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

KARIM RODER,1 MING-SHIU HUNG,1 TAI-LIN LEE,1 TZU-YANG LIN,2 HENGYI XIAO,3 KEN-ICHI ISOBE,3 JYH-LYH JUANG,2 AND C.-K. JAMES SHEN1,∗

Institute of Molecular Biology, Academia Sinica, Nankang,1 and Division of Genomic Medicine, National Health Research Institute,2 Taipei, Taiwan, Republic of China, and Department of Basic Gerontology, National Institute for Longevity Sciences, Oga, Aichi 474-8522, Japan3

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 m5CpG. 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, DmMBD2/3Δ could preferentially recognize m5CpG-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 m5CpG, 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 m5CpG 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 m5CpG residues. Indeed, m5CpG 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 m5CpG-containing DNA, recruit histone deacetylase (HDAC)-containing complexes, and thus cause chromatin condensation in the vicinity of m5CpG (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 syncyntial *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 interacts with 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 m5CpG (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 m5CpG-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.
before the transfection, 226), pCI-Gal4-MBD2b, or pSG424-p45. pSG424 and/or pCI-Gal4 was used as carrier, and different amounts of an expression plasmid [pSG424-dMBD2/3 or p21-GCmut], 0.3 μg of salmon sperm DNA as the internal control for Northern blotting data suggest that the alternative splicing scheme is likely an essential gene for the early development of Drosophila, since a line with P-ele-
FIG. 1. Structure and expression of the Drosophila dMBD2/3 gene. (A) Genomic and cDNA organization of dMBD2/3. The chromosomal location (85D) of the gene and the three P1 clones spanning the region are indicated. Shown below the P1 clones are the cDNAs for dMBD2/3 and dMBD2/3D, which were derived from the cDNA clones LD46808 and LD03777, respectively. The base numbering of the two cDNAs corresponds to that of the genomic sequence shown in panel B. (B) Genomic sequence of dMBD2/3. The sequence of the dMBD2/3 gene was deduced from the three P1 clones. The putative site of transcriptional initiation, as inferred from the sequences of the two cDNA clones, is designated 1. The two introns and the regions upstream and downstream of the gene are shown with lowercase letters. The corresponding amino acids are also shown, with those homologous to the conserved amino acids among the vertebrate MBDs indicated by bold letters. The putative TATA box in the upstream promoter region and the poly(A) signal, AATAAA, are underlined. (C) Expression of dMBD2/3 and dMBD2/3D mRNA. Drosophila poly(A) RNAs isolated from different stages of development were analyzed by Northern blot hybridization with the cDNA clone LD03777 as the probe. The two transcripts (1,270 and 890 nt long) are indicated.
proteins identified thus far could recognize and preferentially bind to DNA containing m^5CpG (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 used 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 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 to the methylated DNA probe was approximately 20-fold lower than that of the mouse MBD2b. This reduced affinity of binding of dMBD2/3 to methylated DNA is likely due to the relatively divergent sequence and consequently the structure of the MBD in dMBD2/3. The preferential binding of dMBD2/3 to DNA containing m^5CpG was also noted by others (A. Wolffe, personal communication).

**MBD is required for dMBD2/3 binding to methylated DNA.** The specificity of the dMBD2/3-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, 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 protein(s) could also act as transcriptional 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 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 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,
stimulated the TK promoter activity (Fig. 4B, see data for pSG424-p45).

The strong repression effects by the MB proteins were also observed in a separate set of experiments. In this case, a minimal promoter of the human α-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-α-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-α-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 McCP2 (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 m5CpG recognition and transcriptional repression by these MB 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-WAF1 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-WAF1, the activities of the TK promoter in plasmid pG0-TK-CAT or pG5-TK-CAT (Fig. 5B) and the α-globin promoter in pG0-α-Luc or pG5-α-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 MB proteins such as dMBD2/3Δ also are 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-α-Luc or pG0-α-Luc was cotransfected with pSG424 or pSG424-dMBD2/3Δ. In addition, pPacSp1 was also cotransfected since high-level activities of both the α-globin promoter in pG5-α-Luc or pG0-α-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 α-globin promoter in pG5-α-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 pC1-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, m5CpG, 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, m5CpG 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 plasmids

- pG0-TK-CAT
- pG5-TK-CAT
- pG0-α-Luc
- pG5-α-Luc
- p21
- p21-GCmut

expression plasmids

- pSG424
- pSG424-dMBD2/3Δ
- pSG424-MBD2b
- pCI-Gal4
- pCI-Gal4-dMBD2/3Δ
- pCI-Gal4-dMBD2/3Δ (36-226)
- pCI-Gal4-MBD2b
- pSG424-p45
- pPacSp1
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 β-galactosidase activity from cotransfected pCMV-β-gal. Activity with 3.5 μg 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-α-Luc or pG0-α-Luc in human 293T cells cotransfected with increasing amounts of different recombinant expression plasmids. The activation from pG5-α-Luc or pG0-α-Luc was measured and standardized with the β-galactosidase activity from cotransfected pCMV-β-gal. The luciferase activity from pG5-α-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.
[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 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 $\text{m}^3\text{CpG}$ residues. The capability of dMBD2/3 to recognize and preferentially bind $\text{m}^3\text{CpG}$-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...
repression by dMBD2/Δ3 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/Δ3 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, i.e., repression of transcription, not requiring the presence of m5CpG residues.

In summary, we have shown that Drosophila MBD proteins such as dMBD2/Δ3 could recognize m5CpG through the MBD motif. dMBD2/Δ3 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 Wolfe for communicating unpublished results. We also thank Duen-wei Hsu, Xin Chen, Mark Ptashne, Narendr Gavva, Bon-chu Chung, and Guntram Suske for some of the results. We also thank Duen-wei Hsu, Xin Chen, Mark Ptashne, Na-nan, H.-H. Ng, C. A. Johnson, C. D. Laherty, B. M. Turner, N. R. Eisenman, and A. Bird. 1998. Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. Nature 393: 386–389.


