

# Transcriptional Repression by *Drosophila* Methyl-CpG-Binding Proteins

KARIM RODER,<sup>1</sup> MING-SHIU HUNG,<sup>1</sup> TAI-LIN LEE,<sup>1</sup> TZU-YANG LIN,<sup>2</sup> HENGYI XIAO,<sup>3</sup>  
KEN-ICHI ISOBE,<sup>3</sup> JYH-LYH JUANG,<sup>2</sup> AND C.-K. JAMES SHEN<sup>1\*</sup>

*Institute of Molecular Biology, Academia Sinica, Nankang,<sup>1</sup> and Division of Genomic Medicine, National Health Research Institute,<sup>2</sup> Taipei, Taiwan, Republic of China, and Department of Basic Gerontology, National Institute for Longevity Sciences, Ogu, Aichi 474-8522, Japan<sup>3</sup>*

Received 6 January 2000/Returned for modification 22 February 2000/Accepted 6 July 2000

**C methylation at genomic CpG dinucleotides has been implicated in the regulation of a number of genetic activities during vertebrate cell differentiation and embryo development. The methylated CpG could induce chromatin condensation through the recruitment of histone deacetylase (HDAC)-containing complexes by methyl-CpG-binding proteins. These proteins consist of the methylated-DNA binding domain (MBD). Unexpectedly, however, several studies have identified MBD-containing proteins encoded by genes of *Drosophila melanogaster*, an invertebrate species supposed to be void of detectable m<sup>5</sup>CpG. We now report the genomic structure of a *Drosophila* gene, *dMBD2/3*, that codes for two MBD-containing, alternatively spliced, and developmentally regulated isoforms of proteins, dMBD2/3 and dMBD2/3Δ. Interestingly, in vitro binding experiments showed that as was the case for vertebrate MBD proteins, dMBD2/3Δ could preferentially recognize m<sup>5</sup>CpG-containing DNA through its MBD. Furthermore, dMBD2/3Δ as well as one of its orthologs in mouse, MBD2b, could function in human cells as a transcriptional corepressor or repressor. The activities of HDACs appeared to be dispensable for transcriptional repression by dMBD2/3Δ. Finally, dMBD2/3Δ also could repress transcription effectively in transfected *Drosophila* cells. The surprisingly similar structures and characteristics of the MBD proteins as well as DNA cytosine (C-5) methyltransferase-related proteins in *Drosophila* and vertebrates suggest interesting scenarios for their roles in eukaryotic cellular functions.**

In vertebrates, including the mammals, the chromosomal DNAs are modified by C methylation at a limited number of CpG dinucleotides, resulting in m<sup>5</sup>CpG, methylated at position 5 of the C residues. This methylation at CpG has been implicated in the regulation of a number of genetic activities during mammalian cell differentiation and embryo development. These activities include tissue-specific gene transcription, X chromosome inactivation, genomic imprinting, cellular defense against viral agents, and tumorigenesis (references 1, 2, 11, 13, 15, and 34 and references therein). The level of m<sup>5</sup>CpG in the vertebrate cells is presumably balanced by the combined actions of DNA cytosine (C-5) methyltransferases (CpG MTases) (2, 26) and DNA demethylases (3).

It is generally accepted that regulation of the different cellular activities by DNA methylation is modulated mainly through the m<sup>5</sup>CpG residues. Indeed, m<sup>5</sup>CpG could be recognized by a specific class of proteins that consist of the methylated-DNA binding domain (MBD) (21). These “MBD proteins,” in particular MeCP2 and MBD2/3, could preferentially bind to m<sup>5</sup>CpG-containing DNA, recruit histone deacetylase (HDAC)-containing complexes, and thus cause chromatin condensation in the vicinity of m<sup>5</sup>CpG (12, 22, 23, 33). This is very likely one major, but not the only, scheme involved in the functioning of vertebrate DNA methylation.

Unlike the vertebrates, several invertebrate species, including *Drosophila melanogaster* (27, 31), do not have apparent DNA methylation in their genomes. Nor has CpG MTase been reported to occur in these invertebrate animals. It was thus surprising to find mammalian CpG MTase-related proteins as well as MBD proteins expressed in *Drosophila*. A *Drosophila*

protein, DmMTR1, is characteristically similar to the vertebrate maintenance CpG MTase, dnmt1, in several aspects (9). In particular, DmMTR1 molecules are located outside the nuclei during interphase of the syncytial *Drosophila* embryos, as is dnmt1 in the mouse blastocyst (19). However, DmMTR1 molecules appear to be rapidly transported into and then out of the nuclei again, as the embryos undergo the mitotic waves (9). Immunofluorescence data further indicate that DmMTR1 molecules “paint” the whole set of condensed *Drosophila* chromosomes throughout the mitotic phase, suggesting that it may play an essential role in the cell cycle-regulated condensation of the *Drosophila* chromosomes. In addition to DmMTR1, another *Drosophila* polypeptide, DmMT2, exhibiting high sequence homology to mammalian dnmt2, a relatively short homolog of dnmt1 but without detectable methylation activities (25), has been identified through a database search (9, 30). Expression of DmMT2 in *Drosophila* is developmentally regulated (9).

Interestingly, several recent reports have also noted the identification of an MBD protein encoded by the *Drosophila* genome (9, 30, 33). Like the mammalian MBD (12, 22, 23, 33), the *Drosophila* MBD protein, which was termed the “*Drosophila* MBD-like sequence” (33) or dMBD2/3 (30), interacts with HDAC in vitro and in vivo (30). However, it did not appear to recognize m<sup>5</sup>CpG (30). Here we report the complete genomic structure of the *Drosophila* *dMBD2/3* gene, which encodes two alternatively spliced and developmentally regulated MBD proteins. We also provide the first evidence that the *Drosophila* MBD protein(s) could recognize m<sup>5</sup>CpG-containing DNA, albeit with a lower affinity than the mammalian MBD proteins. Remarkably, the protein(s) could also function as a corepressor via an HDAC-independent pathway.

## MATERIALS AND METHODS

**Recombinant clones.** The three P1 clones containing the *dMBD2/3* gene were identified by the hybridization of filter-containing *Drosophila* P1 genomic clones

\* Corresponding author. Mailing address: Institute of Molecular Biology, Academia Sinica, Nankang, Taipei, Taiwan, Republic of China. Phone: 011-886-2-27821436. Fax: 011-886-2-2788-4177. E-mail: ckshen@ccvax.sinica.edu.tw or cjshen@ucdavis.edu.

(Genome Systems Inc.) with an expression sequence tag (EST) clone, LD03777. The nucleotide sequences of the P1 clones, LD03777, and another EST clone, LD46808, were determined by automated sequencing of both strands.

All recombinant procedures were done by the methods of Sambrook et al. (29). The cDNAs of dMBD2/3 $\Delta$  and mouse MBD2b were cloned in frame into pGEX-4T-2 and pGEX-4T-1 (Amersham Pharmacia Biotech), respectively. To generate the plasmid pGEX-dMBD2/3 $\Delta$  (36-226), encoding the glutathione transferase (GST) fusion protein of dMBD2/3 $\Delta$  without the MBD, the cDNA of dMBD2/3 $\Delta$  was cut with *Hae*II and blunt ended, and the 3' fragment was cloned into pGEX-4T-1.

