A mammalian protein with specific demethylase activity for mCpG DNA

Sanjoy K. Bhattacharya*, Shyam Ramchandani*, Nadia Cervoni & Moshe Szyf

Department of Pharmacology and Therapeutics, McGill University, 3655 Drummond Street, Montreal, QC, H3G 1Y6, Canada
** These authors contributed equally to this work.*

DNA-methylation patterns are important for regulating genome functions, and are determined by the enzymatic processes of methylation and demethylation. The demethylating enzyme has now been identified: a mammalian complementary DNA encodes a methyl-CpG-binding domain, bears a demethylase activity that transforms methylated cytosine bases to cytosine, and demethylates a plasmid when the cDNA is translated or transiently transfected into human embryonal kidney cells *in vitro*. The discovery of this DNA demethylase should provide a basis for the molecular and developmental analysis of the role of DNA methylation and demethylation.

Many lines of evidence indicate that DNA methylation is important in differential control of gene expression[1]. The molecular mechanisms through which methylation of cytosines in the dinucleotide CpG repress transcription are now being determined[2–4]. The process that determines which CpGs are methylated is critical for proper development[5] and possibly maintenance of somatic biological functions. Two enzymatic processes are involved in laying down the pattern of methylation during development, namely methylation and demethylation[6–9]. The cDNA encoding the enzyme responsible for methylation, DNA methyltransferase, has been purified and cloned from several organisms[10,11], but the nature of the enzymatic process responsible for demethylation has remained controversial. As the removal of methyl groups from cytosines has been presumed to be energetically unlikely, different alternative biochemical pathways have been proposed to explain active demethylation, such as pathways involving glycosylase[12–14] or a nucleotide-replacement reaction[15,28]. To resolve this issue, we need to know the molecular identity of the enzyme responsible for demethylation. We now show that mammalian cells contain an enzyme that demethylates DNA by catalysing the removal of the methyl group on the 5 position of cytosines residing in the dinucleotide sequence mdCpdG, releasing the methyl group in the form of methanol (our unpublished observations). However, the identification of the gene encoding demethylase and the characterization of its molecular properties will be essential for establishing its presence in vertebrate biology and for further understanding its role in development and in the formation of the DNA-methylation pattern, as well as its possible role in pathological states.

Identification of a DNA-demethylase candidate

As the purification of demethylase indicates that it may be of very low abundance (our unpublished observations), we opted to clone the cDNA encoding the demethylase by making use of the fact that the enzyme specifically demethylates methylated CG dinucleotides (our unpublished observations). We assumed that it should be able to recognize methylated CG dinucleotides. Proteins interacting with methylated DNA share a common domain, the methyl-CpG-binding domain (MBD)[16]. We performed a TBLASTN[17] search of the dbEST database using the MBD of the methyl-CpG-binding protein MeCP2, and identified a new expressed sequence tag (EST) cDNA (from a T-cell-lymphoma *Homo sapiens* cDNA 5′ end) (gb/AA361957/AA361957 EST71295) and a mouse homologue (gb/W97165/W97165 mf90g05.r, from Soares mouse embryo NbME13.5). These cDNAs encode proteins of unknown function that show homology to the MBD of MeCP2 (Fig. 1a). The same cDNA has been cloned and identified as encoding the methylated-DNA-binding protein MBD2b (ref. 18). Alignment of our new EST with the cDNAs encoding MeCP2 and an MeCP1-associated protein revealed no homology beyond the previously characterized MBD, which is consistent with a different function for this methylated-DNA-binding protein. We cloned a 1.36-kilobase (kb) cDNA from HeLa cells. A 'virtual translation' of the cDNA identified an open reading frame (ORF) of 262 amino acids (Fig. 1b). The ORF may extend further in the 5′ direction, as no in-frame stop codon was found upstream of the ATG used. A longer cDNA encoding mouse MBD2 has been cloned[18].

A BLAST search[19,20] of the candidate protein using the Predict protein server[21] against a database of protein-domain families widentified only the MBD. No other functional motifs were identified by Prosite analysis[21,22]. This is consistent with a new function for this protein. A coiled-coil prediction[23] of the sequence identified a coiled-coil domain that is involved in protein–protein interactions[24].

The identified cDNA encodes a messenger RNA that is broadly expressed in human cells; this mRNA was revealed by northern blot analysis of human poly(A)$^+$ mRNA (Fig. 1c) to be one major transcript of ~1.6 kb, which is close to the size of the cloned cDNA, confirming that the cloned cDNA does not represent a highly repetitive RNA, but rather a mRNA encoded by a single or low-copy-number gene.

