHORI database.

Further details of Design and Methods are available in the *Online Supplementary Appendix.*

Results and Discussion

Identification of Setd3 from a NHEJ/p53-deficient peripheral B cell lymphoma

We initially observed a t(12;16) translocation by SKY in the CXP163 tumor (*Online Supplementary Figure S1*).[16] Next, we attempted to map the junction of this translocation by FISH. Cytogenetic analysis using bacterial artificial chromosomes

 doi:10.3324/haematol.2012.066977
The online version of this article has a Supplementary Appendix.
Manuscript received on March 28, 2012. Manuscript accepted on September 26, 2012.
Correspondence: jing.wang@ucdenver.edu

(BACs) that contain mouse genomic sequences covering Bcl11b or Cyclin k (Ccnk) and whole chromosomal paint showed that the translocation breakpoint of t(12;16) in CXP163 was between Bcl11b and Ccnk, and SET domain containing 3 (*Setd3*) is the only gene in this region (Figure 1A). In addition, these analyses revealed a complex t(16;12;16) translocation in CXP163, which involved chromosome 16 (Chr16) and appeared karyotypically normal by SKY and chromosomal painting. FISH using serial BACs covering a region on Chr12 containing *Setd3* and Bcl11b, and other analyses showed that an approximately several hundred kb sequence from this portion of Chr12 was inserted into the Igλ locus on Chr16 to generate the complex t(16;12;16) translocation (Figure 1A). Thus, CXP163 contains a t(12;16) and t(16;12;16) complex translocation, both of which involve *Setd3*.

Southern blotting analyses showed that the CXP163 lymphomas had Igλ rearrangements (Figure 1B). To clarify the nature of these Igλ rearrangements, we cloned them from genomic phage libraries generated from CXP163 tumor DNA. Molecular cloning revealed the translocation breakpoint that fused a portion of the *Setd3* gene into the Jλ1-Cλ1 region containing the 3'Igλ enhancer (Figure 1C). The translocation junction occurred within intron 8 of *Setd3* and fused *Setd3* exons 9 through 13 to the intronic sequences just upstream of the Jλ1-Cλ1 region in a head-to-head configuration (Figure 1C and D). In this configuration, sense strands of these two genes are in opposite ori-

