Structure and regulation of the mDot1 gene, a mouse histone H3 methyltransferase

Wenzheng ZHANG*, Yoshihide HAYASHIZAKI† and Bruce C. KONE*[1]

*Division of Renal Diseases and Hypertension, Department of Internal Medicine, The University of Texas Medical School at Houston, 6431 Fannin, MSB 4.148, Houston, TX 77030, U.S.A., and †Laboratory for Genome Exploration Research Group, The Institute of Physical and Chemical Research (RIKEN) Genomic Sciences Center, Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Recently, a new class of histone methyltransferases that plays an indirect role in chromatin silencing by targeting a conserved lysine residue in the nucleosome core was described, namely the Dot1 (disruptor of telomeric silencing) family [Feng, Wang, Ng, Erdjument-Bromage, Tempst, Struhl and Zhang (2002) Curr. Biol. **12**, 1052–1058; van Leeuwen, Gafken and Gottschling (2002) Cell (Cambridge, Mass.) **109**, 745–756; Ng, Feng, Wang, Erdjument-Bromage, Tempst, Zhang and Struhl (2002) Genes Dev. **16**, 1518–1527]. In the present study, we report the isolation, genomic organization and *in vivo* expression of a mouse Dot1 homologue (mDot1). Expressed sequence tag analysis identified five mDot1 mRNAs (mDot1a–mDot1e) derived from alternative splicing. mDot1a and mDot1b encode 1540 and 1114 amino acids respectively, whereas mDot1c–mDot1e are incomplete at the 5′-end. mDot1a is closest to its human counterpart (hDot1L), sharing 84 % amino acid identity. mDot1b is truncated at its N- and C-termini and contains an internal deletion. The five mDot1 isoforms are encoded by 28 exons on chromosome 10qC1, with exons 24 and 28 further divided into two and four sections respectively. Alternative splicing occurs in exons 3, 4, 12, 24, 27 and 28. Northern-blot analysis with probes corresponding to the methyltransferase domain or the mDot1a-coding region detected 7.6 and 9.5 kb transcripts in multiple tissues, but only the 7.6 kb transcript was evident in mIMCD3-collecting duct cells. Transfection of mDot1a–EGFP constructs (where EGFP stands for enhanced green fluorescent protein) into human embryonic kidney (HEK)-293T or mIMCD3 cells increased the methylation of H3-K79 but not H3-K4, -K9 or -K36. Furthermore, DMSO induced mDot1 gene expression and methylation specifically at H3-K79 in mIMCD3 cells in a time- and dose-dependent manner. Collectively, these results add new members to the Dot1 family and show that mDot1 is involved in a DMSO-mediated signal-transduction pathway in collecting duct cells.

Key words: alternative splicing, chromatin, disruptor of telomeric silencing (Dot1), DMSO, histone, nucleosome.

INTRODUCTION

In the nucleus of eukaryotic cells, DNA is coiled around the histone octamer to form the nucleosome, which is further compacted into higher-order chromatin structures. Dynamic changes in the chromatin structure govern accessibility to regulatory transcription factors and, thus, play important roles in many biological processes, such as DNA replication, repair, recombination and transcription. One of the mechanisms by which cells modulate their chromatin structure is the covalent modification of histones by acetylation, methylation, phosphorylation and ubiquitination. Histone methylation occurs on arginine and lysine residues. Histone arginine methylation is typically involved in the activation of genes in which methylases are recruited as co-activators to target promoters. The CARM1/PRMT1 (co-activator-associated arginine methyltransferase, protein arginine methyltransferase) family of HMTs (histone methyltransferases) functions in this manner and principally targets histones H3 or H4 [1–4]. In contrast, gene silencing has been ascribed to the SET (suppressor of variegation, enhancer of zeste, and trithorax) domain family of lysine HMTs. The Suvar39 enzyme was the first such enzyme described, and it was found to methylate K9 of histone H3 and to be recruited by co-repressors [5]. More recently, a third class of HMTs represented by yeast Dot1 (disruptor of telomeric silencing) and its human homologue, hDot1L, was described [6,7]. These HMTs lack a SET domain and affect gene silencing by methylation of a conserved lysine residue (K79) in the globular domain of histone H3. Dot1-mediated methylation of H3-K79 *in vivo* strongly depends on the intact nucleosomal structure and Rad6-dependent ubiquitination of histone H2B at K123 [8].

Dot1 was originally identified as a gene affecting telomeric silencing in *Saccharomyces cerevisiae* and is highly conserved from yeast to human [6,9]. Dot1 methylase is important for transcriptional silencing [9], meiotic checkpoint control during the cell cycle [10] and for localization and association of the Sir silencing proteins in yeast [8]. yDot1 (yeast Dot1) encodes a 582-amino-acid protein containing a lysine-rich domain with at least two consensus nuclear-localization signals in its N-terminal region. Dot1 transcription is cell-cycle-dependent with peak expression at G1 [11]. It remains unknown whether H3-K79 methylation patterns are conserved among eukaryotic species.

In the present study, we report the isolation, genomic organization, alternative splicing isoforms and tissue distribution of a mouse Dot1-like homologue (mDot1) that shares with its human counterpart, hDot1L, a high degree of structural homology and intrinsic H3-K79-specific HMT activity *in vivo*. Furthermore, we demonstrate that the cell-differentiating agent DMSO strongly and specifically induces H3-K79 methylation as a consequence of

Abbreviations used: Dot1, disruptor of telomeric silencing; EST, expressed sequence tag; EGFP, enhanced green fluorescent protein; HMT, histone methyltransferase; SAM, *S*-adenosyl L-methionine; 3′-UTR, 3′-untranslated region; for brevity, the one-letter system for amino acids has been used: e.g. K79 means Lys[79].

[1] To whom correspondence should be addressed (e-mail Bruce.C.Kone@uth.tmc.edu).

The nucleotide sequence data reported have been deposited in the DDBJ, EMBL, GenBank® and GSDB Nucleotide Sequence Databases under accession numbers AY196089, AY196090, AY376663, AY377920 and AY376664.

increased mDot1 mRNA expression in mIMCD3 cells. Thus, our studies add a new orthologue of the Dot1 gene family of HMTs, present new alternative splicing variants, define for the first time the tissue expression patterns in a mammalian species, provide new structural information needed for the design of genetically modified mice to model mDot1 function and regulation and reveal a novel component of the DMSO-mediated signalling pathway in collecting duct cells.

EXPERIMENTAL

Reagents

Antibodies specifically recognizing histone H3 methylated K4, K9, K36 or K79 were purchased from Upstate Biotechnology (Lake Placid, NY, U.S.A.). Anti-α-tubulin was purchased from Santa Cruz Biotechnology (Santa Cruz, CA, U.S.A.). pEGFP and the Living Color A.v. monoclonal antibody recognizing EGFP (enhanced green fluorescent protein) were obtained from Clontech (Palo Alto, CA, U.S.A.). DNA primers were custom-synthesized by Genosys (The Woodlands, TX, U.S.A.). LIPOFECTAMINE™ 2000 reagent was obtained from Invitrogen (Carlsbad, CA, U.S.A.), the bicinchoninic acid protein-estimation kit from Pierce Chemical (Rockford, IL, U.S.A.) and enhanced chemiluminescence reagents from Amersham Biosciences (Piscataway, NJ, U.S.A.). DMSO, forskolin and cycloheximide were purchased from Sigma.

Cell culture and treatment

Human embryonic kidney (HEK)-293T cells and mIMCD3 cells, an immortalized cell line derived from the mouse inner medullary collecting duct [12], were cultured in Dulbecco's modified Eagle's medium, supplemented with 100 μg/ml streptomycin, 100 units/ml penicillin and 10 % foetal bovine serum at 37 °C in a 5 % CO_2 environment. For studying the effects of the test reagents on mDot1 gene expression, mIMCD3 cells, grown to approx. 70 % confluence on 100 mm plates, were treated with different concentrations of forskolin, cycloheximide, DMSO or ethanol for various time periods as indicated in the text and Figure legends. Cells were then briefly washed with PBS and used to prepare acid extracts or total RNA.

EST (expressed sequence tag) and genomic sequence analyses

Initial BLAST searches with the conserved methyltransferase motif of yDot1 against the mouse EST database were performed to identify the EST clones that may be derived from the mDot1 gene. The EST sequences were further analysed using Ensembl Mouse Genome Server (http://www.ensembl.org/Mus_musculus/). Candidate EST clones were obtained from Invitrogen and RIKEN (Yokohama Institute, Japan) for complete sequence analysis. Sequence contig assembly and primer design were performed with MacVector™ 6.5.3. Amino acid sequence alignment was performed with ClustalW 1.8 and BoxShade 3.2.1 from the EMBNet server (http://www.ch.embnet.org/index.html). Searches for candidate transcription start sites were performed with Dragon Promoter Finder software (version 1.4; http://sdmc.lit.org.sg/promoter/promoter1_4/DPF.htm). Consensus binding elements for transcription factors were identified with MatInspector Professional (http://www.genomatix.de/cgi-bin/matinspector/matinspector.pl) [13], using a matrix core identity of 1.0 and an optimized matrix similarity of 1.0 as search criteria.