Demethylation activity of the translated candidate cDNA

A conclusive proof for the existence of a single protein that demethylates DNA is the demonstration that an *in vitro*-translated candidate cDNA can transform a methyl cytosine to cytosine in an isolated system. We subcloned the candidate demethylase cDNA into a pcDNA3.1/His Xpress expression vector in the putative translation frame (producing pcDNA3.1His A) and in a single-base frameshift (producing pcDNA3.1His B), and the cDNA was transcribed *in vitro* and translated in the presence of [^{35}S]methionine. The resulting translation products were resolved by SDS–polyacrylamide gel electrophoresis (PAGE). Autoradiography revealed a protein of relative molecular mass 40,000 (M_r 40K) (Fig. 2a).

The *in vitro*-translated protein (purified on a Ni^{2+}-charged agarose resin) transformed CH_3-cytosine residing in [α-^{32}P]deoxy (d)GTP-labelled plasmid DNA or in (methyl-dC^{32}pdG)$_n$ double-stranded oligomer DNA to cytosine, whereas a frameshift *in vitro*-translated demethylase did not demethylate DNA (Fig. 2b). Thus the demethylase activity is dependent on the demethylase translation product and not on a contaminating activity found in the *in vitro*-translation kit that co-purifies with the putative demethylase. These results show that the candidate cDNA encodes a *bona fide* demethylase.

Transfected demethylase cDNA demethylates DNA

The pcDNA3.1His A and pcDNA3.1His B vectors were transiently transfected into human embryonal kidney cells to test whether the cDNA would direct expression of demethylase activity in human cells. The histidine-tagged proteins were bound to Ni^{2+}-agarose resin and eluted from the resin with 700–1,000 mM imidazole. The expression of the transfected demethylase was verified by western blot analysis (Fig. 3a). The transiently expressed demethylase encoded in pcDNA3.1His A transformed methylated cytosine in two different DNA substrates to cytosine (Fig. 3b, c), in a protein-dependent manner (Fig. 3b, d). The transiently transfected demethylase demethylated a methylated SK plasmid, as indicated by the sensitivity of the plasmid to the restriction enzyme *Hpa*II following its incubation with demethylase but not after incubation in the presence of the protease-K-pretreated demethylase or buffer alone (Fig. 3d). The reaction showed substrate dependence and saturability (Fig. 3e). We loaded transiently expressed demethylase onto a non-denaturing glycerol gradient to determine its native relative molecular mass. Like demethylase purified from human cells (our unpublished observations), our cloned and purified demethylase activity fractionated in the 160K–190K range (data not shown). This is consistent with self-association of our demethylase, possibly mediated by the coiled-coil domain.

A demethylase activity from human cancer cells

To determine whether a demethylase activity like the one shown by the synthetic protein could be isolated from human cells, we subjected nuclear extracts from non-small-cell human lung carcinoma A549 cells to sequential chromatography (see Methods). An active demethylase fraction transformed methyl-cytosine in a (methyl-$dC^{32}pdG)_n$ double-stranded oligomer DNA to cytosine (Fig. 4a). We then subjected demethylase-treated DNA to remethylation with the CpG methyltransferase M.*Sss*I, which can transfer a methyl group exclusively to deoxycytosine[25]. The demethylated product of demethylase must be deoxycytosine, as it is completely remethylated by M.*Sss*I (Fig. 4a). The identity of the demethylated product was confirmed by two-dimensional thin-layer chromatography (TLC), which showed that the product of demethylase migrated with a cold dCMP standard in both dimensions (Fig. 4b).