The maps of the recombinant plasmids used in transfection are depicted in Fig. 4A. pSG424-dMBD2/3 $\Delta$  contains the sequence for amino acids (aa) 1 to 226 of dMBD2/3 $\Delta$  cloned in frame in the expression vector pSG424 (28). pSG424-MBD2b contains the sequence for aa 1 to 226 of the mouse MBD2b protein (8). pSG424-p45 was a gift from N. Gavva at the University of California at Davis. pCI-Gal4 harbors the sequence for the Gal4 DNA binding domain downstream of the cytomegalovirus (CMV) promoter in pCI (Promega Biotech). pCI-dMBD2/3 $\Delta$  contains the sequence for aa 1 to 226 of dMBD2/3 $\Delta$  cloned in frame with Gal4 of pCI-Gal4. Similarly, pCI-Gal4-MBD2b contains the sequence for aa 1 to 262 of the mouse MBD2b protein. To obtain pCI-Gal4-dMBD2/3 $\Delta$  (36-226) without the MBD, the respective cDNA was released from pGEX-dMBD2/3 $\Delta$  (36-226) by *Eco*RI/*Xho*I digestion and cloned into *Eco*RI/*Sal*I-cut pCI-Gal4. pG5-TK-CAT was a gift from M. Ptashne. It contains five Gal4-binding sites in front of the thymidine kinase (TK) basic promoter. pG0-TK-CAT is the derivative without Gal4-binding sites. pG0- $\alpha$ -Luc contains the proximal human  $\alpha$ 1 globin promoter (-87 to +41) upstream of the luciferase gene of pGL3-basic (Promega Biotech). To create pG5- $\alpha$ -Luc, five Gal4-binding sites were cloned upstream of the  $\alpha$ 1 globin promoter in pG0- $\alpha$ -Luc. p21 (35) contains the proximal promoter region (-60 to +13) of the murine p21<sup>waf1</sup> gene upstream of the luciferase gene in pGL3-basic. p21-GCmut is a derivative containing a mutated GC box. pPacSp1 was a generous gift from Guntram Suske (Institute of Molecular Biology and Tumor Research, Philipps University, Marburg, Germany). pAc5.1/V5-His/lacZ was obtained from Invitrogen.

**Northern blot analysis.** Total *Drosophila* RNAs were isolated from different stages using the TRIZOL reagent (Life Technologies). Poly(A) RNAs were then isolated with a kit from Qiagen. One microgram of the RNAs was loaded in each lane, blotted onto nitrocellulose filter paper, and hybridized with <sup>32</sup>P-labeled LD03777 probe.  $\alpha$ 1 tubulin RNA patterns were used as the internal control for the Northern blot analysis.

**EMSA.** The sequence of the  $\beta$ -TATA oligonucleotide used in the electrophoretic mobility shift assay (EMSA) is 5'-GATCTCACTGTGTCGACCTTGGGCATAAAAGGCAGAGCAGCTGGAGCTGCTGCTTAA-3' (complement, 3'-CTAGAGTGACACAGCTGGAACCCGTATTTCCGCTCTCGTACCTC GACGACGAATT-5'). To obtain methylated DNA, the oligonucleotide or pBluescript II KS (pBSKS; Stratagene) was fully methylated with *Sss*I methylase according to the instructions of the manufacturer (New England Biolabs). The oligonucleotides were labeled with <sup>32</sup>P at their 5' ends and gel purified prior to EMSA. The GST fusion proteins were all generated from recombinant pGEX-4T plasmids (Amersham Pharmacia Biotech) transformed into *Escherichia coli* BL21 cells. The proteins were purified with glutathione-Sephadex beads (Amersham Pharmacia Biotech). The expression plasmid pGST-ankyrin was a gift from Xin Chen. For EMSA, different amounts of the fusion proteins were preincubated with 100 ng of pBSKS and 8  $\mu$ g of bovine serum albumin at room temperature for 5 min in 20 mM HEPES (pH 7.9)-3 mM MgCl<sub>2</sub>-1 mM EDTA-10 mM  $\beta$ -mercaptoethanol-0.1% Triton X-100-10% glycerol in a total volume of 20  $\mu$ l. The labeled oligonucleotides (40,000 cpm, approximately 1 ng) were then added, and the samples were further incubated for 20 min at room temperature. To separate the protein-DNA complexes, the reaction mixtures were loaded onto a running nondenaturing 4% polyacrylamide gel, which had been prerun in 0.5 $\times$  Tris-borate-EDTA buffer for 30 min at 4°C and 200 V. Electrophoresis was further carried out at 4°C and 200 V for 2 1/2 h. Gels were dried and exposed on X-ray film at -80°C. For competition experiments, a 100-fold molar excess of the competitor DNAs was included in the binding reactions.

**Transient DNA transfection.** The human kidney epithelial cell line 293T was grown in Dulbecco modified Eagle medium-10% fetal calf serum and transfected at 20% confluence in 2-cm "Cell Wells." The DNA samples for transfection were introduced as a calcium phosphate coprecipitate consisting of 1  $\mu$ g of a reporter plasmid (pG0-TK-CAT, pG5-TK-CAT, pG0- $\alpha$ -Luc, pG5- $\alpha$ -Luc, p21, or p21-GCmut), 0.3  $\mu$ g of pCMV- $\beta$ -gal (18), 10  $\mu$ g of salmon sperm DNA as the carrier, and different amounts of an expression plasmid [pSG424-dMBD2/3 $\Delta$ , pSG424-MBD2b, pCI-Gal4, pCI-Gal4-dMBD2/3 $\Delta$ , pCI-Gal4-dMBD2/3 $\Delta$  (36-226), pCI-Gal4-MBD2b, or pSG424-p45]. pSG424 and/or pCI-Gal4 was used as the control plasmid(s), and either was also added as required to keep the total amount of DNA used in each transfection the same. Sixteen hours after transfection, the medium was changed and cells were incubated for another 48 h. The chloramphenicol acetyltransferase (CAT), luciferase, and  $\beta$ -galactosidase assays were performed as described previously.

For transfection of the *Drosophila* SL2 cell line, the cells were maintained at 23.5°C in a modified M3 medium (Invitrogen) supplemented with 10% fetal bovine serum and penicillin and streptomycin (both from Gibco BRL). At 18 h before the transfection, 2  $\times$  10<sup>6</sup> cells were plated out onto individual 10-cm petri

dishes. On the following day, the cells were harvested and resuspended in medium without the serum and antibiotics. Then 7.5  $\times$  10<sup>5</sup> cells/well were seeded onto a 24-well plate, and DNA transfection was carried out using Lipofectin. Each liposome-DNA mixture contained 1  $\mu$ g of the reporter plasmid pG5- $\alpha$ -Luc or pG0- $\alpha$ -Luc, different amounts of an MBD expression plasmid or its corresponding vector, 0.2  $\mu$ g of pPac-Sp1, and 0.3  $\mu$ g of pAc5.1/V5-His/lacZ. The transfected cells were kept in antibiotic-free and serum-free medium at 23.5°C for 6 h before replacement with normal serum. The luciferase and  $\beta$ -galactosidase activities were assayed after another 72 h of incubation.

For trichostatin A (TSA) treatment, 1  $\mu$ g of the reporter plasmid(s) and 50 ng of the expression plasmid(s) were used in each cotransfection. Twenty hours after transfection, the transfected 293T cells were treated with TSA at various concentrations (see Fig. 5) for 24 h. Luciferase or CAT activities were then assayed and normalized to the extract protein concentrations. In keeping with previous studies (16), we observed a general increase of the promoter activities after the TSA treatment, e.g., six- to eightfold for the CMV promoter and five- to ninefold for the Rous sarcoma virus promoter (data not shown).