Plasmids and constructs

The RIKEN EST clone containing the complete coding region of mouse Dot1a cDNA (accession no. BB678539) was obtained and used as a PCR template to amplify either the entire coding region or the N-terminal fragment containing the putative methyltransferase motif (amino acids 2–478) of mDot1ap. These fragments were cloned into pEGFP at *Kpn*I–*Eco*RI and *Xho*I–*Bam*HI sites respectively. The resulting plasmids were designated as pEGFP-mDot1a and pEGFP-mDot1a 2–478 respectively. Mutation of the three highly conserved residues $Gly^{163}Ser^{164}Gly^{165}$, located in the *S*-adenosyl L-methionine (SAM)-binding pocket [14] to RCR, was introduced into either pEGFPmDot1a or pEGFP-mDot1a 2–478 by overlapping PCR to generate pEGFP-mDot1a^{RCR} and pEGFP-mDot1a 2–478^{RCR} respectively. All inserts were sequenced to verify their authenticity.

Transient transfection

mIMCD3 cells or HEK-293T cells grown on 100 or 150 mm plates were transiently transfected with 5 μg of either pEGFP vector or the constructs containing different mDot1a fragments, using the LIPOFECTAMINE™ 2000 reagent as described in our previous work [15]. After 48 h, cells were washed with PBS and collected to isolate acid extracts or whole cellular proteins.

Western-blot analysis

Acid extracts containing free histones were prepared according to a commercially available method (Upstate Biotechnology). Briefly, cells were lysed in lysis buffer (10 mM Hepes, pH 7.9/1.5 mM $MgCl_2$/10 mM KCl/0.5 mM dithiothreitol/1.5 mM PMSF), followed by the addition of 0.2 M HCl for 30 min on ice. After centrifugation at 11 000 ***g*** for 10 min at 4 °C, the supernatant containing the acid-soluble proteins was dialysed against 0.1 M acetic acid twice for 1–2 h and then against water overnight with three water changes. Whole cellular proteins were obtained by combining cytoplasmic and nuclear extracts prepared using the nuclear/cytosolic fraction kit (BioVision, Mountain View, CA, U.S.A.) according to the manufacturer's instructions. Electrophoretically separated proteins were probed with anti-H3 methylated K4, K9, K36 or K79 antibody at the same dilution (1:1000), anti-α-tubulin antibody (1:1000) or Living Color A.v. monoclonal antibody (1:1000).

Northern-blot analysis

Total RNA prepared from mIMCD3 cells or from mouse uterus, ovary, lung, liver, kidney, heart, colon, brain and bladder was analysed on Northern blots as described previously [16]. The blots were hybridized with a ^{32}P-labelled cDNA probe corresponding to the nucleotides encoding amino acids 2–478 of mDot1ap. The blot was reprobed with the cDNA fragment of the full-length mDot1a-coding region to examine whether additional bands resulting from alternative splicing or expression of other Dot1-like genes could be detected. To confirm further the identity of the detected major bands, an 810 bp fragment specific for exon 27, corresponding to nt 4049–4858 of accession number AY196089, was obtained by PCR and used to hybridize a similar blot. A similar Northern-blot analysis was used to detect the DMSO-induced mDot1 expression in mIMCD3 cells. Equality of loading and transfer was verified by inspection of 28 and 18 S rRNA bands.

Table 1 Genomic structure and alternative splicing of mDot1

No.	Exon		Intron	
	Sequence	Length (bp)	Sequence	Length (bp)
			5′-cccccgcgcgcgctcgccgagtgg	
			ttccgcccgaccccccggctcatt	
1	GTGCCTTCTC.....GCCGGTCTAC	319	gtgagtgtac.....tctctttcag	9201
2	GACAAGCACC.....AGACTATCCG	44	gtgagtactg.....tccatttcag	1517
3*	GTGGGTCTGC.....ACACCAAAAG	75	gtgagcagga.....ccttttcag	2367
4*	TTTTGAAAGC.....CCACCAGCTG	64	gtaagtggct.....ttgtccccag	2469
5	TGGAAGGGCA.....CTGGGCAGTG	229	gtgagtgtct.....ctggctgaag	1737
6	GTGTGGGGCA.....GTACGCAGAG	95	gtgagcagtc.....ttcccctcag	507
7	ACCATGGACC.....AGAATACACA	63	gtgagtgcca.....tcttccttag	3787
8	CTGGAACGAG.....CCAACACGAG	56	gtacggccgg.....tgttttgcag	651
9	TGTTATATTT.....ATGAAGGAAG	80	gtaacgcatc.....tgtgttgcag	2500
10	GCGGCAGAAT.....AACTTGAGTG	69	gtaagctctg.....cttctcatag	456
11	ACATTGGCAC.....CCGCACCATA	107	gtgagtcctc.....tcctctttag	668
12*	CTTGAAAACT.....AAAACTCAGG	42	gtaagtttgc.....ttcctttcag	999
13	GAGGAGCAGG.....GGCCCCCGCG	108	gtgagatagg.....tgttctccag	132
14	GATTCTGGTG.....TCACCCCAGG	232	gtaagcctct.....tctcttgcag	217
15	ATGCGTACAG.....AAGCTTTTAG	114	gtgggtcata.....gtgcctccag	578
16	AGTCCTTCAG.....CCAGGAGAAG	92	gtgcgcaggc.....tctcttgcag	85
17	GAGAAGAACA.....GCTGGATGAG	102	gtgacgtgag.....tcatccccag	206
18	TTGGGCGTGA.....CCTGCGGCTG	138	gtgtgtgcag.....gtacccatag	310
19	CTCAGGGCCC.....GGAGCTGCAG	126	gtagggtatt.....catcctctag	158
20	ATCAGCATCG.....CACACCCTCG	485	gtgaggacaa.....ttcattgcag	87
	CATCGTGGAG.....CACACCCTCG	480		
21	GCCAGAGCAC.....TAGTGAGAAG	136	gtaagggctg.....ctattggcag	514
22	GGTGTGAAGG.....TGAGAGGGCG	147	gtgagttgcc.....tgtcttgcag	779
23	AGGAGCACTC.....GCGCCCACAG	130	gtgagctttg.....tcattgatag	551
24A	GTTTCTATGC.....TCCTGTCGCA	169	gtcccgcaac.....gaccccacag	630
24B*	GTCCCGCAAC.....CAACTCCATG	409	gtaagtgcag.....gaccccacag	221
25	GTCAGCAACA.....GCAGTGCCCG	206	gtgagcacag.....ctttgcacag	800
26	AATTGAGAGA.....CTGGAAAATG	65	gtgagttttt.....ctcccatcag	226
27*	GTGGTGGCTT.....ACCGTTGGAG	951	gtaggcaagg.....cgaggggcag	405
28A*	GCCTCGGCCT.....CTGTTTACAG	175		
28B*	GTGTCTTTAC.....GTTCTGTAAG	631		
28C*	GTCCTGAGTG.....TTCCCTGCAG	733		
28D*	GTAATTAGGA.....CACAAAAACG	940	aaaaaaaaaaaaatggcactggcc-3′	

* Exons involved in alternative splicing.

RESULTS

Cloning and structural characterization of the mouse Dot1-like gene

BLAST searches with the conserved Dot1 methyltransferase motif against the mouse EST database produced three significant alignments: AI627144, AI115276 and BI652486. These three EST clones were purchased (Invitrogen), completely sequenced and found to overlap partially to form a 1.3 kb cDNA contig. This contig was incomplete at both 5′- and 3′-ends. To clone the full-length cDNA, the 1.3 kb contig was first employed to search the mouse EST database and yielded no clues of EST clones containing these ends. Then, the contig was used to search human EST databases. A 6579 bp mRNA sequence (XM_046822) predicted from NCBI contig NT_011268 was retrieved. Finally, the complete sequence of XM_046822 was used to search the mouse EST database. In this way, more than 40 partially sequenced EST clones were identified. After restriction enzyme digestion analysis, the IMAGE clones BE573869 and BM943984 and the RIKEN EST clones BB678539, BB487330 and BB729361 were found to contain inserts with sizes ranging from 1.8 to 5.7 kb. All of these clones were completely sequenced. BE573869 is incomplete at the 5′-end, but has an extended 3′-UTR (3′-untranslated region; new accession no. AY196090). BB678539 contains a 5694 bp cDNA fragment that has a single open reading frame and a shortened 3′-UTR (new accession no. AY196089). The open reading frame is predicted to encode 1540 amino acids. The nucleotide sequence (GACATGG) around the first ATG initiation codon conforms to the Kozak initiation sequence, and there is an immediate upstream stop codon. These results suggest that the coding region encodes the entire length of the protein. We designated this new gene as mDot1a. Alignment of both accessions numbers AY196089 and AY196090 yielded a 5793 bp cDNA contig.