Demethylase did not release a nucleotide, a phosphorylated base

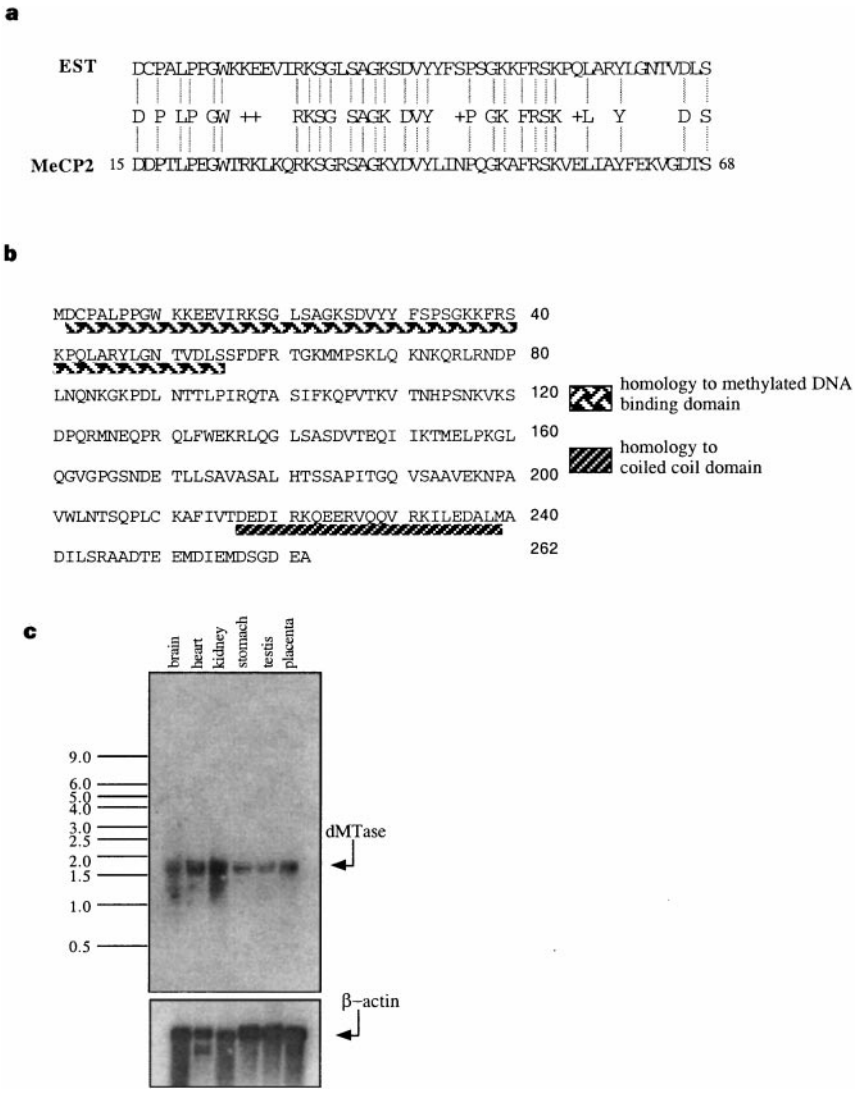

Figure 1 Human demethylase cDNA. **a**, Homology between the methylated-DNA-binding domain (MBD) of MeCP2 (amino acids 15–68) and our putative demethylase (EST). **b**, The sequence of the putative demethylase cDNA was determined and virtually translated by DNASTAR. The 262-amino-acid open reading frame is shown. The region containing homology to the MBD of MeCP2 is highlighted at the amino-terminal of the protein sequence and a predicted coiled-coil domain is highlighted at the carboxy terminus. **c**, A northern blot analysis of 2 μg poly(A)$^+$ mRNA prepared from different human tissues (Origene Technologies) hybridized with demethylase (dMTase) cDNA as a probe.

or phosphate from methylated DNA when incubated with a (^{32}P-methyl-dCpdG)$_n$ substrate that included ^{32}P located 5′ to methyl-deoxycytosine, or when incubated with our standard methylated substrate in which ^{32}P was 3′ to the methyl-5-deoxycytosine (Fig. 4c). Nuclear extracts that contained several glycosylases and nucleases released phosphorylated derivatives in the same assay (Fig. 4c). Demethylase transformed the methyl cytosine in the (^{32}P-methyl-dCpdG)$_n$ substrate to cytosine, as shown when the reacted DNA was digested to 5′ mononucleotides (Fig. 4c, middle panel) and analysed by TLC. As this reaction does not involve release of a ^{32}P derivative (Fig. 4c, left panel), the demethylase must transform methylated cytosines in DNA to cytosines without disrupting the integrity of the DNA substrate by glycosylase or nuclease activity. These results establish that mammalian cells express a *bona fide* demethylase activity, but it remains unknown whether the cloned cDNA encodes the activity purified from A549 cells.

Discussion

An understanding of how methylation patterns are created, maintained and modified and how they are involved in development and pathology has been stifled by the uncertainty surrounding the identity of the enzymes involved in demethylation. We have described here the cloning of a single cDNA that encodes a protein of M_r 40K in both *in vitro* translation and transient transfection