## RESULTS AND DISCUSSION

**Genomic organization of dMBD2/3, a *Drosophila* gene encoding two MBD-containing polypeptides.** Multiple EST clones were identified by Blast search using the mammalian MBD2b sequence (8) as the bait. One of the clones (LD03777) originating from a *Drosophila* embryonic cDNA library was first sequenced. Its insert is 865 bp long, including the poly(A) tail. LD03777 bears an open reading frame corresponding to 226 aa, which has been referred to previously as the "*Drosophila* MBD-like sequence" (33) or dMBD2/3 $\Delta$  (30). The protein contains an MBD-like sequence in the N-terminal domain and a coiled-coil motif near the C terminus. It is apparently the ortholog of the vertebrate MBD2/MBD3 family (8, 30, 33).

dMBD2/3 $\Delta$  is the product of alternative splicing. First, blot hybridization of a P1 *Drosophila* genomic filter with LD03777 identified three P1 clones, 1540, 1769, and 1809 (Fig. 1A). The in situ data located clones 1540 and 1769 to 85D20-85D24 and 85D18-85D27, respectively, but no in situ information was available for clone 1809. Comparison of the nucleotide sequences of LD03777 and the P1 clones revealed the existence of an 891-bp intron (bp 236 to 1126 [Fig. 1B]) downstream of the MBD of dMBD2/3 $\Delta$ . Second, we found another EST clone, LD46808, 1,198 bp long, which contains an extra exon of 339 bp derived from the 891-bp intron (bp 479 to 817 [Fig. 1A and B]). These data indicate that the primary transcript of the dMBD2/3 gene shown in Fig. 1B is alternatively spliced to generate two mRNAs coding for dMBD2/3 $\Delta$  and a longer protein, dMBD2/3 (Fig. 1A).

**Developmental regulation of the alternative splicing of dMBD2/3.** The two dMBD2/3 mRNAs are differentially expressed during *Drosophila* development, as shown by Northern blot analysis using LD03777 as the hybridization probe (Fig. 1C). In overnight embryos and in adult flies, two mRNAs of lengths similar to those of dMBD2/3 $\Delta$  (890 nucleotides [nt]) and dMBD2/3 (1,270 nt), respectively, were present (Fig. 1C, lanes 2 and 6). On the other hand, only the long transcript was expressed in embryos younger than 2 h (Fig. 1C, lane 1), and the short mRNA was the predominant species at the other stages of development (Fig. 1C, lanes 3 through 5). The Northern blotting data suggest that the alternative splicing scheme generating dMBD2/3 and dMBD2/3 $\Delta$  mRNAs is developmentally regulated and that the two proteins play different roles during development. dMBD2/3 is likely an essential gene for the early development of *Drosophila*, since a line with P-element insertion in the putative promoter region of dMBD2/3 is lethal to homozygous embryos (M.-S. Hung, unpublished data).

**dMBD2/3 $\Delta$  could preferentially recognize m<sup>5</sup>CpG-containing DNA.** Whether the *Drosophila* MBD protein binds to m<sup>5</sup>CpG is an interesting question. All vertebrate MBD pro-



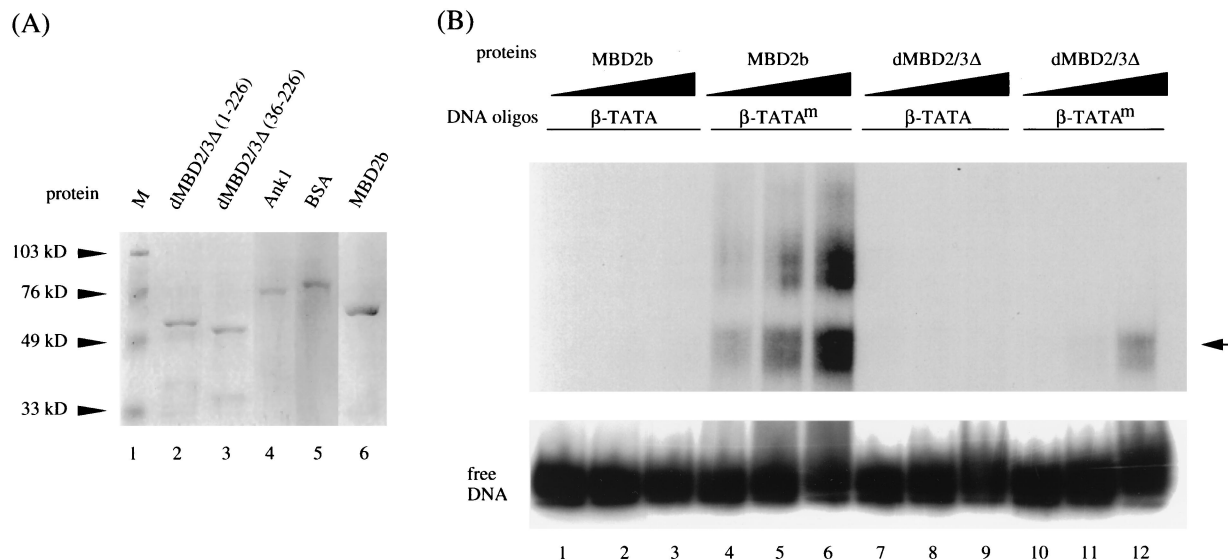


FIG. 2. Preferential binding of *Drosophila* and mammalian MBD proteins to  $m^5$ CpG-containing DNA. (A) Coomassie blue staining patterns of recombinant GST fusion polypeptides. One microgram each of GST fusions with the intact dMBD2/3 $\Delta$  (lane 2), dMBD2/3 $\Delta$  (36-226) (lane 3), a mouse ankyrin polypeptide (lane 4), and mouse MBD2b (lane 6) was electrophoresed on a denaturing sodium dodecyl sulfate-polyacrylamide gel and stained with Coomassie blue. Also run on the gel were molecular mass markers (lane 1) and bovine serum albumin (lane 5). (B) EMSA of recombinant GST fusions. The abilities of the *Drosophila* dMBD2/3 $\Delta$  and mouse MBD2b proteins to recognize  $m^5$ CpG were compared by EMSA. A 56-bp DNA probe ( $\beta$ -TATA) containing the human  $\beta$  globin TATA box with a single CpG was electrophoresed after incubation with competitor DNA and increasing amounts of proteins (2, 20, and 200 ng). Another probe,  $\beta$ -TATA $^m$ , which contains the same sequence as  $\beta$ -TATA but with the CpG site methylated at position 5 of the C residues, was analyzed in the same way. The oligonucleotides  $\beta$ -TATA and  $\beta$ -TATA $^m$  were used as the probes in an EMSA after binding with increasing amounts (2, 20, and 200 ng) of recombinant GST fusions of MBD2b or dMBD2/3 $\Delta$ . The arrow points to the complex formed between the  $\beta$ -TATA $^m$  oligonucleotide and both MBD2b and dMBD2/3 $\Delta$ . The GST protein did not form a complex either with  $\beta$ -TATA or with  $\beta$ -TATA $^m$  (data not shown).

teins identified thus far could recognize and preferentially bind to DNA containing  $m^5$ CpG (reference 4 and reference therein). Close sequence examination indicates that the MBD in dMBD2/3 is similar to those in vertebrates but with many amino acid differences, including relatively long deletions (Fig. 1B) (30, 33). Recombinant GST fusions of different MBD proteins (Fig. 2A) were thus compared for their abilities to bind methylated or unmethylated DNA (Fig. 2B). As expected from the previous studies (8), little binding of mouse MBD2b to unmethylated probe could be detected (Fig. 2B, lanes 1 through 3). On the other hand, it bound to the methylated DNA probe with relatively high affinity (Fig. 2B, compare lanes 4 through 6 to lanes 1 through 3).