Ensembl BLAST search (http://www.ensembl.org/Mus_musculus/blastview) with this contig revealed that the cDNA fragment is derived from 28 exons and a genomic region that spans 39.5 kb on chromosome 10qC1 (Table 1 and Figure 1). Additional EST clone analysis (see below) revealed that exon 24 is further divided into two parts, termed 24A and 24B, whereas exon 28 contains four fragments, termed 28A–D. mDot1a expresses exons 1–27, followed by exon 28D. That is, exons 28A–C are not expressed in this isoform. The open reading frame in mDot1a starts in exon 1 and ends in exon 28D.

Sequencing of BB487330 revealed that it contains a 4599 bp insert produced by alternative splicing. We designated it as mDot1b. Exons 3, 4, 12, 27 and 28A–C were skipped in this variant (new GenBank® accession no. AY376663). As a result, the open reading frame defined in mDot1a (accession no. AY196089) using an ATG in exon 1 was shifted in exon 5, owing to the exclusion of exons 3 and 4, and prematurely terminated after encoding a 49-amino-acid peptide by the stop codon TGA. However, the open reading frame is restarted by another ATG, which shares the first 2 bp TG with the above-mentioned stop codon TGA and

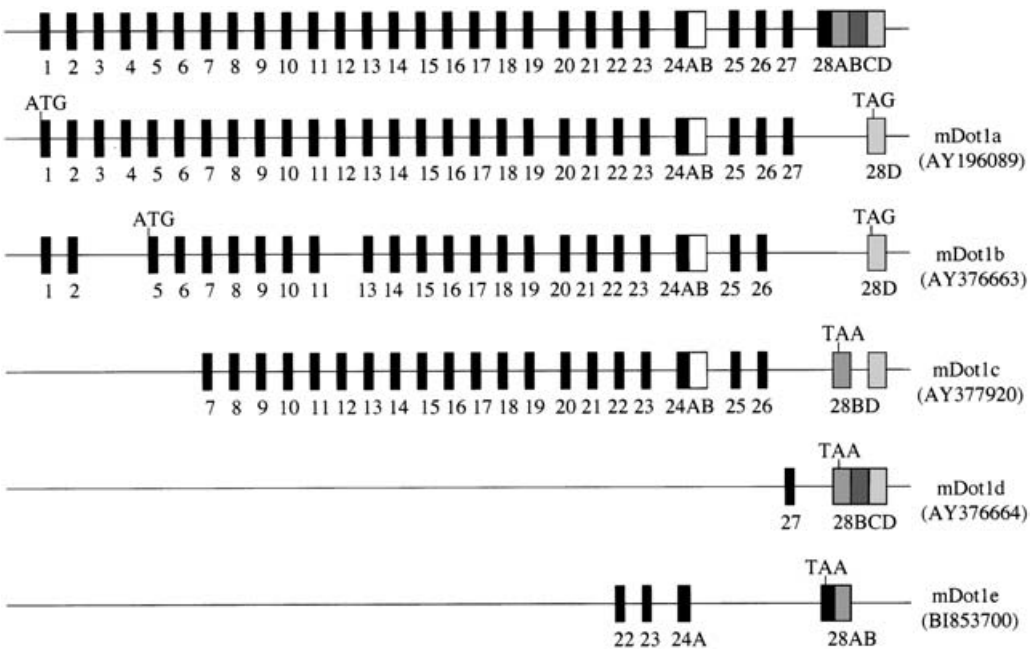

Figure 1 Genomic organization of the mDot1 gene and its alternative splicing variants

Exons are indicated by vertical rectangles and are numbered. Exons with multiple partitions are indicated by differences in shading. The different splicing isoforms are named as mDot1a–mDot1e, with the accession numbers for the cDNA sequences given below. The coding regions of mDot1b–mDot1e were predicted according to mDot1a. The same TAG is used in both mDot1a and mDot1b, whereas both mDot1c and mDot1d share the same TAA. The drawing is not to scale. The exact lengths of all exons and introns are presented in Table 1.

encodes the internal methionine residue at 96 of mDot1a. The sequence CCCATGA around the ATG does not conform to the Kozak sequence. However, the cDNA contains an upstream stop codon. Interestingly, even though exon 27 was not expressed in mDot1b, the open reading frame uses the same TAG stop codon in exon 28D as that in mDot1a. Compared with mDot1ap, mDot1bp does not have amino acids 1–95, 322–334 and 1222–1540. Further analysis is required to determine whether the latter transcript is translated *in vivo*.

The third splicing variant, mDot1c (accession no. AY377920), was documented by sequencing EST clone BB729361. This clone has an insert of 4541 bp, starting in exon 7 and terminating in exon 28D. Similar to mDot1b, mDot1c excludes exon 27, but includes exons 12 and 28B. Using the same open reading frame as that defined in mDot1a, amino acids 215–1222 of mDot1a were encoded by this clone. Addition of exon 28B resulted in an extension of 197 amino acids at the C-terminus. The open reading frame ends in exon 28B by stop codon TAA, which differs from those of mDot1a and mDot1b. This means that the 3′-UTR of this cDNA variant contains an alternative splicing event by skipping exon 28C. Obviously, this clone is incomplete at the 5′-end. We are not sure whether mDot1c skips exons 3 and 4 or not.

mDot1d (accession no. AY376664) was identified by sequencing BM943984. It contains the last 102 bp fragment of exon 27 and a 1.6 kb fragment spanning from exon 28B to 28D. There are four small open reading frames with sizes ranging from 354 to 636 bp. The 636 bp open reading frame is the same open reading frame used in all other mDot1 isoforms and shares the same stop codon TAA in exon 28B as mDot1c. In this way, the last part (amino acids 1504–1539) of mDot1a is joined by a 210-amino-acid extension.

BI853700 was partially sequenced from the 5′-end by the original submitter. We tried to obtain this clone from a commercial vendor (Invitrogen) several times, and each time found that the clone was not the correct one after completely sequencing it. Sequence analysis of the published sequence from this clone revealed that it represents another mDot1 alternative splicing variant, which we designated mDot1e. This variant is very unique. First, although it starts a fragment derived from exons 22 and 23, it includes only the first 169 bp of exon 24, which we termed as 24A. This means that the last 409 bp of exon 24 (24B) is omitted. Instead, a fragment of 175 bp derived from exon 28A is inserted before 28B. We do not know whether the cDNA extends into exon 28C or 28D.

It is noteworthy that the classical GT-AG rule of the intron/exon structure is perfectly observed in all exon/intron boundaries present in all mDot1 isoforms, and the exon sequences identified in the EST clones are perfectly matched with the corresponding publicly available genomic sequences (Table 1). Therefore each of these EST clones contains a single authentic cDNA insert derived from a specific mRNA species, rather than an artifact resulting from the fusion of different cDNA fragments. Further EST clone sequencing will be required to determine the complete cDNA sequences of mDot1c, mDot1d and mDot1e.

Having established that the EST clones are authentic, and that they appear to represent different splicing isoforms of mDot1 mRNAs and that two major transcripts are identified on Northern blots of RNA from all tissues examined (see Figure 3), we intended to clone the entire length of mDot1c, mDot1d and mDot1e and examine the expression of all isoforms by reverse transcriptase–PCR. Unfortunately, we were unable to accomplish this after many unsuccessful attempts with different sources of total RNAs and different primer pairs. Although primers for sequencing or PCR were easily designed from exon 1 or exons 28A–D, no primers within exon 1 can be paired with any primer from exons 28A–D without forming a significant secondary structure between the primers.

Taken together, these results indicate that mDot1 is subject to extensive alternative splicing that can occur in exons 3, 4, 12, 24, 27 and 28. Exon 28 is further divided into four sections with

