▌慶北大農學誌(Agric. Rex. Bull. Kyungpook Natl. Univ.) 28(2010. 12) 63~68

Cloning of the Setd1b gene of *Mus musculus*, a novel histone methyl transferase target in the epigenetic therapy of cancers

Masayo Morishita, Minju Cho, Juhee Ryu, Damiaan E. H. F. Mevius
And Eric di Luccio*

[1]*School of Applied Biosciences, Kyungpook National University, Daegu, 702-701, Republic of Korea*

Abstract

The epigenetic therapy of cancers is emerging as an effective and valuable approach to both chemotherapy and the chemoprevention of cancer. The utilization of epigenetic targets that include histone methyltransferase (HMTase), Histone deacetylatase, and DNA methyltransferase, are emerging as key therapeutic targets. SET containing proteins such as the HMTase Setd1b has been found significantly amplified in cancerous cells. In order to shed some light on the histone methyl transferase family, we cloned the Setd1b gene from Mus musculus and build a collection of vectors for recombinant protein expression in E.coli that will pave the way for further structural biology studies. We prospect the role of the Setd1b pathway in cancer therapy and detail its unique value for designing novel anti-cancer epigenetic-drugs.

Key words : Transcription factor; Histone Methyl Transferase; HMTase inhibitor; Setd1b; Epigenetic Cancer Therapy

INTRODUCTION

Both genetic and epigenetic alterations of transcription factors are key features in carcinogenesis onset with aberrant gene functions and changes in gene expression levels (Varier et al, 2010). Transcription coactivators (TCs) are tightly regulated and can be the primary targets of hormonal control and signal transduction. TCsare responsible for biochemical activities such as modifying the chromatin, unwinding the DNA and recruiting the RNA polymerase. TCs can either activate or silence the transcription through the acetylation/de-acetylation, ubiquitination/de- ubiquitination, and/or methylation/de-methylations of specific amino acid residues on the histone tails H1, H2A, H2B, H3, and H4. Thus, among the superfamily of TCs, the family of histone methyl transferases (HMTase) has been focusing attention in the epigenetic therapy of various cancers.

Many HMTases such as NSD1, NSD2/MMSET/WHSC1, NSD3/WHSC1L1, Setd1b, Suv39h, Ezh2, Mll, Nsd1, Riz, and others have been implicated in numerous tumor developments and are generally found over-expressed in tumorous tissues (Schneider et al, 2002). Usually, reverting HMTase levels or inhibiting HMTase activity appears to be promising in suppressing cancer growth. The development of HMTase inhibitors

*Corresponding author. E-mail : diluccio@knu.ac.kr,
Phone : 82-53-950-5756, Fax : 82-53-950-6750
(Received October 14. 2010; Examined November 13, 2010; Accepted November 22, 2010)

is therefore emerging as an effective strategy in the epigenetic therapy of cancers. An increasing number of studies are revealing the role of NSD1/NSD2/NSD3in cancer development and proving as potential novel targets in cancer therapy.

Histone lysine methylation and the transcriptional regulation

Lysine methylation is one prominent feature of the posttranslational histone modifications in the regulation of chromatin structure and function. Lysine-HMTases target specific histone residues on H3 and H4, and can transfer one, two or three methyl groups on specific lysines on the histone tails. (Krogan et al, 2003). A large body of work has correlated the status of histone lysine methylation (mono, di, or tri) to certain cellular processes including transcriptional regulation. Histone lysine methylation (or any of the other histone modifications) can have both activating and repressive functions on the gene transcription. The various covalent histone modifications constitute an array of signals whose integration determines the fate of the transcription. However, this language of covalent histone modifications and its effect on the transcription remains unclear.