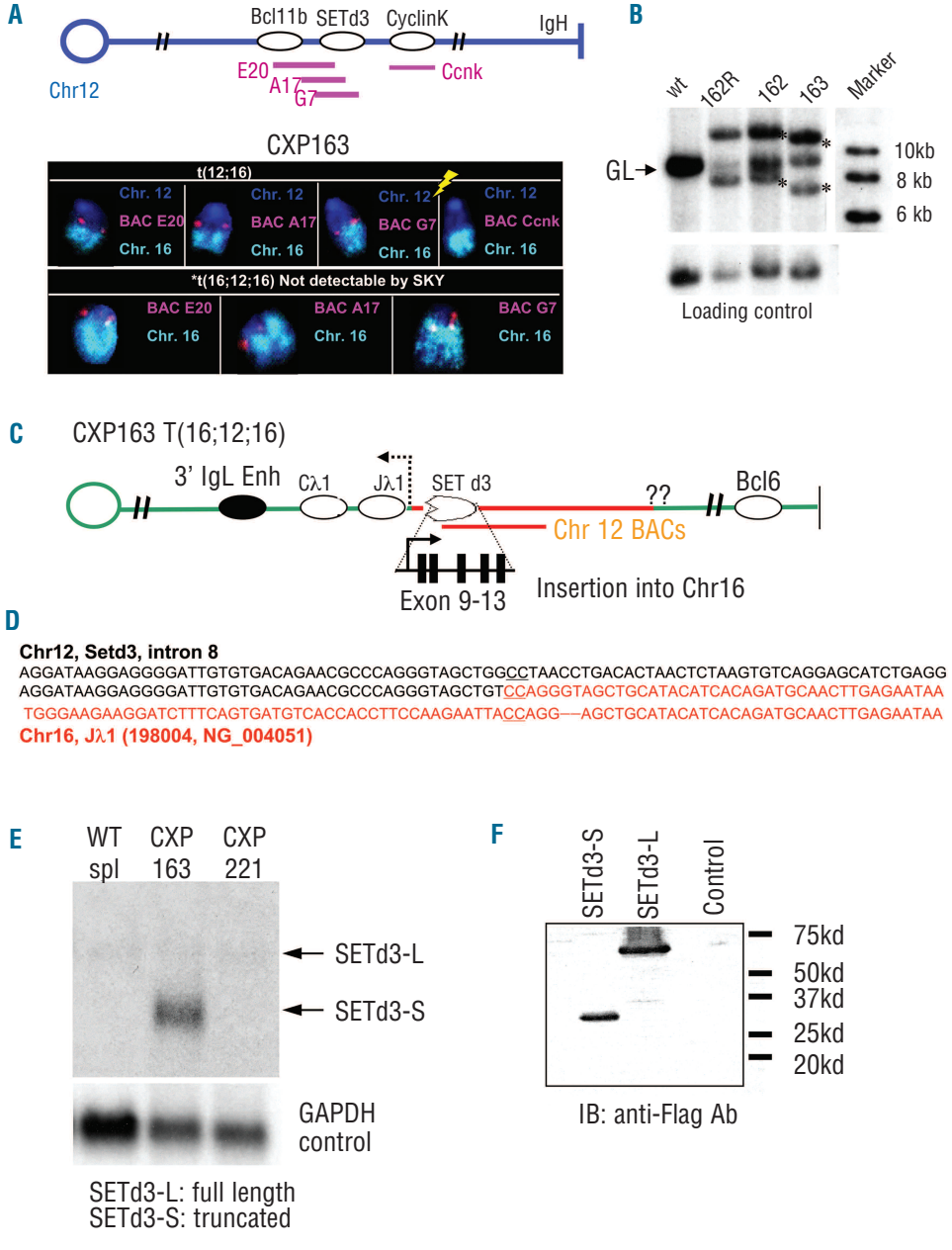

Figure 1. Molecular cloning of *Setd3* from an Igλ translocation derived from Xrcc4/p53 peripheral B-cell lymphoma. (A) Cytogenetic analysis of t(12;16) and complex t(16;12;16) translocations in CXP163. 3 BACs covering the *Setd3* region (E20, A17 and G7) and 1 BAC covering the *CcnK* region were used for FISH analysis. t(12;16) translocation occurred between BAC G7 and *CcnK*, involving the *Setd3* gene. The complex t(16;12;16) translocation has a large region of Chr12 including all 3 BACs (E20, A17 and G7) covering the *Setd3* gene inserted into Chr16. Chr16 was painted light green by whole chromosomal painting. (B) Southern blot analysis of CXP162 and CXP163. Clonal Igλ rearrangements were detected via Cλ1 probe. Germline (GL) bands are indicated by arrow. (C) Schematic map of t(16;12;16) translocation depicts the translocation breakpoints located in the sequence upstream of Jλ1 and in intron 8 of *Setd3* gene. The inserted region of Chr12 is shown in red and Chr16 in green. Arrow indicates the transcriptional orientation of the truncated *Setd3*. Dotted arrow indicates the transcriptional orientation of the germline Jλ1. Question marks indicate that the translocation breakpoint at this junction remains unknown. (D) The translocation junction was cloned by phage library and sequenced to map the breakpoints in the Jλ1-Cλ1 cluster and in intron 8 of the *Setd3* gene. Microhomology at the junctions is underlined. There were mutations and deletions identified at the junction. (E) Upregulation of *Setd3-S* in the CXP163 lymphoma by Northern blot analysis. RNA from wt splenocytes, CXP163, and CXP221 was purified and employed for Northern blot with *Setd3-S* cDNA as the probe. Arrows indicate the *Setd3-L* and *Setd3-S* transcripts, respectively, with the GAPDH transcript as a loading control. (F) Western blot was performed using cell extracts from 293T cells transfected with vectors expressing Flag-tagged *SETD3-S*, *SETD3-L*, or control vector with anti-Flag antibody (Ab).