Interestingly, *Drosophila* dMBD2/3 $\Delta$  also formed a specific complex with the methylated probe but not with unmethylated DNA (Fig. 2B, lanes 7 through 12). This complex comigrates with the fast band of the mouse MBD2b-DNA complexes (Fig. 2B, compare lanes 6 and 12). Quantitation of the respective complexes by PhosphorImager analysis (Fig. 2B) and experiments using lower concentrations of the recombinant proteins (data not shown) indicated that the affinity of binding of dMBD2/3 $\Delta$  to the methylated DNA probe was approximately 20-fold lower than that of the mouse MBD2b. This reduced affinity of binding of dMBD2/3 $\Delta$  to methylated DNA is likely due to the relatively divergent sequence and consequently the structure of the MBD in dMBD2/3 $\Delta$ . The preferential binding of dMBD2/3 $\Delta$  to DNA containing  $m^5$ CpG was also noted by others (A. Wolfe, personal communication).

**MBD is required for dMBD2/3 $\Delta$  binding to methylated DNA.** The specificity of the dMBD2/3 $\Delta$ -DNA complex was further confirmed by experiments shown in Fig. 3. First, as exemplified for a GST-ankyrin fusion polypeptide (Fig. 3A, lanes 1 and 2), proteins without MBD do not bind to the probe, methylated or unmethylated. Similarly, GST alone could not

form a complex with the same set of DNA oligonucleotides (data not shown). Second, the complex was effectively competed away upon the addition of a 100-fold molar excess of methylated probe or methylated pBluescript DNA (Fig. 3A, compare lane 3 to lanes 4 and 6) but not upon addition of the same amount of unmethylated competitors (Fig. 3A, lanes 5 and 7). Finally, removal of aa 1 to 35 from dMBD2/3 $\Delta$ , which comprise a major portion of the MBD-like sequence, abolished its ability to bind methylated DNA, as demonstrated by EMSA (Fig. 3B, compare lanes 2 and 3) and by Southwestern analysis (data not shown). It should be noted here that Tweedie et al. (30) have found no methyl-CpG-binding activity for either dMBD2/3 or dMBD2/3 $\Delta$  in their EMSA. However, in those cases, a different DNA oligonucleotide was used. It is likely that the recognition of  $m^5$ CpG by these two *Drosophila* MBD proteins is dependent on the DNA sequence context.

**dMBD2/3 $\Delta$  and mouse MBD2b could function as transcription corepressors or repressors in human 293T cells.** The mammalian MBD2a protein repressed promoter activity in a cotransfection assay (23). To examine whether the *Drosophila* MBD protein(s) could also act as a transcriptional corepressor, we first carried out side-by-side comparison of the effects of MBD2b, which is an isoform of mouse MBD2a, and of *Drosophila* dMBD2/3 $\Delta$  on the expression in human 293T cells of a cotransfected TK promoter with five Gal4-binding sites (Fig. 4B). The Gal4 fusions of either protein significantly repressed the TK promoter in a dosage-dependent manner (Fig. 4B, see data for pSG424-MBD2b and pSG424-dMBD2/3 $\Delta$ ) when being recruited to the promoter via the upstream Gal4-binding sites in pG5-TK-CAT. In contrast, expression of only the Gal4 (Fig. 4B, see data for pSG424) had no obvious effect and that of a fusion polypeptide between Gal4 and the activation domain of p45, a subunit of the erythroid enriched factor NF-E2,

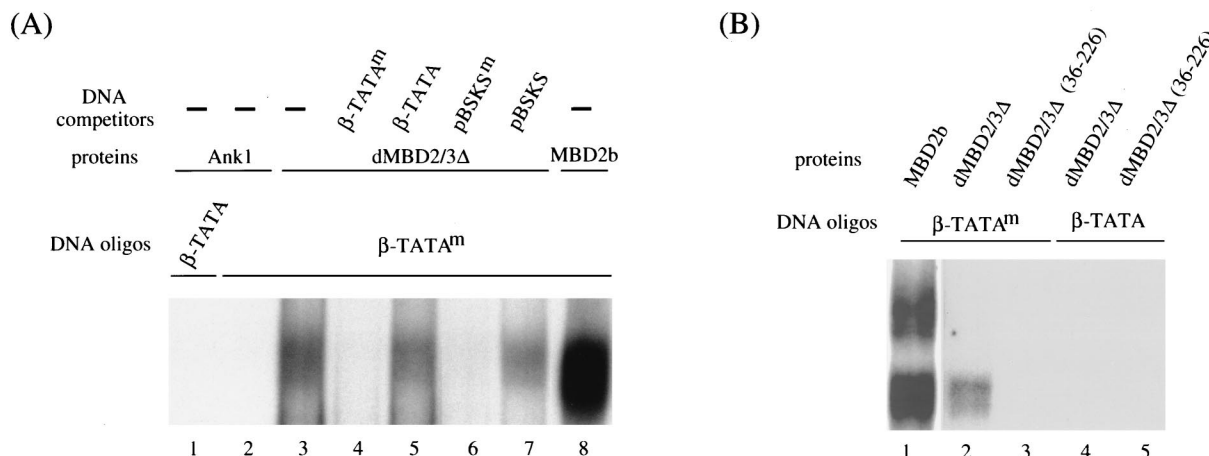


FIG. 3. (A) Oligonucleotide competition EMSA of dMBD2/3Δ.  $\beta$ -TATA and  $\beta$ -TATA<sup>m</sup> oligonucleotides were allowed to form complexes with 200 ng of the GST fusion of dMBD2/3Δ, 20 ng of the GST fusion of MBD2b, or 200 ng of the GST fusion of a mouse ankyrin polypeptide. The binding reactions were carried out without (–) or with the inclusion of 100 ng of competitor DNA oligonucleotide  $\beta$ -TATA,  $\beta$ -TATA<sup>m</sup>, plasmid pBSKS, or pBSKS methylated at all of the CpG sites (pBSKS<sup>m</sup>). The reaction products were then analyzed by EMSA. (B) Requirement of the MBD of dMBD2/3Δ for complex formation. The oligonucleotides were allowed to form a complex with 200 ng of the GST fusion of intact dMBD2/3Δ or with 200 ng of the GST fusion of a truncated dMBD2/3Δ.

stimulated the TK promoter activity (Fig. 4B, see data for pSG424-p45).

The strong repression effects by the MBD proteins were also observed in a separate set of experiments. In this case, a minimal promoter of the human  $\alpha$  globin gene with five Gal4-binding sites (Fig. 4A, see map) was used as the reporter. The expression of either Gal4-dMBD2/3Δ (Fig. 4C, lanes 2 through 4) or Gal4-MBD2b (Fig. 4C, lanes 6 through 8) significantly repressed, by more than 90%, the luciferase activity of pG5- $\alpha$ -Luc. Interestingly, removal of the MBD from dMBD2/3Δ did not affect its function as a transcription corepressor (Fig. 4C, compare lanes 4 and 5). When a reporter plasmid without the Gal4-binding sites, pG0- $\alpha$ -Luc, was used, the two MBD proteins could still repress the promoter activity but to a much smaller extent (approximately 50%) (Fig. 4C, compare lanes 10 through 16 and 2 through 8). This is consistent with our finding that, like for MeCP2 (20) but unlike for MBD1 (24), the repression effect of dMBD2/3Δ is dependent on its distance from the promoter (K. Roder, unpublished data).

The data in Fig. 4 indicate that in human cells, both *Drosophila* dMBD2/3Δ and mouse MBD2b could function as transcriptional corepressors of at least two different basic promoters. Furthermore, as exemplified for dMBD2/3Δ, the m<sup>5</sup>CpG recognition and transcriptional repression by these MBD proteins are exerted through two separate domains.