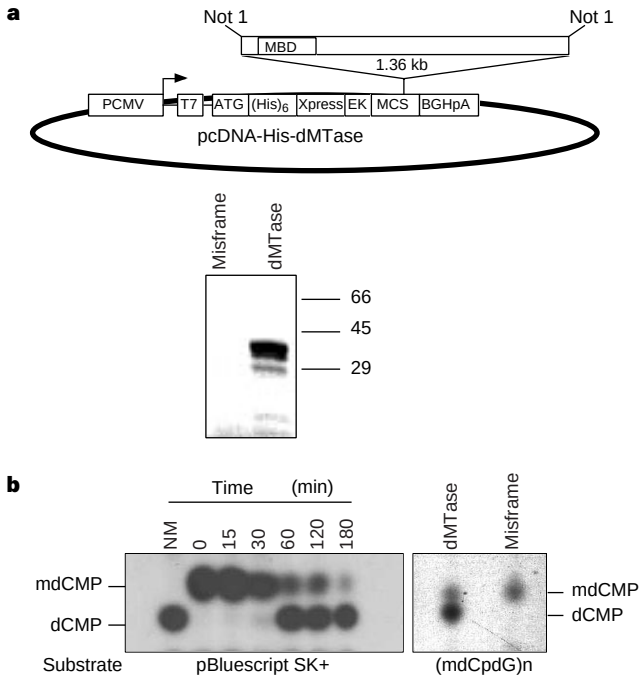

Figure 2 *In vitro* expression and demethylase activity of the human demethylase encoded by the cloned cDNA. **a**, T7-coupled histidine$_6$ tagged mammalian expression construct containing the entire cloned demethylase cDNA in-frame (dMTase) or a single-base-pair frameshift (misframe) were transcribed and translated *in vitro*, revealing an in-frame polypeptide with an apparent M_r value of ~40K (bottom). **b**, Left, demethylation of 2 ng [α-^{32}P]dGTP-labelled methylated SK DNA by incubation with 10 μl Probond-purified (700 mM imidazole fraction) *in vitro*-translated demethylase for varying times. Plasmid pBluescript SK (5 ng) was methylated with M.*Sss*I methylase. *In vitro*-methylated pBluescript SK was used as a template for cDNA synthesis using DNA polymerase and [α-^{32}P]dGTP, methyl-dCTP, dTTP, dATP and hexanucleotide primers. The treated DNA was digested to 3′ mononucleotides with micrococcal nuclease and subjected to TLC. The resulting methyl-dCMP (mdCMP) and dCMP (non-methylated cytosine) mononucleotides are shown. NM, non-methylated substrate control. Right, *in vitro*-translated DNA demethylase, but not misframe demethylase, demethylated cytosines in (methyl-dC^{32}pdG)$_n$ ((mdCpdG)$_n$). The mononucleotide products mdCMP and dCMP are shown. kb, kilobases.