entation, thereby excluding the possibility of forming a fusion product. Identifying the breakpoint of this complex Igλ translocation in CXP163 was part of a systematic study to clone the Igλ translocations/rearrangements in CXP lymphomas.[16] These studies eventually showed that CXP lymphomas arose from peripheral B cells that had attempted secondary V(D)J recombination of Igκ and Igλ light chain genes. Correspondingly, CXP tumors frequently harbored large chromosomal deletions or translocations involving Igκ or Igλ.[16] The t(12;16) translocation in the CXP163 lymphoma is not a balanced translocation, and the remainder of the distal portion of Chr12 was lost, as shown by SKY (*Online Supplementary Figure S1*). Notably, the genomic configuration of the region surrounding *Setd3* on mouse Chr12 is highly conserved with human Chr14q32.2, a critical region frequently involved in lymphomagenesis.[17] Thus, it is likely that there is a potential oncogenic consequence of the distal 12q deletion caused by the t(12;16) translocation.

Characterization of the Setd3 gene in tumor and normal tissues

The mouse *Setd3* gene is normally located on the telomeric end of Chr12 containing 13 exons (*Online Supplementary Figure S2A*). This Igλ translocation disrupted the *Setd3* gene, and the breakpoint occurred in intron 8, which is 1897bp upstream of exon 9. Using Northern blot analysis, we identified a truncated *Setd3* transcript (designated as *Setd3-S*) which accumulated to a much higher steady state level in CXP163 than the full length *Setd3* transcript (designated as *Setd3-L*) in wild-type (wt) splenic B cells or in other CXP tumors lacking this translocation (Figure 1E). Presumably, the putative promoter that drives the expression of *Setd3-S* is located within the 1897bp upstream of exon 9 of *Setd3*. In addition, the 3'Igλ enhancer in close vicinity might also contribute to the high expression of *Setd3-S* transcript. Next, we cloned the cDNA of *Setd3-S* using primers specific to the 5' end of exon 9 and 3' end of exon 13. This particular cDNA contains an open reading frame with the start codon (ATG) located within exon 10. Thus, we concluded that this Igλ translocation led to the production of a truncated *Setd3* transcript encoding the C-terminal portion of *Setd3*.

Lack of further material precluded the characterization of SETD3 protein in CXP163. To investigate whether both *Setd3* transcripts encode proteins with proper size, a Flag-tag was added to the C-terminal end of the *Setd3-L* and *Setd3-S* cDNAs. Cell extracts from *Setd3-L*, *Setd3-S*, or control vector transfected 293T cells were employed for Western blot analysis to detect the SETD3 proteins. Our