**TSA treatment could not inhibit the repressor or corepressor function of dMBD2/3Δ.** Whether the corepressor and repressor functions of dMBD2/3Δ or MBD2b are mediated through HDAC(s) has been tested with the use of TSA. In general, TSA treatment could increase the activities of many endogenous and transfected promoters (10, 32). We have observed the same phenomenon. As shown in Fig. 5A, the activity of p21, the promoter of the murine p21<sup>waf1</sup> gene, was greatly increased in a TSA concentration-dependent manner. This is consistent with the finding by Xiao et al. (35). Furthermore, a mutant promoter, p21-GCmut, with a mutation in the proximal GC box of p21, exhibited a much smaller magnitude of TSA inducibility (Fig. 5A). Similar to p21<sup>waf1</sup>, the activities of the TK promoter in plasmid pG0-TK-CAT or pG5-TK-CAT (Fig. 5B) and the  $\alpha$  globin promoter in pG0- $\alpha$ -Luc or pG5- $\alpha$ -Luc (Fig. 5C) were all increased after TSA treatment, 2- to 4-fold and 15- to 16-fold, respectively. However, it is obvious from

Fig. 5B and C that TSA treatment could not relieve the repression by dMBD2/3Δ, whether or not the reporter plasmids contained the Gal4-binding sites.

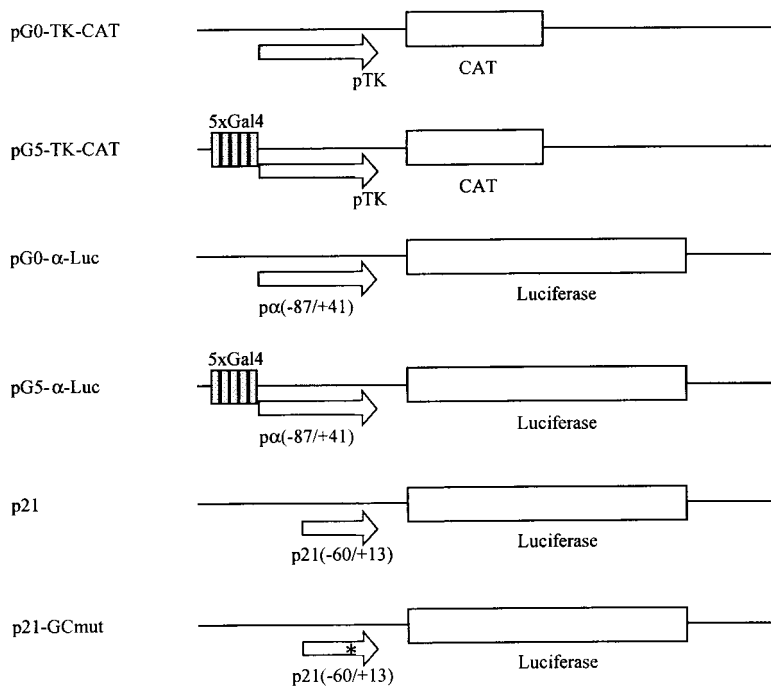
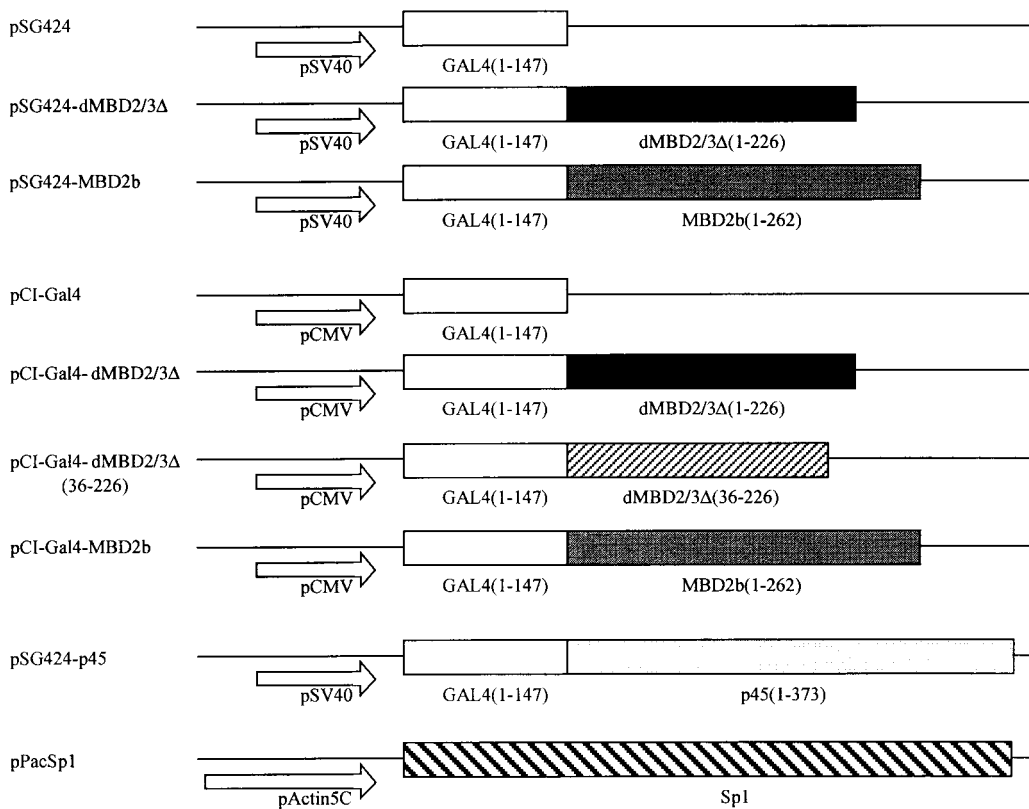
These results together indicated that dMBD2/3Δ and MBD2b could function in vivo as transcription corepressors without the involvement of HDAC. However, dMBD2/3Δ could associate with HDAC in vitro and in vivo (30). Thus, like mammalian MBD2a (23), *Drosophila* MBD proteins such as dMBD2/3Δ are also capable of functioning as transcription corepressors through either HDAC-independent or HDAC-dependent repression pathways in a promoter context-specific manner. This utilization of two alternative pathways to repress transcription has been implicated for at least one other *Drosophila* corepressor, Groucho (5).

**dMBD2/3Δ also could effectively repress transcription in *Drosophila* cells.** To test whether dMBD2/3Δ could function as a corepressor and/or repressor in a *Drosophila* cellular environment, we carried out DNA transfection in an SL2 cell culture. As in those experiments performed for Fig. 4B and C, pG5- $\alpha$ -Luc or pG0- $\alpha$ -Luc was cotransfected with pSG424 or pSG424-dMBD2/3Δ. In addition, pPacSp1 was also cotransfected since high-level activities of both the  $\alpha$  globin promoter in pG5- $\alpha$ -Luc or pG0- $\alpha$ -Luc and the simian virus 40 promoter in pSG424 require the binding of Sp1 factor, which *Drosophila* cells lack (6).

Indeed, as shown in Fig. 6, dMBD2/3Δ could effectively repress the  $\alpha$  globin promoter in pG5- $\alpha$ -Luc. Furthermore, as was true for 293T cells, dMBD2/3Δ also repressed transcription in the absence of the five Gal4-binding sites, although the magnitude of repression was smaller (Fig. 6). Experiments using pCI-Gal4-dMBD2/3Δ instead of pSG424-dMBD2/3Δ gave similar results (data not shown).