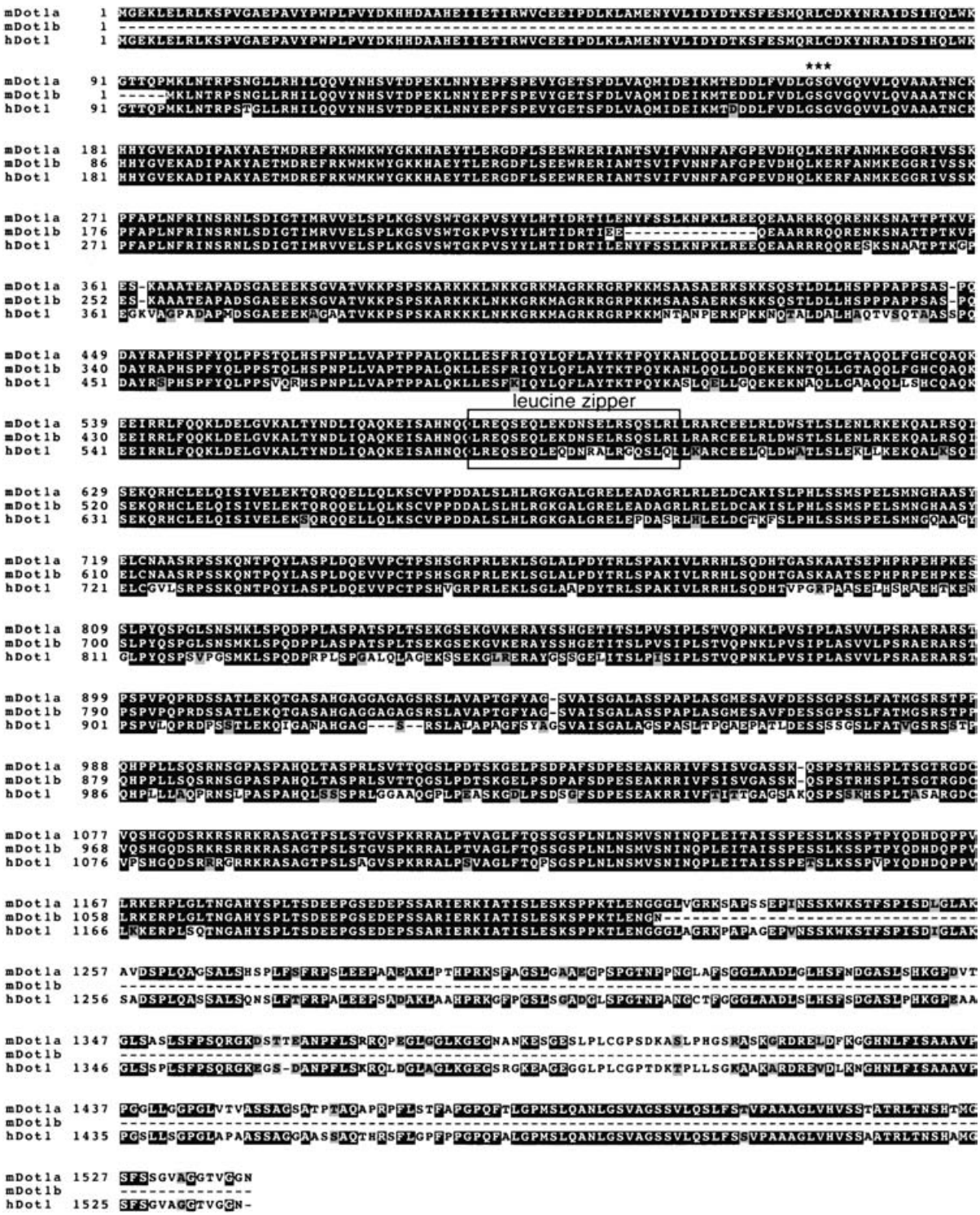

Figure 2 Sequence alignment of mouse and human Dot1 orthologues

ClustalW 1.8 sequence alignment of mDot1ap, mDot1bp and hDot1Lp, with identities shaded in black and similarities shaded in grey. The mutated amino acids located in the SAM-binding site are marked with asterisks. The box indicates the putative leucine zipper motif.

isoform-specific exclusion or inclusion of each section. The five EST clones representing five splicing variants were identified from five different cDNA libraries prepared from different mouse tissues or cells, suggesting that the different isoforms may have overlapping, but also unique, functions in cell growth and differentiation.

Comparison of mDot1a cDNA with hDot1L and yDot1 demonstrates that they encode proteins consisting of 1540, 1537 and 582 amino acids respectively. Both mDot1ap and hDot1Lp share 84 % identity and 88 % similarity (Figure 2), but are similar to yDot1p only around the methyltransferase motif. The N-terminal region (amino acids 1–353) is extremely conserved with

only one amino acid change between mDot1ap and hDot1Lp. In addition, a putative leucine zipper motif is present in mDot1ap and hDot1p (amino acids 574–595 in mDot1p and 576–597 in hDot1Lp), as well as in at least mDot1b and mDot1c, but is absent from yDot1p. This result suggests that the mammalian Dot1 proteins may have obtained additional functions ascribed to leucine zippers, such as DNA-binding activity, through evolution from yeast. Further studies are under way to test this possibility. The mDot1ap and hDot1p are most divergent in two small regions: amino acids 789–840 and 1439–1469 (Figure 2; number corresponding to mDot1ap). These have only 57 % identity and 63 % similarity in the former region. The identity and similarity further decreases to 54.5 and 57.6 % respectively in the latter region. It is also interesting to mention that there are five single-amino-acid deletions, two double-amino-acid insertions and one five-amino-acid insertion in mDot1ap when compared with hDot1p. Whether these differences contribute to differences in function between the two methyltransferases remains to be examined.

We identified the proximal 3 kb of the 5′-flanking region and analysed this region in detail for potential transcription start sites and consensus binding sites for transcription factors. Dragon Promoter Finder software predicted four candidate promoter regions containing transcription start sites. We arbitrarily selected the 3′-most predicted region and arbitrarily assigned the predicted transcription start site in it, which resides 302 bp upstream of the translation-initiation codon for mDot1a, as +1 to provide a reference point for numbering other elements. The true transcription start site remains to be experimentally defined. MatInspector analysis of the mDot1 5′-flanking region using stringent search parameters revealed numerous consensus binding elements for transcription factors known to play a role in cell differentiation and proliferation, as well as DMSO inducibility (Table 2).

mDot1 is widely expressed among tissues

The distribution of Dot1 mRNAs in mammalian species has not been reported previously. We also attempted to determine how many mDot1 isoforms contain the methyltransferase domain. Therefore a ^{32}P-labelled DNA fragment encoding amino acids 2–478 of mDot1ap was used as a probe. As shown in Figure 3, two major bands of approx. 7.6 and 9.5 kb were revealed. Both bands of approx. 7.6 and 9.5 kb bands are longer than any mDot1 cDNA isoforms identified and are even longer than the assembled mDot1a cDNA contig mentioned above. Using the full-length coding region of mDot1a as a probe, we did not detect additional bands. Next, we tried to determine further which isoforms of mDot1 mRNA the two major products represented. A similar blot was hybridized with a probe specific for exon 27; both of the 7.6 and 9.5 kb bands were still detected (results not shown). This result excluded the possibility that these two products were corresponding to mDot1b, mDot1c and mDot1e. On the contrary, they may represent mDot1a and mDot1d. Given that exon 28B and 28C were expressed in mDot1d, but not in mDot1a, which would result in the addition of 1365 bp, it is probable that the approx. 7.6 kb band is mDot1a and the approx. 9.5 kb band is mDot1d. Surprisingly, when RNA from mIMCD3 cells was used, only the approx. 7.6 kb band was observed when the N-terminal probe or the full-length coding region mDot1a probe or exon 27-specific probe was used (Figures 7A–7D).

A BLAST search of the assembled 5793 bp cDNA contig against mouse EST database revealed 70 significant alignments of EST clones. These EST clones are derived from a variety of other tissues or cell types that are not included in our Northern-blot analysis, such as lymph node, diaphragm, retina, testis, mammary gland, placenta and Wolffian duct. EST clones produced from an embryo as early as the two-cell stage or even *in vitro* fertilized eggs to 19.5 days post-conception total foetus were also detected, suggesting that mDot1 is expressed at different developmental stages and may play an important role in cell growth and differentiation.

mDot1 protein has methyltransferase activity towards H3-K79 *in vivo*

As mDot1ap bears closest structural homology to hDot1Lp, we chose mDot1ap for further analysis. To examine whether mDot1ap has methyltransferase activity towards H3-K79 *in vivo*, we used an approach similar to that used by Ng et al. [8] in the study of hDot1Lp. pEGFP, pEGFP-mDot1ap, pEGFP-mDot1ap 2–478, pEGFP-mDot1ap^{RCR} or pEGFP-mDot1a 2–478p^{RCR} vector was transiently transfected into HEK-293T cells. Acid extracts rich in free histones were prepared from the transfected cells and analysed by Coomassie Blue staining of SDS/polyacrylamide gels and by Western-blot analysis using antibodies specific for methylated H3-K4, -K9, -K36 or -K79. The expression level of each fusion protein was determined by isolating whole cellular protein and probing with antibodies specific for EGFP (to detect the fusion proteins) or α-tubulin (to verify equal loading and transfer). As seen in Figure 4(A), overexpression of full-length mDot1ap or mDot1ap 2–478 strongly promoted H3-K79 methylation. In contrast, mDot1ap harbouring a mutation of the putative SAM-binding site encoded by pEGFP-mDot1ap 2–478^{RCR} or pEGFP-mDot1ap^{RCR} or transfection of the vector pEGFP failed to increase H3-K79 methylation (Figure 4A). This result is consistent with previous reports stating that mutating the glycine residues of GSG in the SAM-binding pocket abolishes the methyltransferase activity in both yDot1p and hDot1Lp [7,8]. It is important to note that both of the constructs expressing the intact or mutated N-terminal part, as well as the vector control, were expressed at similar levels, which are much higher than those expressing the full-length protein with or without the mutation (Figure 4F). Similar results were obtained from the experiments using mIMCD3 cells (results not shown). Overexpression of full-length mDot1ap or mDot1ap 2–478, with or without the mutation did not have any obvious effects on the methylation of H3-K4, -K9 or -K36 (Figures 4B–4D). This result is not only consistent with the previous reports [6] stating that overexpression of hDot1p in HEK-293T cells does not alter the methylation level of H3-K4 and -K9, but also extends these results by showing that mDot1a is also not involved directly or indirectly in the methylation of H3-K36. On the basis of these results and the fact that mDot1ap is highly homologous with hDot1Lp, we conclude that mDot1ap, similar to its human orthologue, is a methyltransferase specifically targeting H3-K79 *in vivo*.

