A novel DNMT3B subfamily, ΔDNMT3B, is the predominant form of DNMT3B in non-small cell lung cancer

LUO WANG1, JIE WANG5, SHIYONG SUN3, MARIVONNE RODRIGUEZ2, PING YUE3, SE JIN JANG4 and LI MAO1

1Molecular Biology Laboratory, Department of Thoracic/Head and Neck Medical Oncology, 2Department of Molecular Genetics, The University of Texas M.D. Anderson Cancer Center, Houston, TX 77030; 3Department of Oncology, Winship Cancer Institute, Emory University, Atlanta, GA, USA; 4Department of Pathology, Hanyang University, Kuri Hospital, 471-701 Seoul, Korea; 5Department of Thoracic Medical Oncology, Peking University School of Oncology and Beijing Cancer Hospital, Beijing, P.R. China

Received February 3, 2006; Accepted February 27, 2006

Abstract. De novo promoter DNA methylation represses gene transcription and is a common mechanism to inactivate tumor suppressor genes in tumorigenesis. DNMT3B plays an important role in de novo DNA methylation. We report here the identification of a novel DNMT3B subfamily, termed ΔDNMT3B, whose expression is initiated through a promoter located at intron 4 and exon 5 of the DNMT3B gene. At least 7 transcriptional variants of ΔDNMT3B have been observed as the result of alternative pre-mRNA splicing. Predicted proteins derived from these variants suggest that 4 of the variants share a conservative enzymatic domain due to predicted premature translational termination. In non-small cell lung cancer (NSCLC) cell lines, ΔDNMT3B variants are frequently expressed and are the predominant forms of ΔDNMT3B. Similarly, ΔDNMT3B variants are frequently expressed in primary NSCLC but are not detectable or are expressed at low levels in corresponding normal lung tissue. Our results indicate that ΔDNMT3B is the major expression form of DNMT3B in NSCLC and may play an important role in the development of aberrant promoter methylation during lung tumorigenesis.

Introduction

DNA methylation plays an essential role in normal development of mammalian embryo by regulating gene transcription through genomic imprinting, X chromosome inactivation, and genomic stability (1-4). It is believed that DNA methylation patterns in somatic cells are established during gametogenesis and early embryonic development via consecutive waves of demethylation and de novo methylation (5). DNMT3 consists of DNMT3A and DNMT3B and has been shown to be the major de novo DNA methyltransferase (6,7) that preferentially methylates cytosine in CpG sites. The methylation in CpG-rich promoter regions would result in transcriptional silencing of the corresponding genes.

Okano et al found that murine Dnmt3a and Dnmt3b are highly expressed in the undifferentiated ES cells but are down-regulated during development and maintained at a low level in somatic cells (6). They further revealed that, at E7.5, Dnmt3b was highly expressed in the embryonic ectoderm, neural ectoderm, and chorionic ectoderm while a weak expression was detected in mesodermal and endodermal cells (8). The expression of Dnmt3a was moderate in embryonic ectoderm and weak in mesodermal cells until E8.5 and E9.5, at which point Dnmt3a expression became ubiquitous with increased expression in the somites and the ventral part of the embryo (8). These observations suggest that the two types of enzymes may function differently in development.

Human DNMT3B is highly homologous to the mouse gene and contains 24 exons spanning approximately 47 kb of genomic DNA. Two alternative 5' exons of DNMT3B, both resulting in full-length DNMT3B (DNMT3B1 and DNMT3B2), have been reported (9). Three transcription variants resulting from alternative splicing have also been reported (DNMT3B3-5) (9). Some of the variants lacking DNA methyltransferase activity may compete with variants with enzyme activity resulting in DNA hypomethylation (10), suggesting a complex role of DNMT3B variants. Increased expression of DNMT3B has been frequently observed in human cancer cell lines and primary tumors compared to most normal tissues except testis,
pancreas, thyroid, and bone marrow (9-11). However, the level of DNMT3B expression in tumors does not consistently correlate with the promoter methylation status of genes (11-13). We report here the identification of a new class of DNMT3B transcripts expressed through a novel promoter, termed ΔDNMT3B. We further demonstrate that ΔDNMT3B is the predominant form of DNMT3B in NSCLC, suggesting an important role of ΔDNMT3B in DNA methylation control and lung tumorigenesis.