SETD1B

SET1B is a transcriptional co-activator and is a component of the SET1 histone methyltransferase (HMT) complex, at least composed of the catalytic subunit (SETD1BA or SETD1BB), WDR5, WDR82, RBBP5, ASH2L/ASH2, CXXC1/CFP1, HCFC1 and DPY30. SET1B is a histone methyl transferase enzyne that methylates the lysine 4 on the histone 3. However, SET1B becomes inactive in the SET1 HMT complex if the neighboring Lys-9 residue is already methylated. H3 Lys-4 methylation represents a specific tag for the epigenetic transcriptional activation (Butler et al, 2008; Lee et al, 2010). The non-overlapping localization of SET1 HMT with SETD1B suggests that SETD1A and SETD1B make non-redundant contributions to the epigenetic control of chromatin structure and gene expression. SETD1B is known to interact with HCFC1, ASH2L/ASH2 and WDR82 through a via its RNA recognition motif (RRM domain). Moreover, SETD1B interacts, via the RRM domain, with the hyperphosphorylated C-terminal domain (CTD) of RNA polymerase II large subunit (POLR2A) only in the presence of WDR82. Additionally, Setd1b is ubiquitously expressed (Butler et al, 2008; Lee et al, 2010; Tate et al, 2009; Tate et al, 2010; Thomson et

Table 1.

Primer Name (PK)	Primer sequences (5'-> 3') The restriction site is underlined.	mer	Direction	Restriction site
1	AAGGAGCATATGGCCAAGGCCTCACTG	27	Forward	NdeI
2	CCAGGACATATGCTGGAGCTGGACAGC	27	Forward	NdeI
3	GAGGAGCATATGACTGCAATGGCTGCA	27	Forward	NdeI
4	GAACATCATATGGAGTCAGACCTGGAC	27	Forward	NdeI
6	GCCCCACATATGTCACCTGAACCCTCA	27	Forward	NdeI
7	GGAGGAGAATTCCAAGGACTCCTTCTC	27	Reverse	EcoRI
8	CCTGTCGAATTCAAAGTCCTGGCCACC	27	Reverse	EcoRI
9	CGGGCTGAATTCCACCTGCGGGGAGCT	27	Reverse	EcoRI
10	GGACAGGAATTCCGGTGGGGATCTCCG	27	Reverse	EcoRI
12	ACGCGGGAATTCATTGAGGGTCCCCCG	27	Reverse	EcoRI
23	CTAGCTCATATGCCAACCCCCATCAAGAGG	30	Forward	NdeI
24	GTCGTAGAATTCGGTCATCTCCTCAAACTC	30	Forward	EcoRI

al, 2010; Xu et al, 2011).

Although the exact number of Lysine-HMTases in human is still uncertain, a multiple sequence alignment with 22 known human lysine-HMTases highlights the sequence heterogeneity of the HMTase superfamily at the entire sequence level and the Setd1b define a subfamily among them. The SET domains are conserved in other known HMTases (Cf Figure 4B), it is worth mentioning than the lysine HMTase Dot1 does not have a SET domain (Feng et al, 2002).

The SETD1B pathway as a novel target in the epigenetic therapy of cancers

Both genetic and epigenetic events police the initiation and the progression of cancers. Unfortunately, genetic alterations cannot be reverted. However, epigenetic abnormalities or impairments in DNA methylationor histone methylation, for instance, are potentially reversible. The utilization of epigenetic targets is therefore emerging as an effective and valuable approach to chemotherapy as well as chemoprevention of cancer. Since DNA methylation and histone tail modifications are the most studied epigenetic phenomena, numerous drugs (DNA-methylation inhibitors, histone-deacetylase and histone-methyltransferase inhibitors) are being developed, some of which are at clinical trial stages.

The level of histone acetylation is regulated by histone acetyltransferases (HATs) and histone deacetylases (HDACs). It has been shown that inhibitors of HDACs are able to alter the expression of certain genes and HDAC inhibitors are being used in cancer therapy. HMTases are emerging as valuable therapeutic targets as deregulated HMTase activity plays an important role in carcinogenesis events (Schneider et al, 2002). The development of HMTase inhibitors is raising new hopes in building a more effective drug-arsenal against a variety of cancers (Spannhoff et al, 2009). An increasing number of studies link deregulated NSD proteins to an array of cancers (cf. Table 1). Taken together, it appears that NSD1/NSD2/NSD3 can become the focus for a structure-based drug design for a wide variety of cancers using selective HMTase inhibitors. However, very few lead compounds have been described to selectively inhibit HMTase enzymes. This is mostly due to the lack of structural information on HMTase in order to support the design of selective inhibitors. In this study, we report the cloning details of Setd1b domains from Mus musculus.