DMSO promotes H3-K79 methylation in a dose- and time-dependent manner in mIMCD3 cells

As most studies of Dot1 come from yeast and identification of Dot1 methylation of H3-K79 occurred only recently, the function and importance of H3-K79 methylation remains to be defined fully in mammalian systems. In an effort to identify pathways that signal to mDot1 to influence its expression or activity in mammalian cells, we performed pilot experiments in which mIMCD3 cells were treated with a panel of reagents designed to interrogate specific signalling pathways and examined their effects on H3-K79 methylation by Western-blot analysis. Some of these

Table 2 MatInspector analysis of consensus binding sites for transcription factors in the proximal 5′-flanking region of the mouse mDot1 gene

Transcription-factor-binding sites were analysed by MatInspector Professional release 6.22 using a core similarity of 1.0 and a matrix similarly of 1.0 as search filters. The sequence that corresponds to the optimized transcription factor matrix used by MatInspector is shown. Capitalized nucleotides represent the core sequence in the matrix. To assign a reference point for numbering, the transcription start site (+1) was arbitrarily defined by computer analysis of the 5′-flanking region with Dragon Promoter Finder software and remains to be experimentally defined. This putative transcription start site is 302 bp upstream of the translation-initiation codon ATG of mDot1a.

Name	Description*	Position	Sequence	Strand
ZBPF	Zinc finger transcription factor ZBP-89	−2743 to −2726	tcccccaCCCCcagg	(−)
NFAT	Nuclear factor for activated T-cells	−2728 to −2718	ggagGAAAagg	(+)
CEBP	CCAAT-enhancer-binding protein β	−2654 to −2636	cgtgtttgGCAAggaagtg	(+)
ETSF	c-Ets-1-binding factor	−2648 to −2632	tggcaAGGAagtgcttg	(+)
ETSF	c-Ets-1-binding factor	−2400 to −2384	gatgAGGAagtaggtc	(+)
HEAT	Heat-shock factor 1	−2375 to −2365	AGAAcgttcac	(+)
NMP4	Nuclear matrix protein 4	−2552 to −2243	agAAAAagaaa	(+)
GATA	GATA-binding site (consensus)	−2121 to −2109	ctgaGATAaagcc	(−)
GATA	GATA-binding site (consensus)	−2051 to −2039	agttGATAagtag	(−)
NMP4	Nuclear matrix protein 4	−2007 to −1997	ggAAAAaaatt	(−)
MYOD	Myoblast-determining factor	−1923 to −1900	aatCCACCtgcctct	(+)
CRX	Cone-rod homeobox-containing transcription factor	−1896 to −1880	ctgggATTAaagatgca	(+)
GATA	GATA-binding factor 1	−1856 to −1840	acaaGATAacaaa	(+)
CABL	c-Abl src-type tyrosine kinase	−1726 to −1716	aaAACAaacaa	(+)
SORY	Sox-5	−1689 to −1673	aaaaaCAAagagcaat	(+)
GATA	GATA-binding factor 1	−1642 to −1630	gaagGATAatctc	(+)
MYOD	Myoblast-determining factor	−1614 to −1600	tatcCACCtgcctct	(+)
AP4R	Activator protein 4	−1482 to −1466	agcttCAGCtggggaaa	(+)
RBIT	B-cell regulator of IgH transcription	−1417 to −1400	aataaATTAaaaa	(−)
NKKX	Homeodomain proteins MSX-1 and MSX-2	−1417 to −1405	tttTAATttattt	(+)
SATB	Special AT-rich binding protein	−1406 to −1390	acttacataTAATaaaa	(−)
BRNF	POU factor Brn-2	−1348 to −1332	accatccaTAATgagat	(−)
GFI1	Growth factor independence 1 zinc finger protein	−1315 to −1301	tcaAATCCagcaac	(+)
CHRF	Cell-cycle gene homology region	−1308 to −1296	ggatTTGAacttc	(+)
HOXF	Hox-1.3 homeobox protein	−1228 to −1216	ttttcttaATTAgtgat	(+)
MZF1	MZF1	−1210 to −1204	ggGGGGa	(+)
ZBPF	Zinc finger transcription factor ZBP-89	−1212 to −1198	gggacctCCCCcctt	(−)
IKRS	Ikaros 2	−1112 to −1100	tcttGGGAaaaaa	(+)
PAX6	PAX6	−978 to −960	gcctgagttCCAGtcctga	(+)
EBOR	δEF1	−928 to −919	tgtaAGGTgaataaa	(+)
CREB/ATF6	Activating transcription factor 6	−833 to −813	actaaatGACGtggtaaaaga	(−)
TCFF	TCF11/LCR-F1/Nrf1 homodimers	−822 to −816	GTCAttt	(+)
HOXT	Meis1 homeodomain protein	−614 to −606	gtGACAgct	(+)
SF1F	SF1 steroidogenic factor 1	−604 to −592	acctCAAGgtaac	(+)
STAT5	Signal transducer and activator of transcription 5	−593 to −575	aaaccTTCTgagaataagt	(−)
MZF1	MZF1	−573 to −567	ggGGGGa	(+)
ZBPF	Zinc finger transcription factor ZBP-89	−575 to −561	gttccctCCCCccaa	(−)
LEFF	TCF/LEF-1, involved in Wnt signal transduction	−474 to −458	aagagttCAAAgcaatg	(−)
HOMS	Binding site for S8-type homeodomains	−459 to −451	ctccaATTAa	(−)
HOXT	Meis1 homeodomain protein	−419 to −407	aTAGATttattaat	(+)
HOXF	HOX homeobox factor	−420 to −404	gattGATTtattaattaa	(+)
HOMS	Binding site for S8-type homeodomains	−412 to −404	attaATTAa	(+)
NKXH	DLX-1-, DLX-2- and DLX-5-binding sites	−414 to −402	agttAATTaataa	(−)
RBIT	B-cell regulator of IgH transcription	−414 to −402	agttaATTAataa	(−)
MYOD	Myoblast-determining factor	−365 to −351	aatcCACCtgcctct	(−)
RORA	RAR-related orphan receptor α1	−344 to −328	aattcgaGTCagactg	(+)
FKHD	Forkhead-related activator-3 (FOXC1)	−263 to −247	agcatGTAAataaacaa	(+)
FKHD	HNF-3/Fkh homologue 1 (FOXQ1)	−259 to −243	tgtaaatAAACaataag	(+)
SORY	Sox-5	−241 to −225	ataaaCAATaagcagat	(+)
GATA	GATA-1-binding site	−244 to −232	agcaGATAggcta	(+)
NFAT	Nuclear factor of activated T-cells	−202 to −192	agggGAAAaat	(+)
NMP4	Nuclear matrix protein 4	−199 to −189	ggAAAAattga	(+)
NMP4	Nuclear matrix protein 4	−165 to −158	agAAAAagcca	(+)
IRFF	Interferon regulatory factor-3	−154 to −146	cagaaaaaGAAAatt	(+)
NMP4	Nuclear matrix protein 4	−115 to −105	agAAAAagaaa	(+)
WHZF	Winged helix protein	−65 to −55	gggACGCgaag	(+)
GFI1	Growth factor independence 1 zinc finger protein	−45 to −31	tcAATCccagttgc	(−)
CHRF	Cell-cycle gene homology region	−38 to −26	ggatTTGAacttg	(+)
PXRF	Half-site of PXR/RXR-response element	−35 to −25	ttTGAActtga	(+)

* Abbreviations: DLX-1, distal-less 1; HOX, homeobox; RAR, retinoic acid receptor; HNF, hepatocyte nuclear factor.

chemicals were dissolved in DMSO (e.g. forskolin) or ethanol (e.g. cycloheximide), so that these vehicles were included as controls. Acid extracts were prepared from the treated cells and examined for changes in H3-K79 methylation by Western-blot analysis. Representative Western blots are shown in Figure 5. Without any treatment, mIMCD3 cells showed basal H3-K79

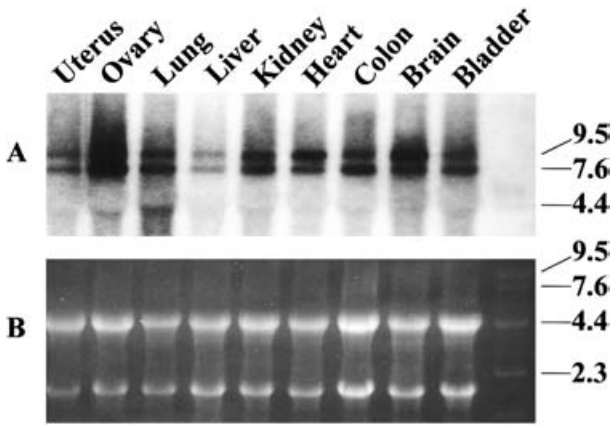

Figure 3 Tissue distribution of mDot1 mRNA in mouse

A representative Northern blot of total RNA harvested from the indicated tissues, which was probed with a ^{32}P-labelled cDNA probe corresponding to the nucleotides encoding amino acids 2–478 of mDot1ap (**A**). Equal loading was demonstrated by staining of 18 and 28 S RNA (**B**). Similar results were obtained when the blot was hybridized with the cDNA fragment of the complete coding region of mDot1a or with a probe specific for exon 27 (results not shown). The samples were run on the same gel as shown in Figures 7(A) and 7(B). $n = 3$.