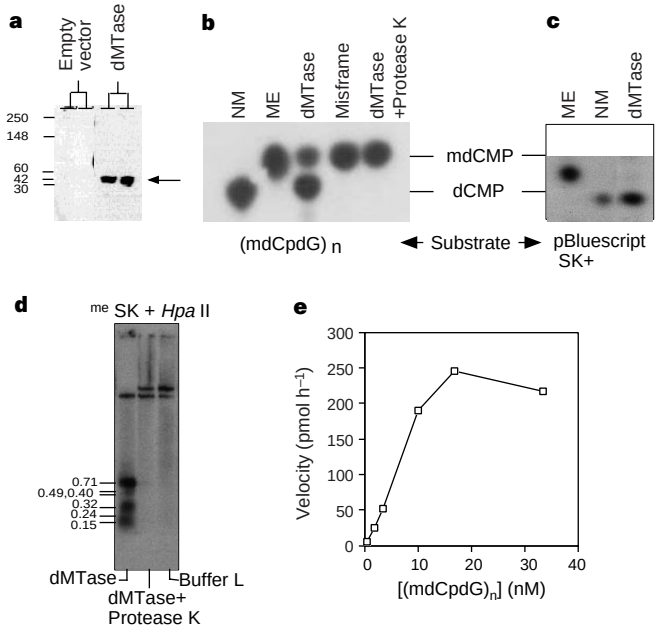

Figure 3 Expression of cloned demethylase cDNA in HEK cells. **a**, A western blot of the purified DNA demethylase (dMTase) and expressed vector control, pcDN3.1HisC (empty vector), from Probond fraction 700 mM imidazole. The western blot was reacted with the anti-Xpress-epitope antibody (Invitrogen). **b**, Demethylation of (methyl-dC^{32}pdG)$_n$ ((mdCpdG)$_n$) DNA treated with 10 μl Probond-purified demethylase (fraction 700 mM; dMTase), Probond-purified misframe demethylase (fraction 700 mM; misframe), or Probond-purified demethylase (fraction 700 mM) pretreated with protease K (20 μg at 50 °C for 2 h; dMTase + protease K). NM, non-methylated substrate control; ME, untreated methylated DNA control. The positions of methyl-dCMP (mdCMP) and dCMP (non-methylated cytosine) are indicated. **c**, Demethylation of [α-^{32}P]dGTP-labelled SK DNA treated with 10 μl Probond-purified demethylase (fraction 700 mM) (dMTase). **d**, Southern blot analysis of *Hpa*II digestion of 5 ng methylated SK DNA (me SK) treated with 5 μl Probond-purified demethylase (fraction 700 mM; dMTase), Probond-purified demethylase (fraction 700 mM) pretreated with protease K (dMTase + Protease K), or incubation buffer alone (Buffer L). Demethylation using these concentrations (5 μl of fractionated dmethylase) is not complete. Under partial demethylation conditions, some molecules are fully demethylated whereas others remain fully methylated. This is consistent with a processive mechanism for the enzyme (N.C. *et al.*, unpublished observations). **e**, Dependence of demethylation activity on substrate concentration. 1 ng methylated [dC^{32}pdG] DNA was supplemented with non-radioactive, methylated [dCpdG] DNA to the indicated concentrations, incubated with 10 μl Probond-purified demethylase (fraction 700 mM) in a 100-μl reaction volume in buffer L for 3 h at 37 °C, and analysed by TLC. Each pmol of DNA substrate bears 500 pmol of 5-methyl-cytosine (the average size of the double-stranded substrate is 500 base pairs). The rate of transformation of methyl-cytosine to cytosine was determined following quantification of the corresponding spots on a phosphorimager.

systems. This single polypeptide demethylates DNA in a defined *in vitro* system, providing conclusive proof for the existence of a *bona fide* demethylation machinery in mammals and presenting us with the molecular tools with which to dissect the biological functions of the demethylase activity. Although our data indicate that a single protein may bear demethylation activities, it is possible that this protein acts in a homomeric complex, as suggested by the sizing of native transiently transfected cloned cDNA.

