Available online at www.sciencedirect.com
ScienceDirect
Biochimica et Biophysica Acta 1769 (2007) 287–294
Biochimica et Biophysica Acta
BBA
www.elsevier.com/locate/bbaexp

Review

Methyl-CpG-binding domain (MBD) proteins in plants

Gideon Grafi [a,*], Assaf Zemach [b], Letizia Pitto [c,*]

[a] Albert Katz Department of Dryland Biotechnologies, Jacob Blaustein Institutes for Desert Research, Ben-Gurion University of the Negev, Midreshet Ben-Gurion 84990, Israel
[b] Department of Plant Sciences, The Weizmann Institute of Science, Rehovot 76100 Israel
[c] Laboratorio di Terapia Genica e Molecolare-Istituto di Fisiologia Clinica-Area della Ricerca CNR, via Moruzzi 1, 56100 Pisa, Italy

Received 23 September 2006; received in revised form 1 February 2007; accepted 16 February 2007
Available online 1 March 2007

Abstract

Cytosine methylation is the most prevalent epigenetic modification of plant nuclear DNA, which occurs in symmetrical CpG or CpNpG as well as in non-symmetrical contexts. Intensive studies demonstrated the central role played by cytosine methylation in genome organization, gene expression and in plant growth and development. However, the way by which the methyl group is interpreted into a functional state has only recently begun to be explored with the isolation and characterization of methylated DNA binding proteins capable of binding 5-methylcytosine. These proteins belong to an evolutionary conserved protein family initially described in animals termed methyl-CpG-binding domain (MBD) proteins. Here, we highlight recent advances and present new prospects concerning plant MBD proteins and their possible role in controlling chromatin structure mediated by CpG methylation.

Keywords: Cytosine methylation; Methyl-CpG-binding domain (MBD) protein; Heterochromatin; Decrease in DNA Methylation 1 (DDM1); Arginine methyltransferase (PRMT); Nucleolar organizing region (NOR); Plant

1. DNA methylation

Methylation is a common epigenetic modification of DNA found in the genomes of most organisms, both prokaryotes and eukaryotes. There are three major types of naturally occurring methylated bases in DNA, namely, N^6-methyladenine, N^4-methylcytosine and 5-methylcytosine [1]; the latter is the most prevalent methylated base in nuclear DNA of higher eukaryotes. Methylation of cytosine is catalyzed by specific DNA methyltransferases [2–4], which transfer a methyl group from the donor *S*-adenosyl-L-methionine to the 5 position of the pyrimidinic ring. Methylation on cytosine is perpetuated by maintenance DNA methyltransferases when occurring at symmetrical contexts, i.e., CpG and CpNpG (where N is any nucleotide) [4]. Cytosine methylation may affect gene expression in two ways: (1) it can render transcription factors incapable of binding their DNA recognition sequence [5,6], and (2) it can be targeted by specific 5-methylcytosine binding proteins to induce nucleosome repositioning or the formation of repressive chromatin [7,8].

2. DNA methylation in plants

While cytosine methylation in animals is prevalent in symmetrical CpG dinucleotides, in plants, it is often found in symmetrical CpG and CpNpG contexts as well as in non-symmetrical CpHpH (H=C, A, T) [9,10]. The distribution of cytosine methylation within the nucleus both in plants and animals is not random but rather sequence/domain specific, that is, most methylated cytosines are found at heterochromatic regions, which are often enriched with repetitive DNAs [11]. Notably, repetitive DNAs such as the centromeric 180 bp repeats as well as rDNAs are mostly methylated at CpG context. However, cytosine methylation, though abundant at chromocenters, is also scattered along chromosome arms; genome-wide high-resolution mapping of DNA methylation in *Arabidopsis thaliana* has recently revealed that over one third of expressed genes contain methylation within transcribed regions [12]. The

* Corresponding authors.
E-mail addresses: ggrafi@bgu.ac.il (G. Grafi), l.pitto@ifc.cnr.it (L. Pitto).

0167-4781/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.bbaexp.2007.02.004

biological importance of cytosine methylation in plants was first inferred from its dynamics during development and in response to pathogen infection [13–15] as well as from its correlation with transposon and transgene silencing [16,17]. Also, treatments with the hypomethylating agent 5-azacytidine as well as genetic manipulation of DNA methylation levels have demonstrated the central role played by cytosine methylation in genome organization, gene expression and plant development. Accordingly, treatment with 5-azacytidine activated imprinted genes in the polyploid hybrid Triticale (a wheat × rye hybrid) [18], induced heritable dwarfism in rice seedlings [19], and caused early flowering in winter wheat and in certain ecotypes of *Arabidopsis* where flowering requires long exposure to cold (vernalization) [20,21]. Likewise, a significant reduction in total genomic cytosine methylation brought about by genetic manipulation of the activity of the DNA methyltransferase MET1 (At5g49160) or by mutating the gene encoding the SWI2/SNF2 nucleosomal remodeling factor DDM1 (At5g66750) has led to transposon activation and to developmental abnormalities including changes in flowering time, and defects in leaf and flower structure [22–26]. In spite of intensive studies of DNA methylation in plants, the way by which the methyl group is interpreted into basic cellular functions has remained unknown until the recent isolation and characterization of the Arabidopsis methyl-CpG-binding domain (MBD) proteins [27–30]. Here we discuss recent findings concerning the Arabidopsis MBD proteins and their role in mediating the effects of CpG-methylation on chromatin structure and gene expression.