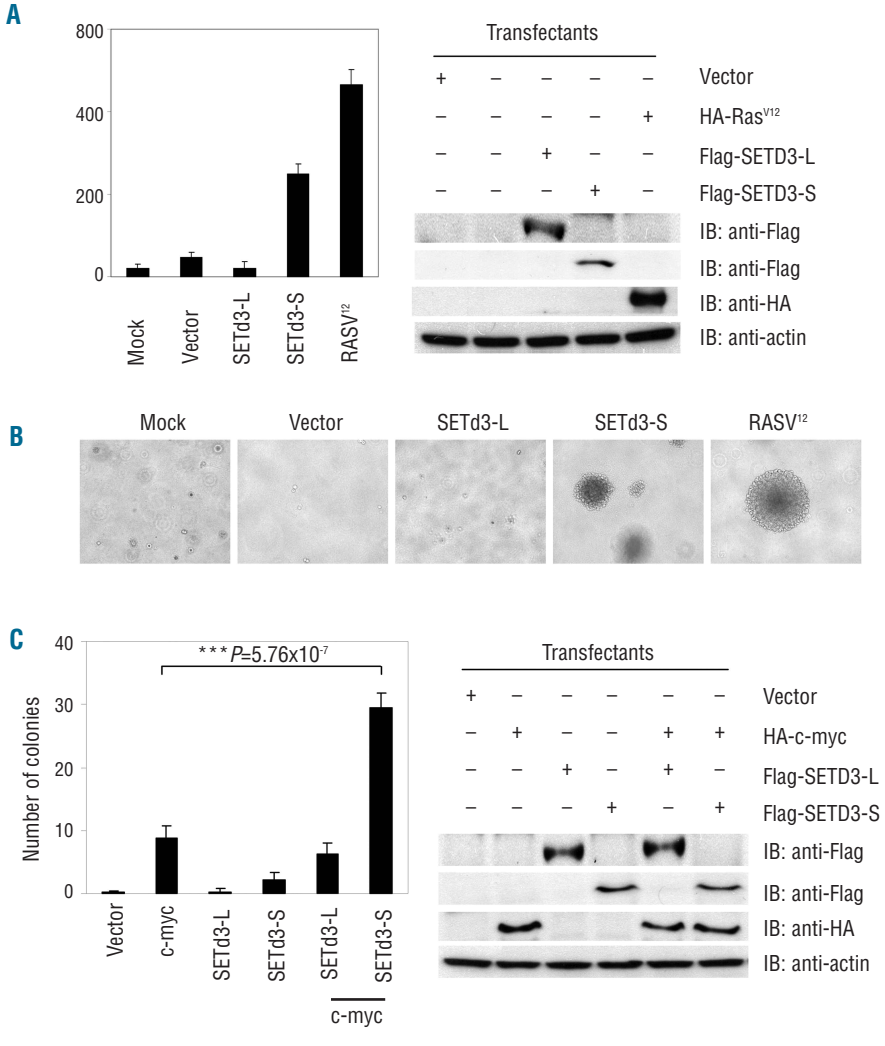

Figure 2. Oncogenic potential of the truncated SETD3-S. (A) Left: NIH3T3 cell lines were infected with different retroviruses expressing either RAS^{V12} (positive control), full length SETD3-L, truncated SETD3-S, or empty vector as a control. Mock infection control is also included. A soft agar assay was performed and colonies were counted 3-4 weeks after plating (n=4 plates per group). Representative data are shown as mean±S.D. from 3 independent experiments. (Right) Western blot for the expression of Flag-tagged SETD3-S or SETD3-L and HA-tagged RAS. (B) Representative images of colonies from all different groups are shown. Only SETD3-S or RAS^{V12} infected 3T3 cells had the ability to form colonies (anchorage-independent growth). (C) *Setd3-S* co-operates with *c-myc* to increase colony formation of primary MEFs. Primary MEFs were isolated from wt embryos and infected with different retroviruses expressing either c-MYC, full length SETD3-L, truncated SETD3-S, or empty vector as a negative control individually, or simultaneously with c-MYC and SETD3-L, or c-MYC and SETD3-S. (Left) A soft agar assay was performed and colonies were counted 3-4 weeks after plating (n=4 plates per group). Data are shown as mean±S.D. from 3 independent experiments. Statistical analysis: t-test, one-tail distribution with equal variance. (Right) Western blot for the expression of Flag-tagged SETD3-S or SETD3-L and HA-tagged c-MYC.

data confirmed that *Setd3-L* and *Setd3-S* cDNAs gave rise to proteins with the predicted size, about 67kd and 30kd, respectively (Figure 1F). There are multiple variants of *Setd3* transcripts identified in the database (see *Online Supplementary Appendix*) which are generated via alternative splicing.

We next examined the *Setd3* transcript in various tissues by Northern blot and found that the *Setd3-L* transcript was present ubiquitously in all the tissues tested using the GAPDH transcript as control (*Online Supplementary Figure S2B and C*). The *Setd3* gene encodes a protein with a highly conserved SET domain located at the N-terminus and a well-conserved RuBisCo lysine methyltransferase (LSMT) C-terminal substrate-binding domain (*Online Supplementary Figure S3*). Interestingly, the truncated SETD3-S protein lacks the highly conserved SET domain. The human homolog of SETD3 protein is 92% identical to its mouse counterpart (NCBI_Blastp). Thus, the Igλ translocation revealed a novel SET domain-containing protein, SETD3, which was also recently identified by independent studies using a bioinformatics approach.[18,19]