Materials and methods

Cell lines and primary tissues. Human NSCLC lines H157, H226, H292, H460, H522, H1299, H1944, and A549 were purchased from the American Type Cell Culture (Rockville, MD). Cells were cultured in DMEM supplemented with 10% heat-inactivated FCS, 2 mM L-glutamine, 100 IU/ml penicillin, and 100 μg/ml streptomycin at 37˚C in the presence of 5% CO2.

Twelve paired primary tumor tissues and corresponding normal lung tissues from patients with primary NSCLC were obtained from surgically resected specimens collected in the Department of Pathology at The University of Texas M.D. Anderson Cancer Center and stored at -80˚C until the experiment. The study was approved by the Institutional Review Boards of The University of Texas M.D. Anderson Cancer Center.

RNA extraction and RT-PCR. The total RNA for each cell line and clinic sample was isolated using Tri reagent (Molecular Research Center, Inc., Cincinnati, OH) according to the manufacturer’s instructions. Approximately 1 μg of total RNA from each sample was used to conduct a reverse transcription reaction in a 20-μl volume using SuperScript II RNase H-reverse transcriptase (Gibco BRL Life Technologies Inc., Grand Island, NY). The synthesized cDNA was either used immediately for PCR amplification or stored at -20˚C for further analysis.

The mRNA expression levels of total DNMT3Bs were detected using a primer set of the forward primer, S1 (5’-GAG TTG GGC ATA AAG GTA GG-3’), and the reverse primer, 1AS1 (5’-TGA GGT ACA CGG TAT GAC C-3’), located at exon 17 and the 3’-untranslation region of DNMT3B1, respectively. The 5’-end forward primers, E1 (5’-CAT GAA GGG AGA CAC CAG GC-3’), E3 (5’-ATG CCA AAG CTC TTC CGG GA-3’), and E5 (5’-TGG AGA TGG AGA CAG TTC AG-3’), and the reverse primer, 1AS1, were also used to detect ΔDNMT3Bs (Fig. 1A).

PCR was performed in a 12.5-μl volume containing 0.5 μl of reverse transcription products, 7% dimethylsulfoxide, 1.5 mM deoxynucleotides (dNTPs), 6.7 mM MgCl2, 16.6 mM (NH4)2SO4, 67 mM Tris, 10 mM β-mercaptoethanol, 6.7 μM ethylene diamine tetra acetate acid (EDTA), 0.5 μM of both the forward and the reverse primer, and 0.625 U of HotStar Taq DNA polymerase (Qiagen, Inc., Chatsworth, CA). Amplification was carried out in a thermal cycler (PCR Express, Hybaid, Middlesex, UK) with an initial denaturing step at 95˚C for 15 min, followed by 35 cycles at 95˚C for 30 sec, 58-62˚C for 1 min, and 72˚C for 1 min, with a last extension step at 72˚C for 10 min. The PCR products were mixed with 6X loading buffer containing 0.5 mg/ml ethidium bromide and separated by electrophoresis on a 2% agarose gel.

Primer extension and nuclease S1 mapping. To determine the exact starting site of the ΔDNMT3B transcript, standard primer

Figure 1. Identification of ΔDNMT3B. (A) Expression of the DNMT3B 5’-region measured using RT-PCR with different 5’ primers located at exon 2 (E1), exon 4 (E2), and exon 6 (E3) of DNMT3B1, respectively. (B) Primer extension analysis showing transcription starting points of ΔDNMT3B.
extension and S1 mapping methods were used with the [$\gamma$-$^3$P]-ATP end-labeled antisense primer, 3B6AS (5'-GGT AGC CGG GAA CTC CAC GC-3'). For primer extension, briefly, 1 μg of total RNA was mixed with [$^3$P]-labelled primer. The mixture was incubated at 70°C for 15 min and then at room temperature for 10 min. Extension reactions (20 μl) consisted of 50 mM Tris (pH 8.3), 75 mM KCl, 3 mM MgCl₂, 10 mM dithiothreitol, 1 mM each dNTP, and 200 U SuperScript™ II reverse transcriptase (Gibco BRL Life Technologies Inc.). Reactions were incubated at 37°C for 15 min. The products were mixed with loading buffer (95% formamide, 20 mM EDTA, 0.05% bromophenol blue and 0.02% xylene cyanole FF), denatured at 98°C for 5 min, and then separated on a 12% acrylamide-7 M urea denatured gel. Radioactive signals were detected by autoradiography.