**Conservation of *Drosophila* proteins related to factors involved in the vertebrate DNA methylation program.** As already mentioned, the end product of the DNA methylation reaction, m<sup>5</sup>CpG, has been recognized as the key determinant responsible for the functioning of CpG MTases and DNA demethylases in vertebrate cells (reference 13 and reference therein). Among the possible mechanisms, m<sup>5</sup>CpG serves as the binding site for MBD proteins, many of which appear to reside within chromatin remodeling complexes that contain HDACs (12, 22, 23, 30). The identification of *Drosophila* proteins (DmMTR1

(A)

reporter plasmidsexpression plasmids

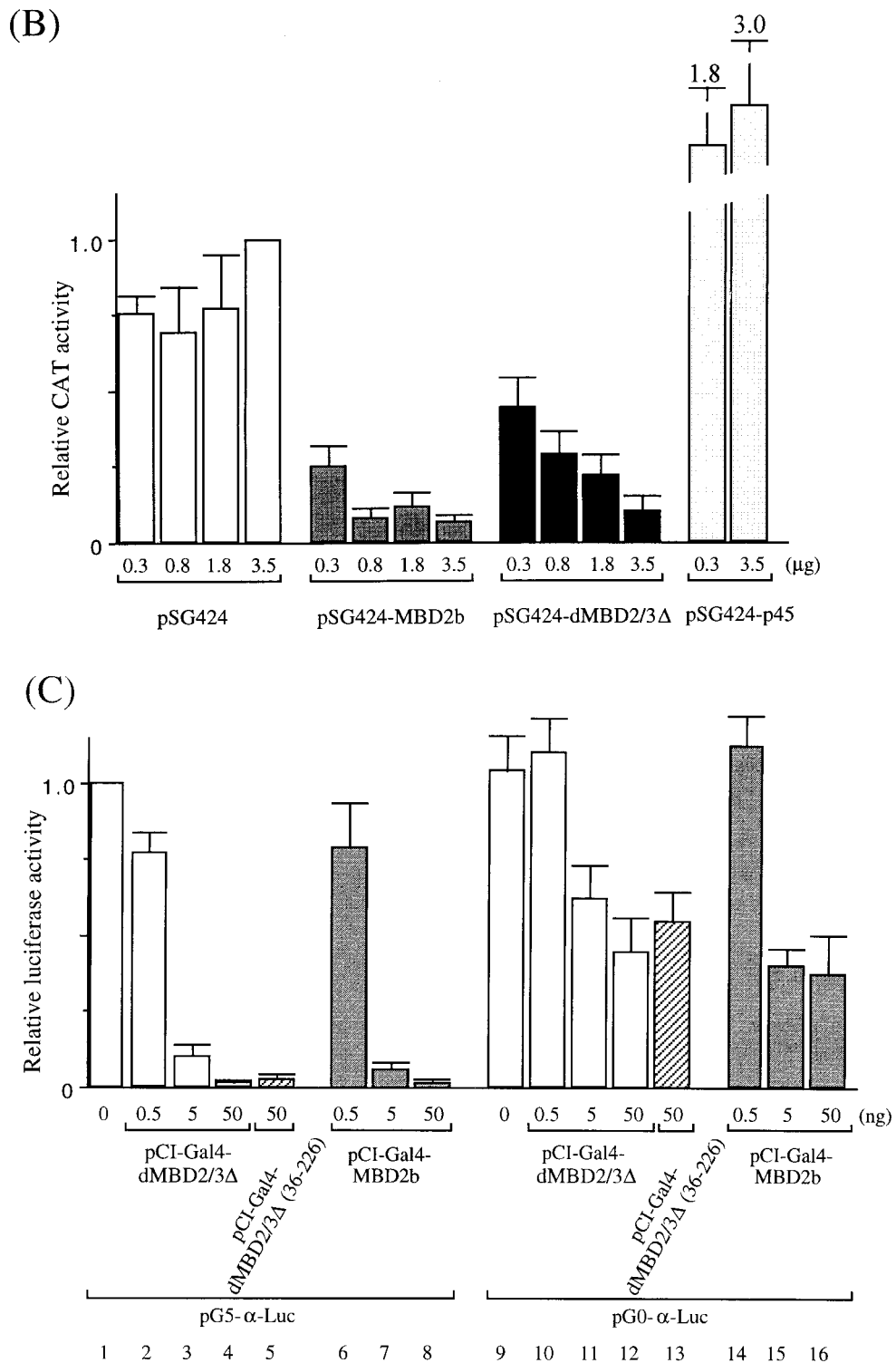


FIG. 4. Transcription repression by *Drosophila* and mammalian MBD proteins in human cells. (A) Maps of reporter and expression plasmids used in the cotransfection assays. For more details of the plasmids, see Materials and Methods. (B) Relative CAT activities from pG5-TK-CAT in human 293T cells cotransfected with increasing amounts of different recombinant expression plasmids. The CAT activities were standardized with the  $\beta$ -galactosidase activity from cotransfected pCMV- $\beta$ -gal. Activity with 3.5  $\mu$ g of cotransfected pSG424 is defined as 1. The bars represent standard deviations of two sets of transfection carried out in duplicate. (C) Relative luciferase activities from pG5- $\alpha$ -Luc or pG0- $\alpha$ -Luc in human 293T cells cotransfected with increasing amounts of different recombinant expression plasmids. The activation from pG5- $\alpha$ -Luc or pG0- $\alpha$ -Luc was measured and standardized with the  $\beta$ -galactosidase activity from cotransfected pCMV- $\beta$ -gal. The luciferase activity from pG5- $\alpha$ -Luc without cotransfected expression plasmid is defined as 1 (lane 1). Note that in contrast with activity shown in panel B, the expression of the Gal fusions in this set of experiments was directed by a CMV promoter in the pCI vector.