Role of Setd3 *in cell growth*

Since the expression of the truncated *Setd3-S* transcript was much greater in the CXP163 tumor, we next tested whether SETD3-S has oncogenic potential using a soft agar assay. Normally, NIH3T3 cells can not undergo anchorage independent growth and do not form colonies in soft agar. To test whether *SETD3-S* or *SETD3-L* can induce colony formation of NIH3T3 cells, we infected these cells with retroviruses expressing either *Setd3-S* or *Setd3-L* cDNA with a Flag-tag at the C-terminal end, or empty vector as a negative control. These analyses revealed that the cells infected with *Setd3-S* can form colonies at a much higher level compared to cells infected with empty vector or uninfected cells (Figure 2A and B). In contrast, the colony number in the *Setd3-L* infected group was similar to that in control groups (Figure 2A and B). The cells infected with retrovirus expressing Ras^{V12} with a C-terminal HA-tag served as a positive control that induced robust colony formation of 3T3 cells (Figure 2A and B). Western blot analysis showed the expression of corresponding exogenous proteins in the different groups of infected cells (Figure 2A). Thus, we conclude that the truncated *SETD3-S* has an oncogenic potential to transform NIH3T3 cells, whereas the full length *SETD3-L* does not. Our data also suggest that the highly expressed *Setd3-S* in CXP163 might contribute to lymphomagenesis. In this context, the CXP163 lymphoma not only highly expresses truncated *Setd3-S* transcript (Figure 1E) but also dramatically up-regulates expression of the c-myc oncogene[16] which is due to a t(12;15) translocation (*Online Supplementary Figure S1*) juxtaposing the c-myc oncogene next to the 3'Igh regulatory region.[16] Thus, we reasoned that the truncated *Setd3-S*, possibly by disrupting *Setd3-L* function, might co-operate with c-myc to promote oncogenesis. Next, we tested whether co-expression of c-myc and *Setd3-S* in primary mouse embryonic fibroblasts (MEFs) promotes colony formation. Primary MEFs were infected with retroviruses expressing *Setd3-S*, *Setd3-L* as described above, or c-myc with a C-terminal HA-tag, or empty vector. Western blot analysis showed the expression of exogenous proteins in the corresponding groups of infected 3T3 cells (Figure 2C). Our data showed that c-myc expression had a moderate effect on colony formation of primary MEFs, whereas both *Setd3-S* and *Setd3-L* had no obvious effects compared to the control group (Figure 2C). In contrast, when c-myc and *Setd3-S* were co-expressed, the number of colonies was significantly increased, whereas co-expression of *Setd3-L* with c-myc had no such effects (Figure 2C). Thus, our data demonstrate that *Setd3-S* may have the potential to function as an oncogene.

Functions of SETD3 *in histone modification*

To investigate whether SETD3 has any SET domain-dependent enzymatic activity on histones, we performed *in vitro* HMT activity assays. First, we expressed and purified recombinant Glutathione S-transferase (GST) only, GST-SETD3-L, and GST-SETD3-S proteins (Figure 3A). These purified proteins were assayed for their putative methyltransferase activity using native HeLa nucleosomes, containing all histones except H1, as substrates.[20] We found that only the purified GST-SETD3-L protein showed robust H3-specific HMT activities on the substrates (Figure 3B). In contrast, GST-SETD3-S or GST proteins displayed no detectable level of HMT activity (Figure 3B). When free histone octamers were used as substrates, we could not detect any HMT activity (*data not shown*), suggesting that SETD3-L only functions on the physiologically relevant substrates, i.e. native nucleosomes, of which H3 is a component. Thus, we conclude that the SETD3 protein is a fully active HMT for H3 with its activity dependent on the SET domain. Recently, independent studies also identified the zebrafish and mouse *Setd3* genes via a bioinformatics approach[18,19] which also

