Gene 420 (2008) 125–134

Contents lists available at ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

cDNA cloning, expression, and characterization of methyl-CpG-binding domain type 2/3 proteins from starfish and sea urchin

Masahito Matsumoto, Tetsuo Toraya *

Department of Bioscience and Biotechnology, Graduate School of Natural Science and Technology, Okayama University, Tsushima-naka, Okayama 700-8530, Japan

ARTICLE INFO

Article history:
Received 26 March 2008
Accepted 13 May 2008
Available online 23 May 2008

Received by I. King Jordan

Keywords:
DNA methylation
Methyl-CpG binding domain
MBD2/3 proteins
Starfish
Sea urchin

ABSTRACT

Two kinds of cDNAs that are highly homologous to mammalian MBD2 and MBD3 cDNAs were cloned from ovary of the starfish *Asterina pectinifera*. They are splicing variants and designated sMBD2/3a and sMBD2/3b cDNAs. sMBD2/3a cDNA spans 1378 bp and consists of a 48-bp upstream untranslated region, a 807-bp open reading frame encoding sMBD2/3a, and a 523-bp downstream untranslated region. sMBD2/3a and sMBD2/3b cDNAs encode proteins with predicted molecular weights of 30,724 and 29,635 consisting of 268 and 260 amino acid residues, respectively. The deduced amino acid sequences of these two are identical from residues 1 to 255, but different from residues 256 to the C-terminal ends. sMBD2/3a is expressed in all the tissues of starfish, whereas sMBD2/3b is highly expressed in ovary and oocytes, slightly in testis, but not in somatic cells. As suggested from the whole-genome sequence of *Strongylocentrotus purpuratus*, a sea urchin MBD2/3 cDNA was cloned from eggs of *Hemicentrotus pulcherrimus* and designated suMBD2/3 cDNA. It encodes a protein with predicted molecular weight of 30,778 consisting of 274 amino acid residues. All the three echinodermal MBD2/3 proteins consist of a methy-CpG-binding domain (MBD) and a coiled-coil domain, and only sMBD2/3a contains a glutamate-rich C-terminal region, a key mark in vertebrate MBD3. The three MBD2/3 proteins expressed in *Escherichia coli* and purified to homogeneity were capable to bind specifically to methylated DNA. It was shown that sMBD2/3a exists as dimer or in the monomer-dimer equilibrium, whereas sMBD2/3b and suMBD2/3 exist as monomer and dimer, respectively.

1. Introduction

Genomic DNA in mammalian cells commonly undergoes methylation of cytosine residues at C5 in CpG sequences. In the human genome, it has been reported that about 85% of CpGs are methylated, whereas the majority of unmethylated CpGs are located in CpG islands (Razin, 1998; Ng and Bird, 1999). However, the methylation status of genomic DNA differs among animal species. The percentage of CpG methylation is 60–90% in vertebrates (Singer et al., 1979), 10–40% in sea anemone (coelenterate), mussel (mollusc), and sea urchin (echinoderm) (Bird and Taggart, 1980; Bird et al., 1979), and very low or not detectable in insects (Rae and Steele, 1979) and nematodes (*Caenorhabditis elegans*) (Simpson et al., 1986). DNA methylation patterns are also different — global methylation in the mammalian genome and mosaic methylation in the invertebrate genome (Suzuki et al., 2007; Simmen et al., 1999; Tweedie et al., 1997; Bird et al., 1979).

DNA methylation is an epigenetic mechanism for transcriptional gene silencing, involved in genomic imprinting (Bartolomei and Tilghman, 1997), X chromosome inactivation (Heard et al., 1997), and tumor-suppressor gene inactivation in cancer cells (Herman and Baylin, 2000), and essential for normal mammalian development. A family of MBD proteins that contain the methyl-CpG-binding domain (MBD) and specifically recognize methyl-CpGs are present in vertebrates. These include MBD1–MBD4 and MeCP2. MBD1, MBD2 and MeCP2 contain transcriptional repression domain (TRD) in addition to MBD and mediate transcriptional repression by targeting chromatin remodeling co-repressor complexes to methylated DNA regions (Nan et al., 1998; Ng et al., 1999). MeCP2 forms the part of the Sin3/HDAC complex, and MBD1 links DNA methylation with histone H3-K9 methyltransferase activity by SETDB1 (Sarraf and Stancheva, 2004). MBD2 and MBD3 are core subunits of the nucleosome remodeling and histone deacetylation (NuRD) complex (Zhang et al., 1999). Mammalian MBD3 contains amino acid substitutions that prevent binding to methyl-CpG, whereas the amphibian MBD3 binds methylated DNA (Wade et al., 1999). MBD4 has a C-terminal glycosylase domain in addition to MBD and was characterized for its role in DNA repair (Hendrich et al., 1999). Among these MBD proteins, MBD2 and MBD3 share extensive sequence similarity outside the MBD motif (Hendrich and Bird, 1998) and are the most highly conserved and widely distributed in Metazoa.