MATERIAL AND METHODS

The Setd1b gene from *Mus musculus* was PCR-amplified from *Mus musculus* cDNA library using the primers listed in (Cf. Table 1 and Figure 1). In each case, the resultant fragment was inserted at NdeI and EcoRI sites of the multi cloning site (MCS) of the plasmid pTYB2 (New England Biolabs) to yield the final bacterial expression vector (Cf. Table 1, Figures 2 and 3). These were designed to produce protein with no extra residues on the C-terminus. Following, the presence of the insert in several candidates of each construct was confirmed by PCR using the forward primer from Setd1b and the Intein reverse primer of pTYB2. All clones were verified by sequencing to ensure that there were no mutations introduced by the PCR reaction. Resultants constructs were transformed into the *E. coli* expression strains ER2566 and BL21.

RESULTS AND DISCUSSION

Little is known about SETD1B analogously to other HMTase, the SETD1B protein is believed to carry multiple functional domains (HMTase, zinc fingers PHD and PWWP domains) separated by linkers of variable length (cf. Figure 1). However, the exact

function of each domain is still unknown. Following the methods used by others to study similar large proteins, we focus on each separate domain at a timeby cloning each functional domain of Setd1b (HMTase, Serine-rich region, Glutamate-rich region, Proline-rich region, Leucine-Proline rich region, histone methyl transferase domain) in order to shed some light on their functions.

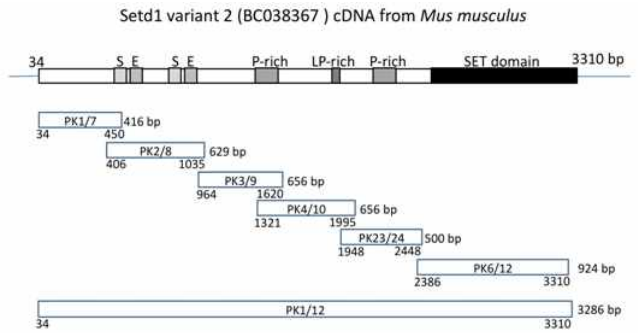

Figure 1. Cloning maps for the mouse Setd1b.

S: Serine-rich region; E: Glutamate-rich region; P: Proline-rich region; LP: Leucine-Proline rich region SET: histone methyl transferase domain

In eukaryotes, the dynamic transition between an extended transcriptionally-active euchromatine and a compact, transcriptionally silent heterochromatin structure is critical for the regulation of gene expression. Post-translational changes within the N-terminal tails of the histones play a key role in the reorganization of the chromatin structure and thus the regulation of gene expression. Several post-translational modifications of the N-terminal tails of the histones including acetylation, phosphorylation, ubiquitinylation, and methylation have been reported. Acetylation of the lysine residues at the N-terminus of histone proteins removes positive charges, thereby reducing the affinity between histones and DNA. This makes RNA polymerase and transcription factors easier to access the promoter region. Histone acetylation enhances transcription while histone deacetylation represses transcription. Histone methylation, carried by the SET domain, induces a tighter affinity between histone and DNA, thus represses the transcription, and is a major player in the regulation of the gene activity.

The PHD zinc finger is typically a domain of about 60 amino acids that fold around one or more zinc ions and is found in over 400 eukaryotic proteins, many of which are involved in the regulation of gene expression and in the maintenance of chromatin structure. PHD zinc finger domains typically show a C4HC3 signature (four cysteines, one histidine, three cysteines) with characteristic cysteine spacing and with additional conserved residues, most notably a tryptophan or other aromatic amino acid preceding the final cysteine pair. Studies have suggested a role for PHD fingers as nucleosome interaction determinants. However the function of the PHD finger is still elusive and controversial, as a variety of functions have been suggested, including phosphoinositide binding and E3 ubiquitin ligase activity (Gurevich et al, 2007; Sue et al, 2004). In addition to their role as a DNA-binding module, PHD zinc finger have been shown to mediate protein-protein and protein-lipid interactions as well