Nuclease S1 mapping was performed with a 1080-bp DNA fragment which was amplified using the forward primer, E4INT-1 (5'-TGC TAA GCT ACA AGG CAC AGG AG-3'), and the reverse primer, E5AS (5'-TCT GTG TCG TCT GTG AGG TC-3'). After cloning this fragment into a PCR ® 2.1-TOPO® plasmid vector (Invitrogen Corp., Carlsbad, CA), a 320-bp fragment which was amplified using the forward primer, E1 (5'-TGC TAA GCT ACA AGG CAC AGG AG-3'), and 5AS (5'-GAG ACA CAT GTA ACA GCT CC-3'). This fragment contains, as the internal primer 3B6AS. The PCR condition was the same as above. This single-strand 320-bp PCR product was separated in a 2% agarose gel and purified using a QIAquick gel extraction kit (Quagen Inc.) followed by recovering in 50 μl Tris-buffer (10 mM Tris-HCl, pH 8.5).

The total RNA from different samples was co-precipitated with 50 ng of recovered 320-bp [$^3$P]-labelled probe. Samples were dissolved in 30 μl of hybridization buffer (40 mM MOPS, pH 6.4, 1 mM EDTA, 0.4 M NaCl and 80% formamide) and incubated at 85°C for 15 min. After overnight hybridization at 54°C based on the GC content of the projected fragment, the samples were digested for 1 h at 37°C with S1 nuclease (Gibco BRL Life Technologies Inc.) in the buffer containing 30 mM sodium acetate, pH 4.6, 1 mM zinc acetate, 5% glycerol, and 0.28 M NaCl. The resulting products were detected as described in primer extension section.

Construction of DNMT3B6 promoter and luciferase assay. The first 1080-bp $\Delta$DNMT3B promoter was amplified with the primer set of E4INT-1 and E5AS. This fragment contains, 355 bp upstream of the $\Delta$DNMT3B transcription starting site, the first exon and intron of $\Delta$DNMT3B. After inserting the fragment into the pGL3-basic vector (Promega Corp., Madison, WI), the plasmids containing both forward (F) and reverse (R) directions were used for transient transfection. To compare the functional difference of C/T polymorphism in the $\Delta$DNMT3B promoter region, both pGL3 T-type and C-type promoters were constructed.

Lung cancer cell line A549 and H157 (ATCC, Manassas, VA) was used for transient transfection using FuGene 6 transfection reagent (Roche Diagnostics Corp., Indianapolis, IN) according to the manufacturer's instructions. The plasmid, pCH110 (Amersham Pharmacia Biotech, Inc., Piscataway, NJ), was used as the internal control for monitoring the transfection efficiency. The signal was detected using a luciferase assay system (Promega Corp.) in a luminometer (Lumat 9507, EG&G Berthold). The values of luciferase activity were normalized against those of β-galactosidase expressed by plasmid pCH110.