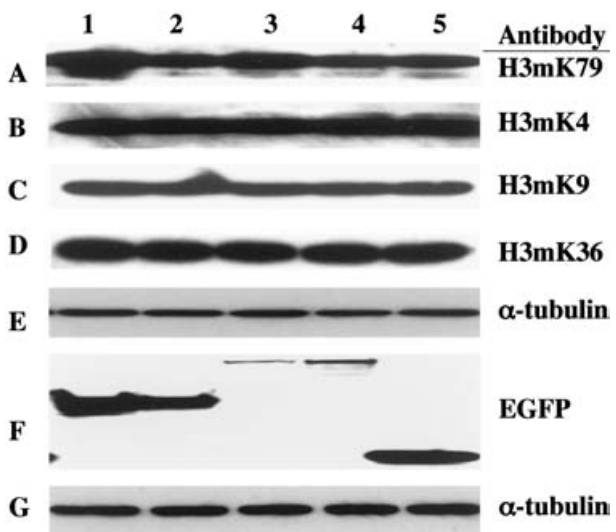

Figure 4 *In vivo* methyltransferase activity of mDot1ap towards H3-K79

HEK-293T cells were transiently transfected with an empty EGFP vector or constructs encoding different versions of mDot1a–EGFP fusion proteins as indicated (see the text for details). Cells were harvested 2 days after transfection, and acid extracts rich in histones and whole cellular proteins were prepared. Identical Western blots loaded with acid extracts rich in histones were probed with an antibody specific for methylated H3-K79 (H3mK79), -K4, -K9 or -K36 (**A**–**D**). Expression of the different versions of EGFP–mDot1a fusion was verified by Western-blot analysis of the whole cellular proteins with anti-EGFP antibody (**F**). All blots were then stripped and reprobed for α-tubulin to verify the equal loading, and similar results were obtained. Reprobed blots of (**A**) and (**F**) are shown in (**E**) and (**G**) respectively. $n = 4$.

methylation. Surprisingly, treatment of 1 % (final concentration in the media) DMSO for 48 h dramatically increased H3-K79 methylation (Figure 5A, cf. lanes 4 and 5), whereas 10 μM forskolin did not have significant additive effects on the H3-K79 methylation (Figure 5A, cf. lanes 2 and 4). Neither 2 % (final concentration in the media) ethanol nor cycloheximide (1 μg/ml) induced significant changes in H3-K79 methylation. Moreover, DMSO-induced H3 methylation appears to be limited to K79, as methylation of the three other lysine residues (K4, K9 and K36)

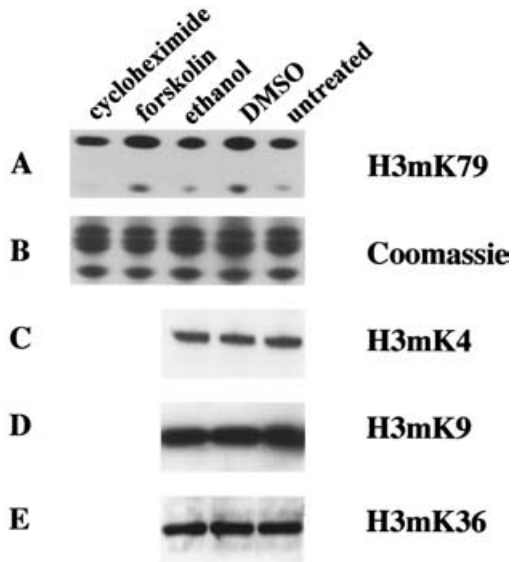

Figure 5 Specific induction of H3-K79 methylation in mIMCD3 cells by DMSO

(**A**, **B**) mIMCD3 cells were incubated for 48 h in the presence of forskolin (10 μM, dissolved in DMSO), cycloheximide (1 μg/ml, dissolved in ethanol) or the respective vehicles (final concentrations in the media: 1 % DMSO and 2 % ethanol). Acid extracts prepared from these cells were subject to Western-blot analysis with the H3mK79-specific antibody (**A**). Equal loading of the extracts was verified by Coomassie Blue staining of a duplicated gel (**B**). $n = 2$. (**C**–**E**) Similar blots as in (**A**, **B**) were probed for H3mK4, H3mK9 and H3mK36. $n = 2$.

did not show any obvious changes by the treatment (Figures 5C–5E). Equal loading of the acid extracts was confirmed by Coomassie Blue staining of the SDS/polyacrylamide gel (Figure 5B). Similar experiments were performed to examine whether DMSO can induce H3-K79 methylation in HEK-293T cells. We found that 1 % DMSO treatment for 48 h did not substantially change the steady-state level of H3-K79 methylation (results not shown). This may be due to the divergence of the 5′-flanking sequences of Dot1 gene between mouse and human or difference between the cell types used.

To analyse the dose–response relationship of H3-K79 methylation by DMSO, mIMCD3 cells were treated with different DMSO doses at final concentrations of 1, 2.5, 5 and 10 % for 48 h. Since treatment of the cells with 5 or 10 % DMSO led to extensive cell detachment from the plate, they were excluded from further analysis. Core histones isolated from cells treated with either 1 or 2.5 % DMSO exhibited a higher level of H3-K79 methylation compared with those from untreated cells, and 2.5 % DMSO was more effective than 1 % DMSO in stimulating the methylation of H3-K79 (Figure 6A). Time-course studies indicated that 1 % DMSO promoted detectable increases in H3-K79 methylation as early as 24 h, with a progressive increase in steady-state H3-K79 methylation up to 72 h (Figure 6C; results not shown).

DMSO induces dose- and time-dependent mDot1 expression in mIMCD3 cells

Since Northern-blot analysis indicated that two abundant mDot1-like mRNA species were present in all tissues examined and that DMSO caused dramatic changes in H3-K79 methylation in mIMCD3 cells, we sought to determine whether these mRNA species were also produced in mIMCD3 cells and whether the DMSO induction of H3-K79 methylation correlated with changes

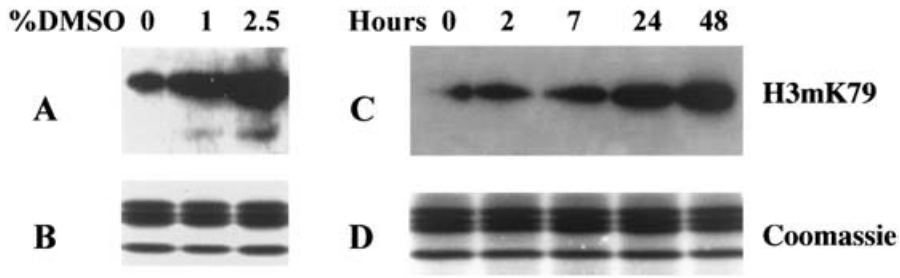

Figure 6 Dose- and time-dependent induction of H3-K79 methylation in mIMCD3 cells by DMSO

(**A**, **B**) Experimental details are the same as in Figures 5(A) and 5(B), except that the cells were treated with 0, 1 and 2.5 % DMSO. (**C**, **D**) Identical with (**A**, **B**), except that the cells were cultured in the presence of 1 % DMSO for 2, 7, 24 and 48 h. $n = 2$.