The PWWP domain is named after a conserved Pro-Trp-Trp-Pro motif. The functions of this domain are still unclear but it appears to be diverse. Analogously to PHD zinc fingers, PWWP can bind either proteins or DNA counterpart. Interestingly, PWWP domains are found associated with other HMTase like MSH6. Msh6 is a human mismatch repair protein and is involve in the correction of DNA errors. Msh6 is an essential component in maintaining the genome stability. Mutations in MSH6-PWWP domain are linked to increased cancer susceptibility particularly in human colorectal cancers. Moreover, a recent study on Pdp1 identified as SET9 (spKMT5) HMTase regulatory partners, showed that the Pdp1-PWWP domain is involved in methyl-lysine recognition (Liu et al, 2008). SET9 (spKMT5) relies on the PWWP domain of Pdp1 to recognize H4K20me and thus is key for Set9 chromatin localization. The functions of the two PWWP domains in the NSD family are still unknown and only

the identification of the NSD binding partners will bring clear answers.

An increasing number of SET containing enzymes are found amplified in a subset of cancers that include acute myeloid leukemia, multiple myeloma, and lung/prostate/breast cancers. HMTases are emerging as valuable therapeutic targets in the epigenetic therapy of cancers. However, very few lead compounds are capable of specifically inhibiting HMTase enzymes. New drugs specifically targeting Setd1b would not only be attractive for the treatment of numerous human cancers, but also bring crucial insights into the design of selective HMTase inhibitors.

Here we report the first essential steps toward the understanding of the Setd1b pathways and the associated dysfunctions, especially cancers. We succeeded in cloning the functional domains along with the whole Setd1b gene from *Mus musculus* (Cf. Figures 2, 3 and 4).

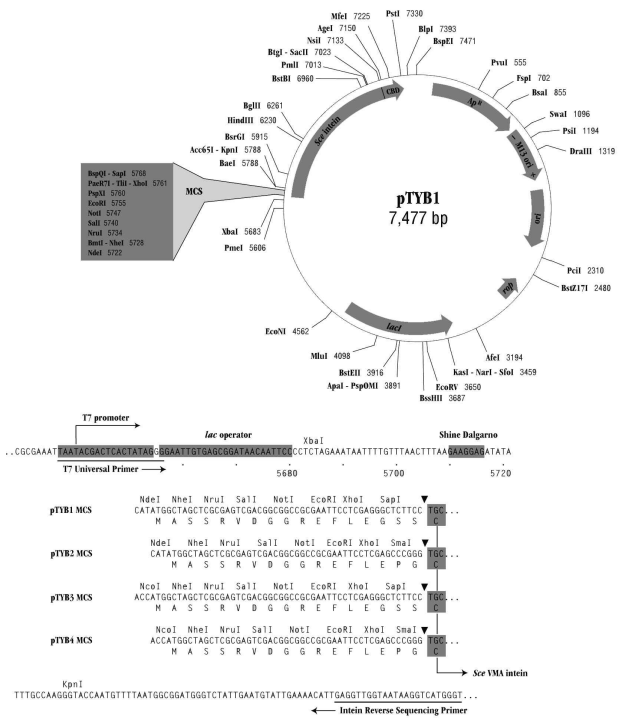

Figure 2. pTYB1 an -2 are E. coli plasmid-cloning-vectors designed for recombinant protein expression and purification using the IMPACT Kit from New England Biolabs (NEB)

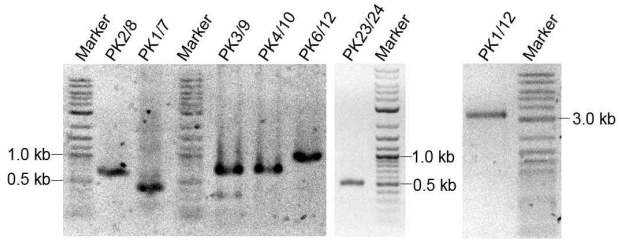

Figure 3. PCR amplification of PK2/8, PK1/7, PK3/9, PK4/10, PK6/12 PK23/24 and PK1/12 sequences corresponding at different fragment of the Nsd1 gene.