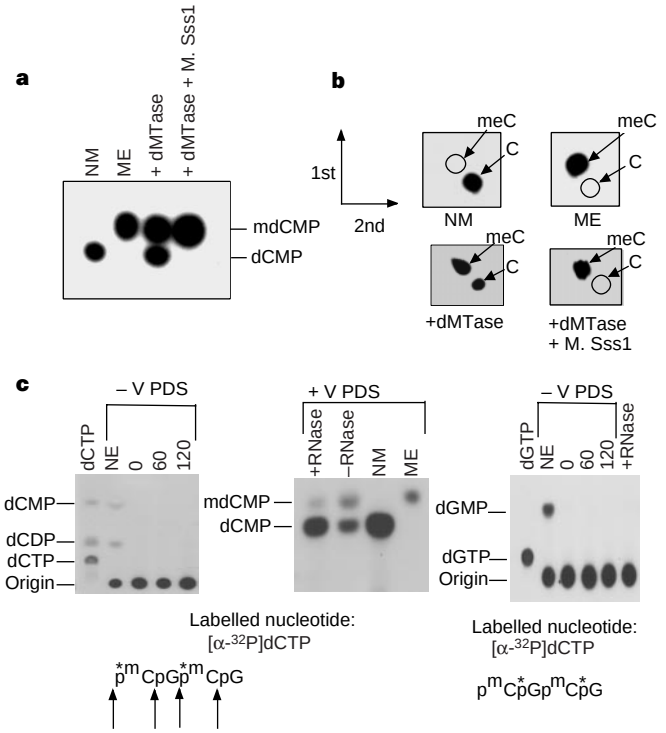

Figure 4 DNA-demethylase activity purified from human cancer cells produces cytosine and exhibits no exonuclease or glycosylase activity. **a**, DEAE–Sephacel demethylase fraction (0.5 units) was incubated with 1 ng (methyl-$C^{32}pG)_n$ substrate in a 50-μl total reaction volume (see Methods) and was divided into two 20-μl aliquots. One aliquot was subjected to methylation with 10 units M.*Sss*I methylase. The remethylated aliquot (dMTase + M.*Sss*1), the non-remethylated aliquot (+dMTase), the (methyl-$C^{32}pG)_n$ substrate that had not been treated with demethylase (ME) and a non-methylated ($C^{32}pG)_n$ substrate (NM) were digested to produce 3′ mononucleotides (methyl-dCMP (mdCMP) and dCMP) and analysed by TLC. **b**, Two-dimensional TLC analysis of the above reaction products in the presence of cold dNMP standards (solvent 1: 66 isobutyric acid:20 H_2O:1 NH_3 solvent 2: 80 $(NH_4)_2SO_4$:2 isopropyl alcohol:2 sodium acetate). The positions of the cold dCMP or mdCMP spots were identified by shining an ultraviolet light on the plate and are indicated in both the one-dimensional (**a**) and the two-dimensional (**b**) analysis. meC, methylated cytosine. **c**, Left, DEAE–Sephacel-purified demethylase (1 unit) was incubated with 1 ng (^{32}p-methyl-dCpdG)$_n$ substrate in 50-μl total reaction volume containing buffer L; 3-μl aliquots were withdrawn at different times, as indicated, and analysed by TLC for the presence of excised ^{32}P residues. As a control, the DNA was incubated for 120 min with A549 nuclear extract (NE). [α-^{32}P]dCTP was run as a control for the absence of non-incorporated dCTP in the labelled DNA (dCTP). Middle, to determine whether demethylase demethylated the DNA in this reaction, the reaction products were digested with a 5′ venom phosphodiesterase (+V PDS) after 2 h of incubation in the presence or absence of 100 μg ml RNase as indicated. (^{32}p-methyl-dCpdG)$_n$ that was incubated in the absence of demethylase (ME) and non-methylated (^{32}pdCpdG)$_n$ (NM) were run as controls. Right, DNA demethylase (dMTase) was incubated with (methyl-dC^{32}pdG)$_n$ substrate, and aliquots were withdrawn at different times and subjected to TLC analysis (shown) and venom phosphodiesterase treatment (data not shown), as described for the middle panel. At the bottom are shown labelled DNA substrates and cleavage sites; asterisks indicate the labelled phosphate groups, and arrows indicate sites of cleavage by venom phosphodiesterase.

We have shown that a general methyl-dCpdG demethylase exists and that this new enzyme can perform the activities that are necessary for demethylation *in vivo*. Our results identify a new enzyme and biochemical reaction that have not yet been described in any organism. The results indicate that the methyl group may be removed in this reaction. Further experiments will be required to unravel the mechanism of the reaction and identify the leaving group. We propose that the reactants of the new reaction are H_2O and methyl-dCpdG-bearing DNA and that the products of the reaction are dCpdG-bearing DNA and methanol. We have shown by TLC (Figs 2b, 3b, c and 4) and restriction enzyme analysis (Fig. 3d) that cytosine is the product of the demethylase reaction. The DNA is intact following the demethylase reaction, as shown by the lack of exonuclease activity (Fig. 4c). Our unpublished preliminary data obtained using votalitization assays, gas chromatography and mass spectrometry indicate that the leaving group may be methanol. As cleavage of a carbon–carbon bond requires high energy, demethylation of 5-methyl-cytosine *per se* has been believed to be thermodynamically unfavourable. If the products of the demethylation reaction are non-methylated cytosine and methanol, the reaction is thermodynamically favourable.

Our data indicate that, like DNA methyltransferase, which exhibits a general recognition of CG dinucleotides, demethylase is an enzyme that can generally recognize methylated CGs. The presence of demethylase mRNA in somatic cells (Fig. 1c) may allow for plasticity in the methylation pattern even in fully developed cells. Altering the level of demethylase in response to extracellular signals[26] or modulating its accessibility to specific sites as a result of the action of *trans*-acting factors can allow for modulation of the covalent modification pattern of the genome in response to physiological signals, and might allow for reversible modulation of gene expression by methylation[27].

Although further experiments will be required to demonstrate the role of demethylase in development and in somatic and cancer tissues, our data provide conclusive evidence that demethylase exists and reveal its primary molecular characteristics. The cloned cDNA will enable the further dissection of the roles of demethylase. □

Methods

Purificaton of demethylase from cells and tissues. Nuclear extracts were prepared from A549 (ATCC) cultures at near confluence as described[26]. A freshly prepared nuclear extract (1 ml, 6 mg) was diluted to a conductivity equivalent to 0.2 M NaCl in buffer L (10 mM Tris-HCl, pH 7.5, 10 mM $MgCl_2$) and applied onto a DEAE–Sephadex A-50 column (Pharmacia) (2.0 × 1 cm) that had been pre-equilibrated with buffer L at a flow rate of 1 ml min^{-1}. Following a 15-ml wash with buffer L, proteins were eluted with a 5-ml linear gradient of NaCl (0.2–5.0 M). 0.5-ml fractions were collected and assayed for demethylase activity. Demethylase eluted between 4.9 and 5.0 M NaCl. Demethylase was purified by 8,000-fold after the DEAE step. Active DEAE–Sephadex column fractions were pooled, adjusted to 0.2 M NaCl by dilution and loaded onto an SP–Sepharose column (Pharmacia) (2.0 × 1 cm). After washing the column, the proteins were eluted with 5 ml of a linear NaCl gradient (0.2–5.0 M). 0.5-ml fractions were collected and assayed for demethylase activity. Active fractions were pooled, adjusted to 0.2 M NaCl by dilution and applied onto a protein Q–Sepharose column (Pharmacia) (2.0 × 1 cm) and proteins were eluted as described. The demethylase activity eluted around 4.8–5.0 M NaCl. The pooled fractions from the protein Q–Sepharose column were loaded onto a 2.0 cm × 2.0 cm DEAE–Sephacel column (Pharmacia) and eluted with 10 ml buffer L. The activity was detected at fraction 4, which is very near the void volume. A batch of 20 purified column fractions was pooled and subjected to cold vacuum evaporation, followed by dialysis in one litre of buffer L (no salt) at 4 °C, with three changes at 6-h intervals. One unit of demethylase activity was defined as the amount of purified enzyme that produced 1 picomol of cytosine using 1 ng methylated ^{32}pCpG substrate in 1 h at 37 °C.