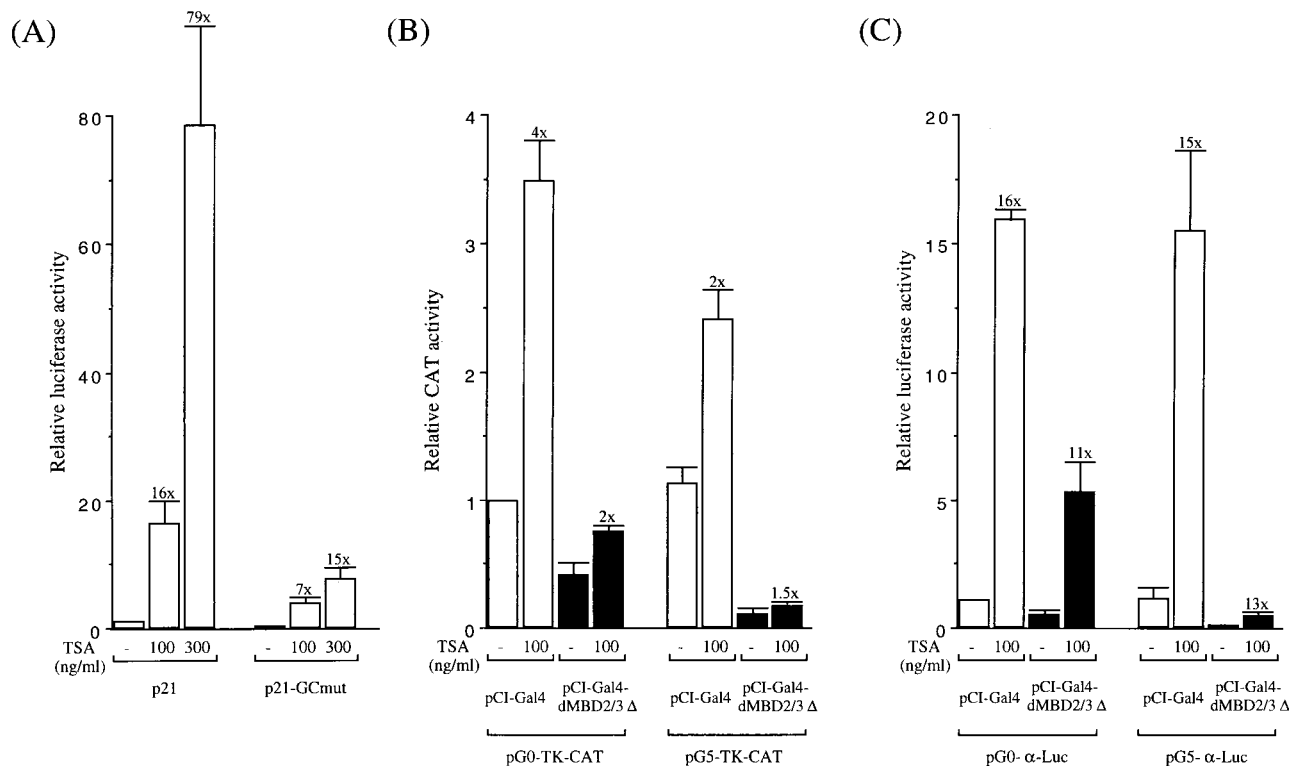


FIG. 5. Effects of TSA treatment on the repression function of dMBD2/3 $\Delta$ . (A) Activity increases of the p21<sup>wt</sup> and p21-GCmut promoters after treatment with the indicated concentrations of TSA. Increases relative to the promoter activity in the absence of TSA treatment are indicated above the bars. (B) Activity increases of the TK promoter by TSA treatment. The blank bars represent the TK promoter activities in the presence of exogenously expressed Gal4 protein, and the black bars represent activities with the Gal4 fusion of dMBD2/3 $\Delta$ . (C) Activity increases of the human  $\alpha$  globin promoter caused by TSA treatment.

[9] and DmMT2 [9, 30]) related to two of the vertebrate CpG MTase family members and of the *Drosophila* MBD proteins (9, 30, 33) has thus raised interesting questions regarding the mechanisms of functioning of these proteins in vertebrates as well as in *Drosophila* (9, 30). Two scenarios could be envisioned to explain the evolutionary conservation of the vertebrate CpG MTase-related proteins in *Drosophila*. First, DNA methylation-demethylation processes actually occur in *Drosophila* but only transiently at specific stages of the development or cell cycle, thereby escaping previous methods of detection (27, 31). Alternatively, these proteins and their mammalian counterparts carry out currently unknown functions not requiring the enzymatic reactions. Our data on dMBD2/3 $\Delta$  have interesting implications for the above models.

Previously, it was shown that ectopically expressed mammalian MBD proteins, in particular MBD1, could effectively repress transfected promoters in *Drosophila* cells, if the promoter DNA was methylated at CpGs prior to the transfection (7, 24). Furthermore, Lyko et al. (17) have reported that ectopic expression of mammalian CpG MTases caused genomic DNA methylation and embryonic lethality in *Drosophila*. The lethality is most likely due to the silencing of essential genes during *Drosophila* development, which in turn, as suggested by the authors, could be caused either by steric inhibition of transcription factor binding or through changes in the chromatin structure near the m<sup>5</sup>CpG residues. The capability of dMBD2/3 $\Delta$  to recognize and preferentially bind m<sup>5</sup>CpG-containing DNA (Fig. 2 and 3) lends support to the first scenario outlined above.

In this study, we have also demonstrated for the first time that *Drosophila* MBD proteins could function as transcription corepressors or repressors for unmethylated promoters in

mammalian as well as *Drosophila* cells (Fig. 4 and 6). This is remarkably similar to the behavior of the human MBD1 proteins in transfected mammalian and *Drosophila* cell cultures (7). Furthermore, TSA inhibition experiments suggested that

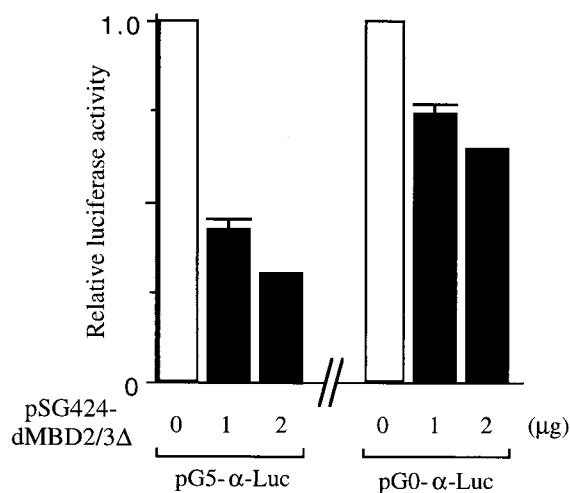


FIG. 6. Transcriptional repression by dMBD2/3 $\Delta$  in *Drosophila* SL2 cells. The relative luciferase activities from pG5- $\alpha$ -Luc or pG0- $\alpha$ -Luc in SL2 cells cotransfected with pSG424 or pSG424-dMBD2/3 $\Delta$  were calculated. Activity with 1  $\mu$ g of cotransfected pSG424 is defined as 1. The  $\alpha$  globin promoter and simian virus 40 promoter in pG0- $\alpha$ -Luc and pG5- $\alpha$ -Luc and in pSG424, respectively, were driven by Sp1 factor expressed by cotransfected pPacSp1. Indeed, the luciferase activities were at least 10 times lower without cotransfection with pPacSp1 (data not shown).



repression by dMBD2/3 $\Delta$  was mostly HDAC independent (Fig. 5), as previously observed for repression of certain promoters by the mammalian MBD proteins (23, 36). In this regard, it should be noted that certain HDACs, such as the Sir2 HDAC, could not be inhibited by TSA (9a). Since dMBD2/3 $\Delta$  could associate with several types of HDACs (30), it may also repress transcription through HDAC-dependent pathways. In any case, our data also provide strong support for the second scenario, i.e., that *Drosophila* CpG MTase- and MBD protein-related factors, as well as their mammalian homologs, carry out functions, e.g., repression of transcription, not requiring the presence of m<sup>5</sup>CpG residues.

In summary, we have shown that *Drosophila* MBD proteins such as dMBD2/3 $\Delta$  could recognize m<sup>5</sup>CpG through the MBD motif. dMBD2/3 $\Delta$  also acts as a potent repressor in both *Drosophila* and mammalian cells. These data together with the previous reports indicate that a nearly full complement of factors related to those involved in the vertebrate DNA methylation program are expressed in *Drosophila*. Interestingly, as already mentioned above, the functional conservation of the MBD proteins is further supported by the finding that a *Drosophila* strain carrying a P-element insertion in the dMBD2/3 promoter is lethal to homozygous embryos (Hung, unpublished). Molecular genetic analysis of these proteins in *Drosophila* will provide interesting and essential insights into the mechanistic aspects of the functions of CpG MTases, MBD proteins, and their related factors in eukaryotic cells.

#### ACKNOWLEDGMENTS

The first four authors contributed equally to this work.

We are grateful to Shigeo Hayashi at the Genetic Strain Research Center, National Institute of Genetics, Japan, for providing the P1 genomic clones and to Alan Wolffe for communicating unpublished results. We also thank Duen-wei Hsu, Xin Chen, Mark Ptashne, Narendra Gava, Bon-chu Chung, and Guntram Suske for some of the vectors and clones used in the study.

K. Roder is the recipient of a DFG postdoctoral fellowship (Ro 2121/1-1). This research was supported by the Academia Sinica, the National Science Council, and the Foundation of Biomedical Sciences, Taipei, Taiwan, Republic of China.