Detection of individual $\Delta$DNMT3B splicing variants. The expression levels of specific DNMT3B and $\Delta$DNMT3B variants in NSCLC cell lines and primary tissues were determined using specific primer sets corresponding to individual DNMT3B and $\Delta$DNMT3B variants. For $\Delta$DNMT3B1, we used 1S, 5'-TGG AAG GCC ACC TCC AAAG C-3', as the forward primer and 1AS, 5'-GCC TGC AGC AGC CAT CG-3', as the reverse primer; for $\Delta$DNMT3B2, 2S 5'-AGA TCA AGG GCT TCT CCT GG-3' and 2AS 5'-GAG TCT TGT TCT GTG GTC GCG-3'; for $\Delta$DNMT3B3, 3S 5'-GTT CAG AGT ATC AGG TCT CTC C-3' and 1AS; for $\Delta$DNMT3B4, 3S and 2AS; for $\Delta$DNMT3B5, 4S 5'-GTT CAG AGT ATC AGA GAA CAA GAC-3', and 3AS 5'-CTG CCA CAA GAC AAA CAG CC-3'; for $\Delta$DNMT3B6, 5S 5'-GTC TTG CTC GAG AGC AAC AGC AC-3', and 4AS 5'-CAG TAA GAC TAG CAA GAC TCG-3'; and for $\Delta$DNMT3B7, 6S 5'-TGC TCT GGA GAG AAC AGC AC-3', and 5AS 5'-GAG ACA CAT GTA ACA GCT CC-3'. A common forward primer, E1 5'-TGC TAA GCT ACA GAC AGG AC-3', was used for the DNMT3B variants, and specific reverse primers were used to distinguish individual variants as follows: 1AS for DNMT3B3 corresponding to $\Delta$DNMT3B1 and $\Delta$DNMT3B3; 2AS for DNMT3B3s corresponding to $\Delta$DNMT3B2 and $\Delta$DNMT3B4; 6AS, 5'-CGA GTC TTG TTC TCT GAT ACT C-3', for $\Delta$DNMT3B6 corresponding to $\Delta$DNMT3B5; 7AS, 5'-CGA GTC TTG TTC TCT GAT ACT C-3', for $\Delta$DNMT3B6; and 8AS, 5'-CGA GTC TTG TTC TCT GCA GAC-3', for $\Delta$DNMT3B7.

Results

To determine the expression levels of DNMT3B1 in normal lung tissue and lung cancer tissue, we analyzed 12 pairs of primary NSCLC tissue and corresponding normal lung tissue using RT-PCR with a set of primers located at exon 17 and exon 23 of DNMT3B respectively. We found that DNMT3B1 expression was either undetectable or at trace levels in the vast majority of normal lung tissue analyzed while it was detectable in all NSCLC tissue with 50% (6/12) of the tumors expressing at a high level. To validate the finding, we designed several additional sets of primers that allowed us to amplify DNMT3B1 mRNA at different exon locations closer to its transcriptional initiation site. We found that the expression level of the gene was much lower when using a primer located at exon 2 (E1) or exon 4 (E4) compared to a primer located at exon 6 (E5) of DNMT3B1 (Fig. 1A), suggesting the presence of additional transcripts excluding exons 2-4. To confirm this observation, we tested other primer sets at these regions and the results were consistent with the previous observation (data not shown). To exclude the possibility of contamination with homologue molecules in the RT-PCR products, we performed direct sequencing analysis of each RT-PCR product. The sequences matched perfectly to the originally reported corresponding transcript sequence of DNMT3B1 (GenBank accession number (AN): AL035071).

To determine the exact starting point(s) of the novel transcripts, we performed primer extension assay using RNA
templates from lung cancer tissue and NSCLC cell lines and a primer (3B6AS) located at exon 5 of DNMT3B1. We identified two major transcriptional initiation sites located at nt 23990 and nt 23994 within exon 5 of DNMT3B1 (GenBank AN: 15306493), respectively (Fig. 1B). This partial exon 5 was then named as the first exon of the novel transcript(s) containing either 28 bp or 24 bp depending on which transcriptional initiation site it derives from. We further validated the finding by using nuclease S1 RNA mapping analysis (data not shown). We designated the new transcript from these

Figure 2. Promoter of ΔDNMT3B. (A) Activity of ΔDNMT3B promoter detected using luciferase assay. (B) Effect of a T→C transition (polymorphism) on promoter activity of ΔDNMT3B. The larger arrow indicates the position of the polymorphism. Based on the polymorphism, the promoters are defined as C-type or T-type. F indicates forward direction of the promoter and R indicates reverse direction of the promoter.