In contrast, invertebrates seem to contain a smaller number of *mbd*-like genes. Only one gene that is similar to the mammalian *mbd2* and *mbd3* genes has been identified and designated *mbd2/3* in

Abbreviations: H6-tag, His_6-tag; IPTG, isopropyl-β-D-thiogalactopyranoside; MBD, methyl-CpG binding domain; SDS, sodium dodecyl sulfate; UTR, untranslated region.

* Corresponding author. Tel.: +81 86 251 8194; fax: +81 86 251 8264.
E-mail address: toraya@cc.okayama-u.ac.jp (T. Toraya).

0378-1119/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.gene.2008.05.007

Drosophila melanogaster (Roder et al., 2000). Whole genome sequences suggest that invertebrates have only MBD2/3 protein, while vertebrates do not. The MBD2/3 is thus considered to be the original MBD protein (Hendrich and Tweedie, 2003). The *D. melanogaster* MBD2/3 protein binds to methylated DNA very weakly (Roder et al., 2000) because its MBD is severely disrupted. It has been proposed that Drosophila MBD could function as transcription corepressor or repressor for unmethylated promoters. In contrast, the nematodes MBD2/3-like protein lacks MBD (Gutierrez and Sommer, 2004, 2007). It has been assumed that nematode MBD2/3 proteins function independent of DNA methylation ranging from the indispensable to the non-essential.

Information about MBD proteins is important to understand the role of DNA methylation. Although the methylation of genomes and genes is different between invertebrate and vertebrates (Tweedie et al., 1997), no papers have been published so far that report on MBD proteins in Echinoderms. As the first step to answer the question whether an epigenetic DNA methylation-dependent mechanism for transcriptional gene silencing is operational in Echinoderms through MBD proteins, we wish to report the cDNA cloning and characterization of starfish MBD proteins. Very recently, whole genome sequence of *Strongylocentrotus purpuratus* has been published that suggests the presence of MBD2/3 protein in sea urchin. So, both starfish and sea urchin MBD proteins were prepared by heterologous expression of their cDNAs in *Escherichia coli*, purified and characterized. Tissue distributions of the MBD2/3 proteins in these animals are also described here.

2. Materials and methods

2.1. Materials

Oligonucleotide primers and oligonucleotides containing 5-methylcytosine were synthesized by Invitrogen and Sigma Aldrich Japan Genosys, respectively. Alkaline-phosphatase-conjugated goat anti-rabbit IgG was purchased from Kirkegaard & Perry Laboratories Inc., USA.

2.2. Animals

The starfish *Asterina pectinefera* were collected near Asamushi, Aomori and Hashirimizu, Kanagawa, Japan, and the sea urchin *Hemicentrotus pulcherrimus* near Ushimado, Okayama, Japan in their breeding seasons. They were kept in laboratory aquaria supplied with circulating artificial seawater (Roht Pharmaceutical Co. Ltd., Osaka, Japan) at 14 °C.

2.3. Bacterial strains, plasmids, DNA manipulations, and nucleotide sequencing

E. coli JM109, BL21(DE3), and DH5αMCR strains were used as hosts for plasmids pCR-XL-TOPO (Invitrogen), pGAD10 (Clontech), and expression plasmids pET15b (Novagen) and pF7KC (Promega). Standard recombinant DNA techniques described by Sambrook et al. (1989) were used. Restriction endonucleases and other enzymes for the construction of plasmids were used according to the manufacturer's instructions. Template plasmid DNA was prepared by an alkaline lysis method and sequenced on an ABI PRISM 310 Genetic Analyzer using a BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems).

2.4. cDNA synthesis

Total RNA was extracted from ovarian tissues, immature oocytes, and other tissues of *A. pectinifera* or eggs of *H. pulcherrimus* by the acid guanidinium thiocyanate-phenol chloroform (AGPC) method. A mixture of heat-denatured total RNA and NotI-$(dT)_{18}$ Primer (Invitrogen) was first incubated at 42 °C for 5 min. Superscript II reverse transcriptase (Invitrogen) was then added, and the mixture was incubated at 42 °C for 50 min. After the reaction was terminated by heating at 70 °C for 15 min, the resulting cDNA-RNA hybrids were digested with ribonuclease H (Invitrogen) at 37 °C for 20 min. The oligo(dT)-primed cDNAs thus obtained were stored at −20 °C and used as templates for PCR. For 5'-RACE, sumbd-2r was used as a gene-specific primer.