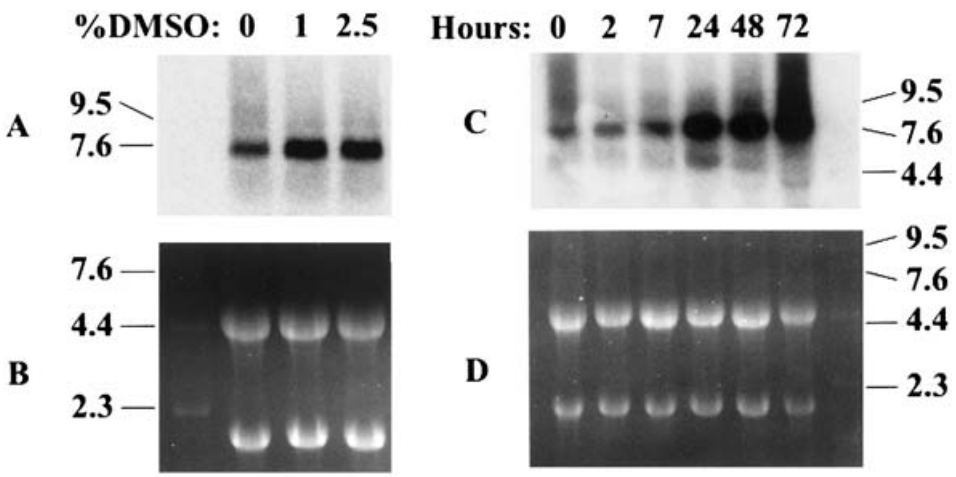

Figure 7 Dose- and time-dependent increase in mDot1 mRNA level in mIMCD3 cells after DMSO treatment

mIMCD3 cells were treated exactly in the same manner as in Figures 6(A)–6(D). Northern-blot analysis of total RNA extracted from these cells was performed as in Figure 3. $n = 3$.

in mDot1 mRNA levels. Northern-blot analysis of total RNA from mIMCD3 cells treated with or without DMSO was performed with the same three mDot1a probes used in the experiments depicted in Figure 3. As illustrated in Figure 7, all three probes detected a major band of approx. 7.6 kb under basal conditions (Figures 7A, lane 1, and 7C, lane 1), whereas the band of approx. 9.5 kb evident in all tissues examined (see Figure 3) was not detected in mIMCD3 cells. These results strongly suggest that the band of approx. 7.6 kb represents the principal mDot1 mRNA in mIMCD3 cells. In addition, DMSO induced a dose- and time-dependent increase in steady-state mDot1 mRNA levels (Figure 7) in a manner that paralleled the delayed induction of H3-K79 methylation. Taken together, we conclude that mDot1 is solely responsible for the basal and inducible methylation of H3-K79 in mIMCD3 cells.

DISCUSSION

The N-terminal tails of core histones are targets of multiple types of covalent modifications. Histone methylation has been shown to be important in regulating chromatin dynamics and gene activity. A number of HMTs have been identified and were found to be involved in gene regulation. Among these is the recently described Dot1 family of HMTs. In the present study, we have cloned and characterized the mouse homologue mDot1, determined its genomic organization, detected multiple exons subject to alternative splicing, established its *in vivo* function as a methyltransferase of H3-K79, characterized for the first time its distribution among mammalian tissues and identified novel induction of its expression by DMSO.

Detailed sequence analysis revealed that mDot1 is expressed from at least 28 exons located in 10qC1. Interestingly, hDot1L mRNA is also expressed from 28 exons from 19p13.3, which is syntenic to mouse 10qC1. Comparison of the genomic structures between mDot1a and hDot1L revealed that the exon structures are well conserved and nearly all are identical in length, whereas the introns are very divergent in terms of sequence and length. The highly conserved exon structures provided the basis for the high identity at the amino acid level and, thus, for the intrinsic H3-K79 methyltransferase activity between the two Dot1s.

EST analysis demonstrated that mDot1 expression is subject to extensive alternative splicing that involves exons 3, 4, 12, 24, 27 and 28 and produces at least five isoforms. The expression of at least two isoforms is further supported by the observation that two major bands are present on the Northern blots with total RNAs isolated from different mouse tissues. By using an exon 27-specific probe, we determined that these two principal transcripts are not mDot1b, mDot1c or mDot1e, since exon 27 is absent from these variants. We are not certain that the approx. 7.6 and 9.5 kb transcripts are mDot1a and mDot1d, as they are much larger than our assembled cDNA contig of mDot1a, and

the entire length of mDot1d remains to be defined. The different isoforms may have differential functions towards cell growth and differentiation, as the smaller transcript is the main, if not the only, mDot1 mRNA species in mIMCD3 cells. It is possible that mIMCD3 cells express all other isoforms in a much lower level that cannot be obviously detected by Northern-blot analysis.

Among the five isoforms of mDot1 mRNA, we determined the complete coding region of mDot1a and mDot1b. Interestingly, compared with mDot1a, mDot1b undergoes three alternative splicing events to skip exons 3, 4, 12 and 27. Multiple regions involving alternative splicing within a single mRNA variant have been reported. For example, the mouse *Ggn* gene gives rise to eight mRNA variants by alternative splicing, and one of these variants (splicing variant 5) is generated by three alternative splicing events [17].

mDot1b lacks both N- and C-terminal fragments as well as amino acids 322–334 that are present in mDot1a. It remains to be established whether mDot1b still possesses H3-K79-specific methyltransferase activity and uses nucleosome as the substrate and to determine the potential functional significance of the N- and C-terminal domains as well as the internal fragment of amino acids 322–334 present in mDot1a. Finally, the spatial expression profile of mDot1b as well as all other mDot1 isoforms during mouse development and their cellular localization regarding their overlapping and unique functions remain to be defined.

Virtually nothing is known about the regulation of Dot1 gene expression in mammalian cells *in vivo*. We found that mDot1 is widely expressed among mammalian tissues and is detectable in databases of expressed genes from various stages of mouse development. This result is consistent with mDot1 having an important and ubiquitous role in cellular proliferation and differentiation. Our finding of time- and dose-dependent induction of mDot1 mRNA expression and H3-K79 methylation by DMSO in mIMCD3 cells also provides new insights into the regulation of this gene, as well as new information about the molecular mechanism of action of DMSO. The delay in the onset of DMSO induction of mDot1 suggests that secondary factors might contribute to the stimulus required for transcriptional activation. The results also suggest an autologous feedback loop in which mDot1-mediated H3-K79 methylation might alter the transcription of mDot1 itself.

The putative 5′-flanking region of mDot1 is dense in consensus sites for transcription factors, and many of these are important in cellular differentiation and/or DMSO induction (Table 2). A GATA-1 site has been shown to confer DMSO inducibility of the mouse protein 4.2 gene [18]. The mDot1 5′-flanking region contains multiple GATA-1 sites distributed throughout the 5′-flanking region (Table 2). In contrast, the mDot1 5′-flanking region that we analysed did not contain an AML1 (acute myeloid leukaemia-1) site, which has been reported to confer DMSO inducibility to the human myeloperoxidase gene [19], an Sp1 (SV40 protein-1) element, which has been identified in the DMSO-responsive enhancers of the β-galactoside α2,6 sialyltransferase gene [20] and the mouse high-affinity neurotensin receptor (Ntr-1) gene [21] or a putative antioxidant response element (5′-CAGCCCCAGGG-ACAGAGCTG-3′), which was shown to be DMSO-responsive to the human apolipoprotein A-1 promoter [22]. It is possible that these or other DMSO-inducible elements are located elsewhere in the gene. Comparison of the proximal 3 kb of the mDot1 and hDot1 5′-flanking regions reveals a significant homology only in two short spans, corresponding to mDot1 -181 to -148 and -138 to -114. Within these regions shared by mDot1 and hDot1, there are no known DMSO-responsive elements. Further studies will be needed to define specific elements of the mDot1 promoter responsible for gene induction.

DMSO is an amphipathic molecule that is commonly used as a solvent in biological studies and as a vehicle for drug therapy. DMSO has multiple cellular effects and serves as a cell-differentiating agent, cryoprotectant and hydroxyl radical scavenger [23]. It has been shown to arrest the cell cycle in several cell lines [24–26]. Clinically, it has been used as a topical analgesic and in the treatment of brain oedema [27], interstitial cystitis [28] and amyloidosis [29]. However, the molecular mechanisms underlying its actions are incompletely understood. As a cell-differentiating agent, DMSO has been shown to induce the differentiation of cell lines of different lineages, such as HL-60 leukaemia cells [30] and P9CL6 cells derived from embryonal carcinoma [31]. Critical events in P9CL6 differentiation into spontaneously beating cardiac myocytes occurred in the first 4 days and required DMSO stimulation. cDNA microarray analysis of global changes in gene expression during P9CL6 differentiation revealed that only 13 known genes and 16 ESTs showed more than 2-fold increase in expression on days 1–3 [31]. mDot1 was not shown in this list. We do not know whether mDot1 was included in the microarrays containing approx. 9000 genes of known sequence in this study. As Dot1 mRNA and, thus, methylation of H3-K79 can be significantly increased by DMSO treatment within 24 h in mIMCD3 cells, it would be interesting to determine whether DMSO induction of mDot1 expression and H3-K79 methylation can occur and is required for the differentiation of P9CL6 cells.

The genomic structural information reported here should prove to be of fundamental importance in the design and interpretation of studies examining the regulation of the mDot1 gene in response to provocative stresses, cell-cycle transitions or developmental stages. This information should prove useful in functional studies of potential *cis* elements that govern differential transcriptional regulation of the mDot1 gene. Since such gene control elements may be dispersed throughout the gene, knowledge of the complete gene sequence, as reported here, should expedite the discovery of mDot1 gene control mechanisms.