Assay for demethylase activity. We assayed the conversion of methyl-dCMP as described[26]. 1 ng of [α-^{32}P]-labelled, fully methylated poly[methyl-dC^{32}PdG]$_n$ or a

control, non-methylated poly[dC^{32}pdG]$_n$ substrate was prepared as described[26], reacted in a 50-μl reaction volume in buffer L containing 10 μg RNase (all demethylase reactions were done in the presence of RNase unless indicated otherwise) with the relevant demethylase fraction for 3 h at 37 °C, purified by phenol/chloroform extraction, and digested by micrococcal nuclease (Pharmacia) (100 μg at 10 μl) and incubated with calf spleen phosphodiesterase (Boehringer) (2 μg; to produce 3′ mononucleotides) for 15 h at 37 °C. The digestion products were loaded onto a TLC plate (Kodak, 13255 Cellulose), developed in medium containing 132 ml isobutyric acid:40 ml water:4 ml ammonia solution, and autoradiographed.

Coupled *in vitro* transcription and translation. Demethylase cDNA was cloned from a HeLa cell cDNA library in λTriplEx phage (Clontech) according to standard procedures. The positive insert in the pTriplEx plasmid was excised from the phage according to manufacturer's protocols and the identity of the insert was verified by DNA sequencing. The insert was excised by *Not*I restriction and subcloned into pcDNA3.1/His XpressA vector (Invitrogen) to generate histidine-tagged demethylase and pcDNA3.1/His XpressB to generate a single-base frameshift (misframe).

The plasmids encoded by the pcDNA 3.1/His Xpress demethylase constructs were transcribed and translated by coupled transcription–translation using Promega T7 coupled-TNT rabbit reticulocyte lysate kit (according to the manufacturer's protocol). The translation products were bound to a Probond nickel column (Invitrogen) and demethylase was eluted, according to the manufacturer's protocol, with increasing concentrations of imidazole in the range 50–1,000 mM (100 μl per step). The imidazole-eluted fractions were dialysed against buffer L as described. To assay demethylase activity, four *in vitro*-translation reactions were pooled and purified as above.

Transfection and expression of demethylase in vertebrate cells. 10 μg of either pcDNA3.1/His XpressA (demethylase) or pcDNA3.1/His Xpress B (misframe demethylase) were introduced into HEK 293 cells (ten 100-mm-diameter plates per plasmid at early log phase density) using the calcium phosphate precipitation method. Total cell extracts were prepared 48 h after transfection, bound to a Probond nickel resin (Invitrogen) and eluted with 500 μl of increasing concentrations of imidazole (50–1,000 mM) as recommended by the manufacturer. The imidazole-eluted fractions were dialysed against buffer L.

Received 12 October 1998; accepted 11 January 1999.