2.5. Construction of starfish ovary cDNA library

The poly(A)$^+$ RNA was purified using oligo(dT)-Latex (Takara) from total RNA from starfish ovary. The starfish ovary cDNA library in the pGAD10 vector was constructed from 5 μg of the poly(A)$^+$ RNA using a two-hybrid cDNA library construction kit (Clonetech) according to the manufacturer's instructions (Miyake et al., 2001).

2.6. Pcr

DNA segments were amplified by PCR with Taq (Sigma), GoTaq Flexi (Promega), KOD Dash (Toyobo), or ExTaq (Takara) DNA polymerases using a template and appropriate primer pairs listed in Table 1. PCR products were cloned into the pCR-XL-TOPO vector using a TOPO-XL-PCR Cloning kit (Invitrogen).

2.7. 3'- and 5'-RACE

The nucleotide sequences of the 3'-UTR and 3'-terminal region of cDNAs were analyzed by 3'-RACE. 3'-Terminal region and 3'-UTR were amplified by PCR with KOD Dash or Taq DNA polymerase using oligo (dT)-primed cDNAs, forward primer smbd23-4f or sumbd-1f in the coding region, and reverse primer 3'-RACEr anchoring at the *Not*I-

Table 1
Oligonucleotides used for cDNA cloning and RT-PCR

Name	Sequence	Location[a] or comments
mbd23-d1r	5'GGYTTRCTICKRAAYTTCTTYCC	121–143
smbd23-1f	5'TTGGTGGTCCGTGAGTACG	−38–20
smbd23-2f	5'CCTTTTTACCTAAAGGGTGGAAG	29–51
smbd23-3f	5'GTGTACCTGAATACAGAGCAGCC	628–649
smbd23-4f	5'GCAAGAGGAGCGTGTTAAGGAGG	693–712
smbd23-5f	5'GACAGACAGCGTCAATCTTTAAGC	308–331
smbd23-6r	5'CCTCTGGTAACTCCATGGATTG	475–496
smbd23-7f	5'GGAGGAAGTTTCGCAGTAA	122–140
AD1	5'CTATTCGATGATGAAGATACCCCAC-CAAACC	From pGAD10 cDNA library kit
AD2	5'GTGAACTTGCGGGGTTTTTCAGT ATCTACGAT	From pGAD10 cDNA library kit
sgfsmbd23-f	5'*GCGATCGC*CATGGAGAAGAAAAAGAC-GGAAGAATG	1–26 containing *Sgf*I site
pmesmbd23b-r	5'*GTTTAAAC*GGAGCCTGCAGTACC TTGTATC	761–782 containing *Pme*I site
ndesmbd23-f	5'AC*CATATG*GAGAAGAAAAAGACGGAAG	1–22 containing *Nde*I site
ecosmbd23a-r	5'*GAATTC*ACTCTGATCCCTCCTC	790–818 containing *Eco*RI site
ecosmbd23b-r	5'G*GAATTC*TAGGAGCCTGCAGTAC	770–786 containing *Eco*RI site
smbd23a-r	5'TCACTCTGATCCCTCCTCTTC	786–817
smbd23b-r	5'AACACTGGTGTAGTCCAAGTCC	790–821
satp6-f	5'CTGGTTATCTATTTCCAAACCATC	96–119 for starfish ATPase 6
satp6-r	5'GAGTATGTTAATGATATGTGACTC	300–323 for starfish ATPase 6
NotI-(dT)18	5'AACTGGAAGAATTCGCGGCCG CAGGAA(T)18	For cDNA synthesis
sumbd-1f	5'GCGCGTTTGACTTCCGAACG	226–246
sumbd-2r	5'CTGTTTGGGAGTTTCCTTTGTGTCTTG	433–459
ndesumbd23-f	5'CTGCATATGGTACCAATGGATTTACC	−1–20 containing *Nde*I site
ecosumbd23-r	5'GAATTCCTAAAGCACCAGGCCTTGTATG	804–825 containing *Eco*RI site
3'-RACEr	5'TGGAAGAATTCGCGGCCGCAG	From *Not*I-(dT)18
AAP	5'GGCCACGCGTCGACTAGTACGGGIIG-GGIIGGGIIG	For 5'RACE
AUAP	5'GGCCACGCGTCGACTAGTAC	For 5'RACE

[a] Numbers from the start sites of cDNAs.