Figure 3. Alternative splicing of ΔDNMT3B. (A) High level expression of ΔDNMT3B variants in lung cancer tissue compared to corresponding normal lung tissue. (B) The structure scheme of DNMT3B1 and ΔDNMT3B variants. N indicates normal tissue and T indicates cancer tissue.
starting sites as ΔDNMT3B because it lacks exons 1-4 of DNMT3B.

To determine the existence of a potential promoter upstream of the newly identified transcript, we constructed a 1080-bp DNA fragment containing 355 bp upstream of the ΔDNMT3B transcriptional initiation site, the first exon and intron of ΔDNMT3B and partial exon 2 into a vector containing a reporter gene. We performed a promoter activity assay and found promoter activity of the DNA fragment (Fig. 2A). We further constructed serial plasmids with both sense and reverse sequences of the DNA fragment and various deletions. Using these constructs, we found that the core promoter activity of ΔDNMT3B is in a 477-bp fragment containing one repressor element and three cis-acting elements (data not shown).

Interestingly, a common T→C transition polymorphism was found in the promoter region of ΔDNMT3B located -286 bp from the transcriptional initiation site, which may change a TFIID (CTcTAT TCCA) binding site to GATA-1 (TCTAT C) binding site. We noticed a stronger promoter activity with the T form than the C form (18-fold vs. 12-fold compared to the control, respectively) (Fig. 2B).

Because of the presence of various sizes of RT-PCR products using these primer sets, we suspected that ΔDNMT3B might contain multiple splicing variants (Fig. 3A). By directly sequencing the individual fragments, we found at least seven variants (by including or excluding different combinations of exons 3, 4, 5, and 6 of ΔDNMT3B), which we designated as ΔDNMT3B1-7 (Fig. 3B). A comparative analysis of the putative amino acid sequences of the variants showed that ΔDNMT3B lacked 199 amino acids at the N-terminal, compared with DNMT3B1, and that ΔDNMT3B1 and ΔDNMT3B2 contained a complete PWWP motif; other variants either contained a partial PWWP motif or no such structure; ΔDNMT3B5-7 lacks the enzymatic domains because of a premature termination that results from frame shifting after alternative splicing (Fig. 3B).

To detect individual ΔDNMT3B variants, we designed specific PCR primer sets on the basis of their splicing patterns (Fig. 4A). We detected ΔDNMT3B1 and ΔDNMT3B2 expression in all 13 NSCLC cell lines analyzed, ΔDNMT3B4 expression in 12 of the 13 cell lines, and ΔDNMT3B6 expression in 7 of the 13 cell lines; in contrast, expression of ΔDNMT3B3, ΔDNMT3B5, and ΔDNMT3B7 were less frequent (Fig. 4B). Interestingly, the DNMT3B variants were expressed less frequently and at lower levels in these cell lines (Fig. 4C). In a multiplex PCR analysis, we determined

---

**Figure 4.** Alternative or aberrant splicing variants of ΔDNMT3B subfamily. (A) Location of the primers used to amplify individual ΔDNMT3B variants in this study. (B) Expression patterns of ΔDNMT3B variants in NSCLC cell lines; 1-7 represent ΔDNMT3B1-7, respectively. (C) Expression patterns of DNMT3B, with more proximal exons corresponding to ΔDNMT3B1-4 and ΔDNMT3B6. (D) Multiplex PCR using primer sets for DNMT3B1 and ΔDNMT3B1 with different ratios in concentration (concentration of DNMT3B1 primer set was serially diluted from 1 to 9 and serially increased from 1 to 9; 1, DNMT3B1 primer set alone; 9, ΔDNMT3B1 primer set alone; 5, equal concentrations for both primer sets). The upper band represents the DNMT3B product and the lower band represents the ΔDNMT3B1 product. The lower panel shows relative intensity of the product bands.
the relative amplification efficiencies of the primer sets for DNMT3B and ΔDNMT3B using DNA templates containing various concentrations of DNMT3B and ΔDNMT3B (Fig. 4D). The robust amplification of DNMT3B1 (Fig. 4D) indicated that the lack or very low levels of RT-PCR products in the cell lines reflected the low level of the corresponding DNMT3B transcripts.