B. C. K. was supported by NIH grant nos. RO1 DK-50745 and P50 GM-20529, a 'DREAMS' Center grant from the U.S. Department of Defense and endowment funds from The James T. and Nancy B. Willerson Chair. We are grateful to Dr Xiangyang Xu and Mrs Sandra Higham for their technical support. We also thank Dr Paul Szauter from The Jackson Laboratory (Bar Harbor, Maine, U.S.A.) for help with mouse genomic sequence search and analysis.

REFERENCES

1 Bauer, U. M., Daujat, S., Nielsen, S. J., Nightingale, K. and Kouzarides, T. (2002) Methylation at arginine 17 of histone H3 is linked to gene activation. EMBO Rep. **3**, 39–44

2 Wang, H., Huang, Z. Q., Xia, L., Feng, Q., Erdjument-Bromage, H., Strahl, B. D., Briggs, S. D., Allis, C. D., Wong, J., Tempst, P. et al. (2001) Methylation of histone H4 at arginine 3 facilitating transcriptional activation by nuclear hormone receptor. Science **293**, 853–857

3 Schurter, B. T., Koh, S. S., Chen, D., Bunick, G. J., Harp, J. M., Hanson, B. L., Henschen-Edman, A., Mackay, D. R., Stallcup, M. R. and Aswad, D. W. (2001) Methylation of histone H3 by coactivator-associated arginine methyltransferase 1. Biochemistry **40**, 5747–5756

4 Strahl, B. D., Briggs, S. D., Brame, C. J., Caldwell, J. A., Koh, S. S., Ma, H., Cook, R. G., Shabanowitz, J., Hunt, D. F., Stallcup, M. R. et al. (2001) Methylation of histone H4 at arginine 3 occurs *in vivo* and is mediated by the nuclear receptor coactivator PRMT1. Curr. Biol. **11**, 996–1000

5 Rea, S., Eisenhaber, F., O'Carroll, D., Strahl, B. D., Sun, Z. W., Schmid, M., Opravil, S., Mechtler, K., Ponting, C. P., Allis, C. D. et al. (2000) Regulation of chromatin structure by site-specific histone H3 methyltransferases. Nature (London) **406**, 593–599

6 Feng, Q., Wang, H., Ng, H. H., Erdjument-Bromage, H., Tempst, P., Struhl, K. and 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

7 van Leeuwen, F., Gafken, P. R. and Gottschling, D. E. (2002) Dot1p modulates silencing in yeast by methylation of the nucleosome core. Cell (Cambridge, Mass.) **109**, 745–756

8 Ng, H. H., Feng, Q., Wang, H., Erdjument-Bromage, H., Tempst, P., Zhang, Y. and Struhl, K. (2002) Lysine methylation within the globular domain of histone H3 by Dot1 is important for telomeric silencing and Sir protein association. Genes Dev. **16**, 1518–1527

9 Singer, M. S., Kahana, A., Wolf, A. J., Meisinger, L. L., Peterson, S. E., Goggin, C., Mahowald, M. and Gottschling, D. E. (1998) Identification of high-copy disruptors of telomeric silencing in *Saccharomyces cerevisiae*. Genetics **150**, 613–632

10 San-Segundo, P. A. and Roeder, G. S. (2000) Role for the silencing protein Dot1 in meiotic checkpoint control. Mol. Biol. Cell **11**, 3601–3615

11 Spellman, P. T., Sherlock, G., Zhang, M. Q., Iyer, V. R., Anders, K., Eisen, M. B., Brown, P. O., Botstein, D. and Futcher, B. (1998) Comprehensive identification of cell cycle-regulated genes of the yeast *Saccharomyces cerevisiae* by microarray hybridization. Mol. Biol. Cell **9**, 3273–3297

12 Rauchman, M. I., Nigam, S. K., Delpire, E. and Gullans, S. R. (1993) An osmotically tolerant inner medullary collecting duct cell line from an SV40 transgenic mouse. Am. J. Physiol. **265**, F416–F424

13 Quandt, K., Frech, K., Karas, H., Wingender, E. and Werner, T. (1995) MatInd and MatInspector: new fast and versatile tools for detection of consensus matches in nucleotide sequence data. Nucleic Acids Res. **23**, 4878–4884

14 Min, J., Feng, Q., Li, Z., Zhang, Y. and Xu, R. M. (2003) Structure of the catalytic domain of human DOT1L, a non-SET domain nucleosomal histone methyltransferase. Cell (Cambridge, Mass.) **112**, 711–723

15 Zhang, W., Kuncewicz, T., Higham, S. C. and Kone, B. C. (2001) Structure, promoter analysis, and chromosomal localization of the murine H^+/K^+-ATPase $\alpha 2$ subunit gene. J. Am. Soc. Nephrol. **12**, 2554–2564

16 Kone, B. C. and Higham, S. C. (1998) A novel N-terminal splice variant of the rat H^+-K^+-ATPase $\alpha 2$ subunit. Cloning, functional expression, and renal adaptive response to chronic hypokalemia. J. Biol. Chem. **273**, 2543–2552

17 Lu, B. and Bishop, C. E. (2003) Mouse GGN1 and GGN3, two germ cell-specific proteins from the single gene Ggn, interact with mouse POG and play a role in spermatogenesis. J. Biol. Chem. **278**, 16289–16296

18 Karacay, B. and Chang, L. S. (1999) Induction of erythrocyte protein 4.2 gene expression during differentiation of murine erythroleukemia cells. Genomics **59**, 6–17

19 Austin, G. E., Zhao, W. G., Regmi, A., Lu, J. P. and Braun, J. (1998) Identification of an upstream enhancer containing an AML1 site in the human myeloperoxidase (MPO) gene. Leuk. Res. **22**, 1037–1048

20 Taniguchi, A., Hasegawa, Y., Higai, K. and Matsumoto, K. (2000) Transcriptional regulation of human β-galactoside $\alpha 2$, 6-sialyltransferase (hST6Gal I) gene during differentiation of the HL-60 cell line. Glycobiology **10**, 623–628

21 Tavares, D., Tully, K. and Dobner, P. R. (1999) Sequences required for induction of neurotensin receptor gene expression during neuronal differentiation of N1E-115 neuroblastoma cells. J. Biol. Chem. **274**, 30066–30079

22 Tam, S. P., Zhang, X., Cuthbert, C., Wang, Z. and Ellis, T. (1997) Effects of dimethyl sulfoxide on apolipoprotein A-I in the human hepatoma cell line, HepG2. J. Lipid Res. **38**, 2090–2102

23 Santos, N. C., Figueira-Coelho, J., Martins-Silva, J. and Saldanha, C. (2003) Multidisciplinary utilization of dimethyl sulfoxide: pharmacological, cellular, and molecular aspects. Biochem. Pharmacol. **65**, 1035–1041

24 Darling, D., Tavassoli, M., Linskens, M. H. and Farzaneh, F. (1989) DMSO induced modulation of c-*myc* steady-state RNA levels in a variety of different cell lines. Oncogene **4**, 175–179

25 Teraoka, H., Sawai, M., Takase, K., Yamamoto, K., Nozaki, N., Okazaki, T. and Tsukada, K. (1991) Expression of c-*fos* and c-*myc* in Raji Burkitt's lymphoma cells during the progression of DMSO-induced G1 cells into S phase. Exp. Cell Res. **195**, 274–276

26 Thomas, C. P., Dunn, M. J. and Mattera, R. (1995) Ca^{2+} signalling in K562 human erythroleukaemia cells: effect of dimethyl sulphoxide and role of G-proteins in thrombin- and thromboxane A2-activated pathways. Biochem. J. **312**, 151–158

27 Ikeda, Y. and Long, D. M. (1990) Comparative effects of direct and indirect hydroxyl radical scavengers on traumatic brain oedema. Acta Neurochir. Suppl. (Wien) **51**, 74–76

28 Parkin, J., Shea, C. and Sant, G. R. (1997) Intravesical dimethyl sulfoxide (DMSO) for interstitial cystitis – a practical approach. Urology **49**, 105–107

29 McCammon, K. A., Lentzner, A. N., Moriarty, R. P. and Schellhammer, P. F. (1998) Intravesical dimethyl sulfoxide for primary amyloidosis of the bladder. Urology **52**, 1136–1138

30 Tohkin, M., Morishima, N., Iiri, T., Takahashi, K., Ui, M. and Katada, T. (1991) Interaction of guanine-nucleotide-binding regulatory proteins with chemotactic peptide receptors in differentiated human leukemic HL-60 cells. Eur. J. Biochem. **195**, 527–533

31 Peng, C. F., Wei, Y., Levsky, J. M., McDonald, T. V., Childs, G. and Kitsis, R. N. (2002) Microarray analysis of global changes in gene expression during cardiac myocyte differentiation. Physiol. Genomics **9**, 145–155

Received 5 June 2003/9 September 2003; accepted 23 October 2003
Published as BJ Immediate Publication 23 October 2003, DOI 10.1042/BJ20030839