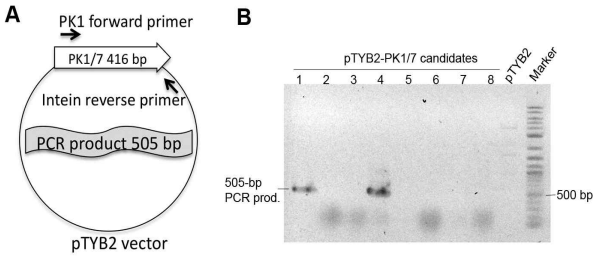

Figure 4. Construct validation and screening of the candidates.

Example of the PK1/7 construct

It is the first time the cloning of the entire gene of a Setd1b family member has ever been reported. Recombinant protein expression and purifications are now well underway.

ACKNOWLEDGEMENTS

The National Research Foundation of Korea and the Kyungpook National University supported the research presented in this paper.

REFERENCES

1. Butler JS, Lee JH, Skalnik DG 2008 CFP1 interacts with DNMT1 independently of association with the Setd1 Histone H3K4 methyltransferase complexes. DNA Cell Biol 27: 533-543
2. Feng Q, Wang H, Ng HH, Erdjument-Bromage H, Tempst P, Struhl K, Zhang Y 2002 Methylation

of H3-lysine 79 is mediated by a new family of HMTases without a SET domain. Curr Biol 12: 1052-1058

3. Gurevich I, Flores AM, Aneskievich BJ 2007 Corepressors of agonist-bound nuclear receptors. Toxicol Appl Pharmacol 223: 288-298
4. Krogan NJ, Kim M, Tong A, Golshani A, Cagney G, Canadien V, Richards DP, Beattie BK, Emili A, Boone C, Shilatifard A, Buratowski S, 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
5. Lee JH, You J, Dobrota E, Skalnik DG 2010 Identification and characterization of a novel human PP1 phosphatase complex. J Biol Chem 285: 24466-24476
6. Liu X, Wang L, Zhao K, Thompson PR, Hwang Y, Marmorstein R, Cole PA 2008 The structural basis of protein acetylation by the p300/CBP transcriptional coactivator. Nature 451: 846-850
7. Schneider R, Bannister AJ, Kouzarides T 2002 Unsafe SETs: histone lysine methyltransferases and cancer. Trends Biochem Sci 27: 396-402
8. Spannhoff A, Sippl W, Jung M 2009 Cancer treatment of the future: inhibitors of histone methyltransferases. Int J Biochem Cell Biol 41: 4-11
9. Sue SC, Chen JY, Lee SC, Wu WG, Huang TH 2004 Solution structure and heparin interaction of human hepatoma-derived growth factor. J Mol Biol 343: 1365-1377
10. Tate CM, Lee JH, Skalnik DG 2009 CXXC finger protein 1 contains redundant functional domains that support embryonic stem cell cytosine methylation, histone methylation, and differentiation. Mol Cell Biol 29: 3817-3831
11. Tate CM, Lee JH, Skalnik DG 2010 CXXC finger protein 1 restricts the Setd1A histone H3K4 methyltransferase complex to euchromatin. FEBS J 277: 210-223
12. Thomson JP, Skene PJ, Selfridge J, Clouaire T, Guy J, Webb S, Kerr AR, Deaton A, Andrews R, James KD, Turner DJ, Illingworth R, Bird A 2010 CpG islands influence chromatin structure via the CpG-binding protein Cfp1. Nature 464: 1082-1086
13. Varier RA, Timmers HT 2010 Histone lysine methylation and demethylation pathways in cancer. Biochim Biophys Acta
14. Xu C, Bian C, Lam R, Dong A, Min J 2011 The structural basis for selective binding of non-methylated CpG islands by the CFP1 CXXC domain. Nat Commun 2: 227