To determine whether the expression profiles observed in the NSCLC cell lines also exist in the primary NSCLC, we analyzed 12 pairs of primary NSCLC tumors and corresponding normal lung tissue. We found that the expression of ΔDNMT3Bs but not DNMT3Bs was frequent in the primary tumors whereas the expression of the variants was mainly non-detectable or weakly expressed in the corresponding non-cancerous lungs (Fig. 5). Nine (75%) of the 12 tumors expressed at least one variant, similar to the cancer cell line data.

Discussion

DNA methyltransferases play an important role in initiation and maintenance of cytosine methylation in human genome. Although the role of DNMT1 is primarily to maintain DNA methylation status and the role of DNMT3s relates to de novo DNA methylation, recent studies have shown that these molecules not only possess distinct activities (6,8,14) but also interact with each other in complex biological processes to regulate patterns of DNA methylation in human genome (15-18). The dominant-negative effect of DNMT3b4, which lacks methyltransferase enzymatic motifs, in competing with DNMT3b3 to result in DNA hypomethylation on pericentromeric satellite regions (10) suggests an important role of the isoforms of the gene. The identification of ΔDNMT3B and its multiple splicing variants in this study further complicate the role of DNMT3B family members in regulating DNA methylation in physiological and pathological conditions.

ΔDNMT3B derives from a novel promoter located upstream of exon 5 of DNMT3B1 with a putative translation initiation site at exon 6 (exon 2 of ΔDNMT3B). As a result, the predicted proteins of ΔDNMT3Bs lack 200 amino acids at the N-terminal of DNMT3B1 but maintain the PWPP domain in several variants (ΔDNMT3B1-4). Therefore, these putative proteins may share some common function, such as DNA methyltransferase activity, with DNMT3B but possess other distinct biological features. Because predicted ΔDNMT3B5-7 lack the enzymatic domain of DNA methyltransferase, their biochemical functions may be more distinct than those of their other family members.

In previous studies, DNMT3B has been found to be more highly expressed in cancer cell lines and primary tumors than in normal tissue; however, an association between the expressional level of DNMT3B and promoter methylation of tumor suppressor genes was not established (11-13,19). One possible explanation was that DNMT3B expression may be regulated in the cell cycle and that the increased expression observed in tumors is merely a reflection of increased cell proliferation (13). Studies have shown that both DNMT3B and DNMT1 genes are necessary for maintaining the methylated promoters of tumor suppressor genes (20,21). Another study showed that DNMT3b4, a DNMT3B variant lacking methyltransferase enzymatic motifs, might act as a dominant-negative factor to reduce DNA methylation (10), suggesting that splicing variants of DNMT3B may play distinctive roles in regulating DNA methylation. The identification of ΔDNMT3Bs in this study adds additional complexity to the current knowledge of DNMT3B in biological systems.

The high expression of ΔDNMT3Bs but not DNMT3Bs in both NSCLC cell lines and primary tumors suggests the
importance of the Δ form in lung tumorigenesis. Because the tumors express different patterns of ΔDNMT3B variants, it is possible that the expression of variable ΔDNMT3Bs rather than the overall expression levels plays a role in promoter methylation in lung tumorigenesis. Further studies are needed to address this issue.

A common sequence polymorphism is found in the promoter of ΔDNMT3B, which affects promoter activity. In accordance with the notion that ΔDNMT3Bs is involved in early lung tumorigenesis, individuals carrying the T-type promoter, which has a higher promoter activity, had a >2-fold increased risk of lung cancer in a large case-control epidemiological study (22). Together, these data support the role of ΔDNMT3Bs in regulating promoter methylation during lung tumorigenesis.

Acknowledgements

This study was supported by the Department of Defense, Grant DAMD17-01-1-01689-1.

References

10. Saito Y, Kanai Y, Sakamoto M, Saito H, Ishii H and Hrohoshi S: Overexpression of a splice variant of DNA methyltransferase 3b, DNMT3b