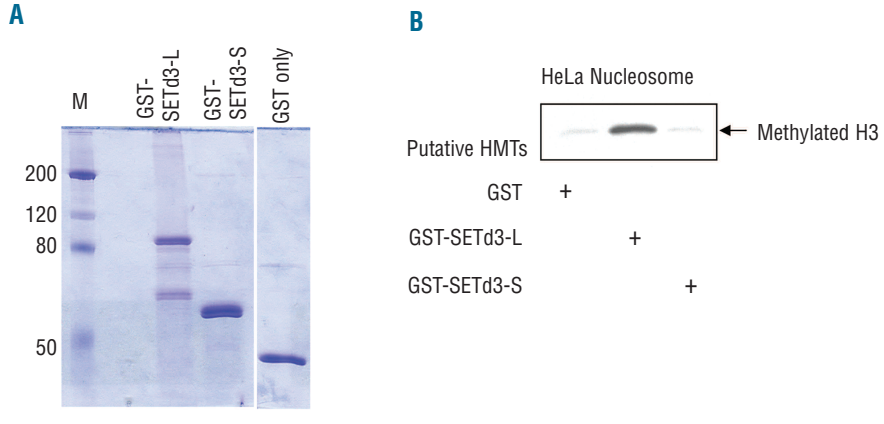

Figure 3. ***In vitro*** **HMT assay for SETD3. (A) Coomassie blue staining of the SDS-PAGE gel shows the purified GST-SETD3-L, GST-SETD3-S, and GST only proteins, with the protein marker loaded on the left (M). (B)** ***In vitro*** **HMT assays for GST-SETD3-L, GST-SETD3-S, and GST. Nucleosomes purified from HeLa cells were used as the substrate. Individual recombinant protein was added into the HMT assay as indicated (+). The final reaction was analyzed by SDS-PAGE followed by autoradiography. Methylation of H3 is indicated with an arrow.**

showed that SETD3 has HMT activity, in particular, on H3K4 and H3K36.[18] These studies used several H3 peptides instead of nucleosomes as substrates for the HMT assay. However, we could not detect HMT activity when free histone octamers were employed as substrates. It is likely that the activity of SETD3 on native nucleosomes, which are physiological substrates *in vivo*, is higher than that on free histone substrates. It remains to be determined whether SETD3 displays the same substrate specificity on native nucleosomes. In addition, SETD3 contains another well-conserved RuBisCo LSMT C-terminal substrate-binding domain. Rubisco LSMT is a chloroplast-localized SET domain-containing protein, which catalyzes the trimethylation of K14 in the large subunit of Rubisco,[21] an essential photosynthetic enzyme.[22] The RuBisCo LSMT C-terminal substrate-binding domain is also present in ribosomal lysine (K) methyltransferase (RKM) 1-4 of Saccharomyces cerevisiae.[23] RKM1 has been shown to methylate the ribosomal 23 subunit.[24] Thus, it remains possible that SETD3 might also have other non-histone substrates in the cytoplasm. Taken together, we propose that SETD3-L might function through its well-conserved SET domain and play an important role in suppressing tumor development while SETD3-S might function as a dominant negative mutant of SETD3-L and promote oncogenesis. Thus, it would be of interest to investigate the potential involvement of the *Setd3* gene in carcinogenesis.

Acknowledgments

We thank Dr. FW Alt (Children's Hospital Boston, HHMI, USA) for his generous support of this project and Dr. Y. Geng (Dana-Farber Cancer Institute, Boston, MA, USA) for advice on the soft agar assay and vectors including pBabe-puro-Ras^{V12} and pBabe-hygro-c-myc. We thank M Murphy for technical support with FISH and SKY. We apologize to those whose work was not cited due to length restrictions.

Funding

This work was supported by University of Colorado School of Medicine startup funds and Leukemia Research Foundation funds (to JHW). JHW is a recipient of the Boettcher Foundation Webb-Waring Biomedical Research Award. CTY is a recipient of the V Foundation Scholar Award, Emerald Foundation New Investigator Award, Kimmel Foundation Scholar Award, and ASH Scholar Award.

Authorship and Disclosures

Information on authorship, contributions, and financial & other disclosures was provided by the authors and is available with the online version of this article at www.haematologica.org.