1. Siegfried, Z. & Cedar, H. DNA methylation: a molecular lock. *Curr. Biol.* **7,** 305–307 (1997).
2. Nan, X., Campoy, F. J. & Bird, A. MeCP2 is a transcriptional repressor with abundant binding sites in genomic chromatin. *Cell* **88,** 471–481 (1997).
3. Nan, X. S. *et al.* Transcriptional repression by the methyl-CpG-binding protein MECP2 involves a histone deacetylase complex. *Nature* **393,** 386–389 (1998).
4. Jones, P. L. *et al.* Methylated DNA and MECP2 recruit histone deacetylase to repress transcription. *Nature Genet.* **19,** 187–191 (1998).
5. Li, E., Bestor, T. H. & Jaenish, R. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* **69,** 915–926 (1992).
6. Razin, A. & Riggs, A. D. DNA methylation and gene function. *Science* **210,** 604–610 (1980).
7. Brandeis, M., Ariel, M. & Cedar, H. Dynamics of DNA methylation during development. *Bioassays* **15,** 709–713 (1993).
8. Kafri, T., Gao, X. & Razin, A. Methanistic aspects of genome-wide demethylation in the preimplantation mouse embryo. *Proc. Natl Acad. Sci. USA* **90,** 10558–10562 (1993).
9. Monk, M., Boubelik, M. & Lehnert, S. Temporal and regional changes in DNA methylation in the embryonic, extraembryonic and germ cell lineages during mouse embryo development. *Development* **99,** 371–382 (1987).
10. Bestor, T., Laudano, A., Mattaliano, R. & Ingram, V. Cloning and sequencing of a cDNA encoding DNA methyltransferase of mouse cells. The carboxyl-terminal domain of the mammalian enzymes is related to bacterial restriction methyltransferases. *J. Mol. Biol.* **203,** 971–983 (1988).
11. Bestor, T. H. & Verdine, G. L. DNA methyltransferases. *Curr. Opin. Cell Biol.* **6,** 380–389 (1994).
12. Razin, A. *et al.* Replacement of 5-methylcytosine by cytosine: a possible mechanism for transient DNA demethylation during differentiation. *Proc. Natl Acad. Sci. USA* **83,** 2827–2831 (1986).
13. Vairapandi, M. & Duker, N. J. Enzymic removal of 5-methylcytosine from DNA by a human glycosylase. *Nucleic Acids Res.* **21,** 5323–5327 (1993).
14. Jost, J. P., Siegmann, M., Sun, L. J. & Leung, R. Mechanisms of demethylation in chicken embryos—purification and properties of a 5-methylcytosine-DNA glycosylase. *J. Biol. Chem.* **270,** 9734–9739 (1995).
15. Weiss, A., Keshet, I., Razin, A. & Cedar, H. DNA demethylation *in vitro*: involvement of RNA. *Cell* **86,** 709–718 (1996).
16. Cross, S. H., Meehan, R. R., Nan, X. & Bird, A. A component of the transcriptional repressor MeCP1 shares a motif with DNA methyltransferase and HRX proteins. *Nature Genet.* **16,** 256–259 (1997).
17. Boguski, M. S., Lowe, T. M. & Tolstoshev, C. M. dbEST-database for expressed sequence tags. *Nature Genet.* **4,** 332–373 (1995).
18. Henrich, B. & Bird, A. Identification and characterization of a family of mammalian methyl-CpG binding proteins. *Mol. Cell. Biol.* **18,** 6538–6547 (1998).
19. Corpet, F., Gouzy, J. & Kahn, D. The ProDom database of protein domain families. *Nucleic Acids Res.* **26,** 323–326 (1998).
20. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215,** 403–410 (1990).
21. Bairoch, A., Bucher, P. & Hofman, K. The PROSITE database, its status in 1997. *Nucleic Acids Res.* **25,** 217–221 (1997).
22. Rost, B. PHD: predicting one-dimensional protein structure by profile based neural networks. *Methods Enzymol.* **266,** 525–539 (1996).
23. Lupas, A., Van Dyke, M. & Stock, J. Predicting coiled coils from protein sequences. *Science* **252,** 1162–1164 (1991).
24. O'Shea, E. K., Rutkowski, R. & Kim, P. S. Evidence that the leucine zipper is a coiled coil. *Science* **243,** 538–542 (1989).
25. Nur, I. *et al.* Prokayotic and eukaryotic traits of DNA methylation in spiroplasmas (mycoplasmas). *J. Bacteriol.* **164,** 19–24 (1985).
26. Szyf, M., Theberge, J. & Bozovic, V. Ras induces a general DNA demethylation activity in mouse embryonal P19 cells. *J. Biol. Chem.* **270,** 12690–12696 (1995).
27. Szyf, M. DNA methylation properties: consequences for pharmacoloty. *Trends Pharmacol. Sci.* **7,** 233–238 (1994).
28. Swisher, J. A., Rand, E., Cedar, H. & Pyle, A. M. Analysis of putative RNase sensitivity and protease insensitivity of demethylation activity in extracts from rat myoblasts. *Nucleic Acids Res.* **26,** 5573–5580 (1998).

Acknowledgements. We thank B. ten Oever and H.-Y. Shiu for technical assistance and O. A. Mamer for help with mass spectrometry. This research was supported by the National Cancer Institute of Canada.

Correspondence and requests for materials should be addressed to M.S. (e-mail: mszyf@pharma.mcgill.ca). The DNA demethylase cDNA sequence has been deposited with GenBank under accession number AF 072242.