#### REFERENCES

- Baylin, S. B. 1997. Tying it all together: epigenetics, genetics, cell cycle, and cancer. *Science* **277**:1948–1949.
- Bestor, T. H., and G. L. Verdine. 1994. DNA methyltransferases. *Curr. Opin. Cell Biol.* **6**:380–389.
- Bhattacharya, S. K., S. Ramchandani, N. Cervoni, and M. Szyf. 1999. A mammalian protein with specific demethylase activity for mCpG DNA. *Nature* **397**:579–583.
- Bird, A., and A. P. Wolffe. 1999. Methylation-induced repression—belts, braces, and chromatin. *Cell* **99**:451–454.
- Chen, G., J. Fernandez, S. Mische, and A. J. Courey. 1999. A functional interaction between the histone deacetylase Rpd3 and the corepressor Groucho in *Drosophila* development. *Genes Dev.* **13**:2218–2230.
- Courey, A. J., and R. Tjian. 1988. Analysis of Sp1 in vivo reveals multiple transcriptional domains, including a novel glutamine-rich activation motif. *Cell* **55**:887–898.
- Fujita, N., S.-I. Takebayashi, K. Okumura, S. Kudo, T. Chiba, H. Saya, and M. Nakao. 1999. Methylation-mediated transcriptional silencing in euchromatin by methyl-CpG binding protein MBD1 isoforms. *Mol. Cell. Biol.* **19**:6415–6426.
- Hendrich, B., and A. Bird. 1998. Identification and characterization of a family of mammalian methyl-CpG binding proteins. *Mol. Cell. Biol.* **18**:6538–6547.
- Hung, M. S., N. Karthikeyan, B. Huang, H.-C. Koo, J. Kiger, and C.-K. J. Shen. 1999. *Drosophila* proteins related to vertebrate DNA (5-cytosine) methyltransferases. *Proc. Natl. Acad. Sci. USA* **96**:11940–11945.
- Imai, S.-I., C. M. Armstrong, M. Kaerberlein, and L. Guarente. 2000. Transcriptional silencing and longevity protein Sir2 is an NAD-dependent histone deacetylase. *Nature* **403**:795–800.
- Jin, S., and K. W. Scotto. 1998. Transcriptional regulation of the MDR1 gene by histone acetyltransferase and deacetylase is mediated by NF-Y. *Mol. Cell. Biol.* **18**:4377–4384.
- Jones, P. A., and M. L. Gonzalzo. 1997. Altered DNA methylation and genome instability—a new pathway to cancer? *Proc. Natl. Acad. Sci. USA* **94**:2103–2105.
- Jones, P. L., G. J. C. Veenstra, P. A. Wade, D. Vermaak, S. U. Kass, N. Landsberger, J. Strouboulis, and A. P. Wolffe. 1998. Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. *Nat. Genet.* **19**:187–191.
- Jost, J. P., and H. P. Saluz (ed.). 1993. DNA methylation: molecular and biological significance. Birkhäuser Verlag, Basel, Switzerland.
- Kudo, S. 1998. Methyl-CpG-binding protein MeCP2 represses SP1-activated transcription of the human leukosialin gene when the promoter is methylated. *Mol. Cell. Biol.* **18**:5492–5499.
- Laird, P. W., and R. Jaenisch. 1996. The role of DNA methylation in cancer genetic and epigenetics. *Annu. Rev. Genet.* **30**:441–464.
- Luo, R. X., A. A. Postigo, and D. C. Dean. 1998. Rb interacts with histone deacetylase to repress transcription. *Cell* **92**:463–473.
- Lyko, F., B. H. Ramsahoye, H. Kashevsky, M. Tudor, M.-A. Mastrangelo, T. L. Orr-Weaver, and R. Jaenisch. 1999. Mammalian (cytosine-5) methyltransferases cause genomic DNA methylation and lethality in *Drosophila*. *Nat. Genet.* **23**:363–366.
- MacGregor, G. R., and C. T. Caskey. 1989. Construction of plasmids that express *E. coli* beta-galactosidase in mammalian cells. *Nucleic Acids Res.* **17**:2365.
- Mertineit, C., J. A. Yoder, T. Taketo, D. W. Laird, J. M. Trasler, and T. H. Bestor. 1998. Sex-specific exons control DNA methyltransferase in mammalian germ cells. *Development* **125**:889–897.
- Nan, X., F. J. Campoy, and A. Bird. 1987. MeCP2 is a transcriptional repressor with abundant binding sites in genomic chromatin. *Cell* **88**:471–481.
- Nan, X., R. R. Meehan, and A. Bird. 1993. Dissection of the methyl-CpG binding domain from the chromosomal protein MeCP2. *Nucleic Acids Res.* **21**:4886–4892.
- Nan, X., H.-H. Ng, C. A. Johnson, C. D. Laherty, B. M. Turner, R. N. Eisenman, and A. Bird. 1998. Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature* **393**:386–389.
- Ng, H.-H., Y. Zhang, B. Hendrich, C. A. Johnson, B. M. Turner, H. Erdjument-Bromage, P. Tempst, D. Reinberg, and A. Bird. 1999. MBD2 is a transcriptional repressor belonging to the MeCP1 histone deacetylase complex. *Nat. Genet.* **23**:58–61.
- Ng, H.-H., P. Jeppesen, and A. Bird. 2000. Active repression of methylated genes by the chromosomal protein MBD1. *Mol. Cell. Biol.* **20**:1394–1406.
- Okano, M., S. Xie, and E. Li. 1998. Dnmt2 is not required for de novo and maintenance methylation of viral DNA in embryonic stem cells. *Nucleic Acids Res.* **26**:2536–2540.
- Okano, M., D. W. Bell, D. A. Haber, and E. Li. 1999. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* **99**:247–257.
- Patel, C. V., and K. P. Gopinathan. 1987. Determination of trace amounts of 5-methylcytosine in DNA by reverse-phase high-performance liquid chromatography. *Anal. Biochem.* **164**:164–169.
- Sadowski, I., and M. Ptashne. 1989. A vector for expressing GAL4(1-147) fusions in mammalian cells. *Nucleic Acids Res.* **17**:7539.
- Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
- Tweedie, S., H.-H. Ng, A. L. Barlow, B. M. Turner, B. Hendrich, and A. Bird. 1999. Vestiges of a DNA methylation system in *Drosophila melanogaster*? *Nat. Genet.* **23**:389–390.
- Urieli-Shoval, S., Y. Gruenbaum, J. Sedat, and A. Razin. 1982. The absence of detectable methylated bases in *Drosophila melanogaster* DNA. *FEBS Lett.* **146**:148–152.
- Van Lint, C., S. Emiliani, and E. Verdin. 1996. The expression of a small fraction of cellular genes is changed in response to histone hyperacetylation. *Gene Expr.* **5**:245–253.
- Wade, P. A., A. Geggone, P. L. Jones, E. Ballestar, F. Aubry, and A. P. Wolffe. 1999. Mi-2 complex couples DNA methylation to chromatin remodelling and histone deacetylation. *Nat. Genet.* **23**:62–66.
- Walsh, C. P., and T. H. Bestor. 1999. Cytosine methylation and mammalian development. *Genes Dev.* **13**:26–34.
- Xiao, H., T. Hasegawa, and K.-I. Isobe. 1999. Both Sp1 and Sp3 are responsible for p21<sup>waf1</sup> promoter activity induced by histone deacetylase inhibitor in NIH3T3 cells. *J. Cell. Biochem.* **73**:291–302.
- Yu, F., J. Thiesen, and W. H. Stratling. 2000. Histone-deacetylase-independent transcriptional repression by methyl-CpG-binding protein 2. *Nucleic Acids Res.* **28**:2201–2206.