References

1. Kornberg RD, Lorch Y. Twenty-five years of the nucleosome, fundamental particle of the eukaryote chromosome. Cell. 1999;98 (3):285-94.
2. Bhaumik SR, Smith E, Shilatifard A. Covalent modifications of histones during development and disease pathogenesis. Nat Struct Mol Biol. 2007;14(11):1008-16.
3. Tschiersch B, Hofmann A, Krauss V, Dorn R, Korge G, Reuter G. 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. 1994;13(16):3822-31.
4. Jenuwein T, Laible G, Dorn R, Reuter G. SET domain proteins modulate chromatin domains in eu- and heterochromatin. Cell Mol Life Sci. 1998;54(1):80-93.
5. Rea S, Eisenhaber F, O'Carroll D, Strahl BD, Sun ZW, Schmid M, et al. Regulation of chromatin structure by site-specific histone H3 methyltransferases. Nature. 2000;406 (6796):593-9.
6. Roguev A, Schaft D, Shevchenko A, Pijnappel WW, Wilm M, Aasland R, et al. The Saccharomyces cerevisiae Set1 complex includes an Ash2 homologue and methylates histone 3 lysine 4. Embo J. 2001;20(24):7137-48.
7. Peters AH, O'Carroll D, Scherthan H, Mechtler K, Sauer S, Schofer C, et al. Loss of the Suv39h histone methyltransferases impairs mammalian heterochromatin and genome stability. Cell. 2001;107(3):323-37.
8. Sanders SL, Portoso M, Mata J, Bahler J, Allshire RC, Kouzarides T. Methylation of histone H4 lysine 20 controls recruitment of Crb2 to sites of DNA damage. Cell. 2004;119(5):603-14.
9. Kim KC, Huang S. Histone methyltransferases in tumor suppression. Cancer Biol Ther. 2003;2(5):491-9.
10. Schneider R, Bannister AJ, Kouzarides T. Unsafe SETs: histone lysine methyltransferases and cancer. Trends Biochem Sci. 2002;27(8):396-402.
11. Ayton PM, Cleary ML. Molecular mechanisms of leukemogenesis mediated by MLL fusion proteins. Oncogene. 2001;20(40): 5695-707.
12. Huang S. Histone methyltransferases, diet nutrients and tumour suppressors. Nat Rev Cancer. 2002;2(6):469-76.
13. Bassing CH, Swat W, Alt FW. The mechanism and regulation of chromosomal V(D)J recombination. Cell. 2002;(Suppl 109):S45-55.
14. Lieber MR. The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. Annu Rev Biochem. 2010;79:181-211.
15. Yan CT, Boboila C, Souza EK, Franco S, Hickernell TR, Murphy M, et al. IgH class switching and translocations use a robust non-classical end-joining pathway. Nature. 2007;449(7161):478-82.
16. Wang JH, Alt FW, Gostissa M, Datta A, Murphy M, Alimzhanov MB, et al. Oncogenic transformation in the absence of Xrcc4 targets peripheral B cells that have undergone editing and switching. J Exp Med. 2008;205(13):3079-90.
17. Kamnasaran D, Cox DW. Current status of human chromosome 14. J Med Genet. 2002;39(2):81-90.
18. Eom GH, Kim KB, Kim JH, Kim JY, Kim JR, Kee HJ, et al. Histone methyltransferase SETD3 regulates muscle differentiation. J Biol Chem. 2011;286(40):34733-42.
19. Kim DW, Kim KB, Kim JY, Seo SB. Characterization of a novel histone H3K36 methyltransferase setd3 in zebrafish. Biosci Biotechnol Biochem. 2011;75(2): 289-94.
20. Dou Y, Milne TA, Tackett AJ, Smith ER, Fukuda A, Wysocka J, et al. Physical association and coordinate function of the H3 K4 methyltransferase MLL1 and the H4 K16 acetyltransferase MOF. Cell. 2005;121 (6): 873-85.
21. Trievel RC, Flynn EM, Houtz RL, Hurley JH. Mechanism of multiple lysine methylation by the SET domain enzyme Rubisco LSMT. Nat Struct Biol. 2003;10(7):545-52.
22. Feller U, Anders I, Mae T. Rubiscolytics: fate of Rubisco after its enzymatic function in a cell is terminated. J Exp Bot. 2008;59(7):1615-24.
23. Wlodarski T, Kutner J, Towpik J, Knizewski L, Rychlewski L, Kudlicki A, et al. Comprehensive structural and substrate specificity classification of the Saccharomyces cerevisiae methyltransferome. PLoS One. 2011;6(8):e23168.
24. Porras-Yakushi TR, Whitelegge JP, Clarke S. Yeast ribosomal/cytochrome c SET domain methyltransferase subfamily: identification of Rpl23ab methylation sites and recognition motifs. J Biol Chem. 2007;282(17): 12368-76.