3. Mammalian MBD proteins

The cloning of MeCP2, a mammalian gene whose product is capable of binding methylated CpG sites [31,32], provided new insight into the molecular mechanisms involved in the interpretation of cytosine methylation into functional states. The minimal domain possessing methyl-CpG binding activity termed MBD (*M*ethyl-CpG-*B*inding *D*omain) was dissected from MeCP2 and found to consist of 85 amino acids [33]. Later on, the solution structure of the MBD motif of the human MBD1 in complex with methylated DNA was resolved by NMR spectroscopy revealing five highly conserved amino acid residues that form a hydrophobic patch that mediates the recognition with methyl-CpG dinucleotides [34]. In humans, there are four MBD proteins named MBD1–4 in addition to the founding member MeCP2; all, but MBD3, specifically recognize and bind methylated CpG sites [reviewed in 35]. In addition to the MBD motif, mammalian MBD1, MBD2 and MeCP2 contain an active transcription repression domain (TRD) capable of long-range repression in vivo [36]. The mammalian MBD4 represents a unique protein having a thymine glycosylase activity that mediates G:T mismatch repair often found at methylated CpG regions [37]; this mismatch results from spontaneous hydrolytic deamination that converts 5-methylcytosine into thymine [38]. In addition to MBD3 whose MBD motif is not functional, the human genome possesses six additional genes encoding putative MBD proteins termed TAM (TIP5, ARBP, MBD), which were predicted to be non-functional in binding methylated CpG sites [35]. Besides MBD proteins, a family of proteins named Kaiso-like lacking the MBD motif were found to bind methylated CpG dinucleotides through their zinc-finger domains [39,40]. MBDs may act by recruiting a variety of histone deacetylases (HDACs), histone methyltransferases, and chromatin remodeling factors to methylated CpG sites leading to the formation of transcriptionally repressed chromatin [41,42]. The biological significance of MBD proteins is demonstrated in the Rett syndrome (after Andreas Rett, an Austrian pediatrician who first described this disorder in 1966), a childhood neuro-developmental disorder resulting from mutations in the gene encoding the MBD transcriptional repressor MeCP2 [43,44].

4. Methylated DNA binding proteins in plants

To understand the way by which the DNA methylation signal in plants is interpreted into basic nuclear functions, attempts have been made to characterize plant nuclear proteins capable of binding 5-methylcytosine. In pea, a 5-methylcytosine specific, DNA binding protein designated DBP-m was found to bind 5-methylcytosine in any context [45]. Partial characterization revealed a 70–90 kDa protein whose high binding affinity to 5-methylcytosine is dependent mainly on the number of 5-methylcytosine residues present [46]. Similarly, methylated DNA binding activities were found in nuclear extracts of soybean seed, cauliflower florets, corn seed, wheat germ and carrot [47,48]. In carrot, two classes of proteins were identified: dcMBP1 (*Daucus carota* methylated DNA binding protein 1), with high affinity for sequences containing 5-methylcytosine in the canonical CpG context, and dcMBP2 (*D. carota* methylated DNA binding protein 2) with lower affinity for 5-methyl CpG dinucleotides but with higher affinity for DNA sequences containing 5-methylcytosine in both CpHpH and CpNpG contexts [48]. Thus far, the molecular identity of any of these proteins remained unknown.

5. Plant MBD proteins

Sequencing of plant genomes and the subsequent identification of genes encoding putative MBD proteins provided new insight into the way cytosine methylation is interpreted in plants. The Plant Chromatin Database (http://www.chromdb.org) lists 13 MBD genes in *A. thaliana*, 17 in rice (*Oryza sativa* ssp. Japonica), 14 in maize (*Zea mays*) and 14 genes in poplar (*Populus trichocarpa*). Based on sequence similarity within the MBD motif, plant MBD proteins were divided into 8 subclasses (classification is based on ChromDB, http://www.chromdb.org [27,49]). Bioinformatics analyses pointed to evolutionary divergence of dicots MBD proteins from those of monocots. Hence, subclasses IV (AtMBD5 and AtMBD6) and VI (AtMBD7) – whose MBD motifs show the highest similarity to human functional MBD motifs – appear to be unique to dicots [49]. Based on sequence similarity and intron position, AtMBD5, AtMBD6 and AtMBD7 were suggested to have originated from a common ancestor [29]. AtMBD7, containing three MBD motifs represents a unique type of MBD protein not