Isolation and Expression Analysis of CaMBD1 Gene Encoding Methyl-CpG-binding Domain Proteins in Red Pepper (Capsicum annum L.)

Yu Jin Jung1, Keun Hyang Lee1, Jang Sun Choi2, and Kwon Kyoo Kang1*

1Department of Horticulture, Hankyong National University, Ansung, 456-749, Korea
2National Institute of Horticultural & Herbal Science, Suwon, 440-706, Korea

ABSTRACT In this study, we isolated cDNA with open reading frame encoding putative methyl-binding domain proteins from red pepper, which was designated as CaMBD1 (HQ171162). BLASTX search and phylogenetic analysis suggested that the CaMBD1 gene belonged to AtMBD10 group (subclass I) of MBD family. The expression profile of the CaMBD1 was studied via Q-RT-PCR and the results indicated that the CaMBD1 were differentially expressed in detected red pepper tissues. It was interesting to note that CaMBD1 was highly expressed in dry seeds and endosperms. Moreover, the differential expression pattern of CaMBD1 was observed in leaves and roots under water-stress. Also a GFP-Ca MDB1 fusion construct introduced into the onion epidermal cells confirmed localization of CaMBD1 into the nuclei. To investigate the biological significance of CaMBD1 proteins, we transformed Arabidopsis using CaMBD1 gene. The resulting 35S::CaMBD1 plants showed a variety of phenotypic effects including aerial rosettes, serrated leaves, abnormal position of flowers, fertility problems and late flowering. Arabidopsis lines involved in chromatin remodeling show similar phenotypes. Our results suggest an important role of CaMBD1 biological in plant growth and development.

Keywords Gene expression, MBD gene, Red pepper, RT-PCR, Transgenic plant

INTRODUCTION DNA methylation is one of the most abundant epigenetic modifications in higher plants and animals. It plays an important role in regulating developmental processes (Finnegan et al. 2000). Cytosine methylation may affect gene expression in two ways: (1) it can render transcription factors incapable of binding their DNA recognition sequence (Tariq and Paszkowski 2004), and (2) it can be targeted by specific 5-methylcytosine binding proteins to induce nucleosome repositioning or the formation of repressive chromatin (Eden and Cedar 1994; Kass et al. 1997). The discovery of a group of proteins containing methyl-CpG-binding domain (MBD) greatly contributed to the understanding of methylation-induced gene silence. MBD proteins appeared to act as structural proteins, which recruit a variety of histone deacetylase (HDAC) complexes and chromatin remodeling factors, leading to chromatin compaction and consequently, to transcriptional repression (Ehrlich 1993; Ballestar and Wolffe 2001; Adams et al. 2007). In plants, recent studies show that genome-wide demethylation has pleiotropic effects on morphology. A pea (Pisum sativum) DNA-binding protein that can recognize m5C was identified and partially purified (Ehrlich 1993), and two groups of methylated DNA-binding proteins, dcMBP1 and dcMBP2, were detected in carrot (Daucus carota) (Nikitina et al. 2007a). However, the molecular identities of these proteins have not been determined. Based on the amino acid sequence homology with animal MBDS, the plant chromatin database (http://www.chromdb.org/) lists 13 genes encoding putative MBD proteins in Arabidopsis, 17 genes in rice (Oryza sativa L.) and 16 genes in maize (Zea mays L.) (Anita et al. 2003). Sequence comparisons suggest that a number of the...
MBD proteins in *Arabidopsis* are unlikely to bind methylated CpGs. Sequence comparisons suggest that a number of the MBD proteins in *Arabidopsis* are unlikely to bind methylated CpGs. In this article, we reported the isolation and expression analysis of *CaMBD1* gene encoding methyl-CpG-binding domain proteins in red pepper (*Capsicum annum* L.). Multiple alignment analysis showed that the *CaMBD1* motifs were most similar to AtMBD10 (Class I) of *Arabidopsis* MBD family. As a first approach to investigate the biological significance of *CaMBD1* gene, *Arabidopsis* was transformed with a 35S::*CaMBD1* construct, resulting in plants with pleiotropic developmental alterations.

**MATERIALS AND METHODS**

**Plant materials**

Red pepper seeds were surface-sterilized in 2% sodium hypochlorite for 15 minutes and washed four times (2 min each) with sterilized water, then germinated in petri dishes with water (16 h photoperiod at 20°C). The seedlings were cultivated under a controlled condition (26°C day/20°C night, 16 h photoperiod, 75% relative humidity) and watered every two days by MS liquid medium. Water deficit was induced by 20% polyethylene glycol (PEG, molecular weight, 6000) for various periods (0.5 h, 1 h, 2 h, 6 h, 12 h, 24 h and 48 h) treatments with three-leaf stage seedlings, and some seedlings were left in normal nutrient solution as control. All materials were prepared by three independent treatments for three replicates. In addition, dry seeds, leaves, stems, roots, inflorescences, embryos and endosperm tissues of the red pepper variety ‘Bukang’ were collected for analyzing the spatial expression pattern of *CaMBD1* with three replicates.

**BLAST-based and database searches**

Database searches were performed using BLASTP, TBLASTN and PSI-BLAST (http://www.ncbi.nlm.nih.gov/BLAST/). Protein domains were identified using the programs RPSBLAST (NCBI) and ProfileScan (http://hits.isb-sib.ch/cgi-bin/PFSCAN) searching the Pfam-A, Prosite profiles and Smart databases (NCBI). Amino acid sequence alignments were created using ClustalX 1.8 (http://www-igbmc.u-strasbg.fr/BioInfo/ClustalX) with default parameters and manual adjustments from GeneDoc 2.6.001 (http://www.psc.edu/biomed/genedoc/).

**RNA isolation and first strand cDNA synthesis**

Total RNAs from various rice tissues were extracted using Trizol reagent (Invitrogen) according to the manufacturer’s protocol. The RNA was sequentially treated with DNase I (Promega, WI, USA) at 37 °C for 15 min in order to remove the remaining genomic DNA. The first strand cDNA was synthesized with 2 μg of purified total RNA using the Access RT-PCR system (Promega, WI, USA) according to the manufacturer’s protocol.

**Isolation of full-length cDNA of *CaMBD1***

The *CaMBD1* fragment was identified in dbEST division of GenBank by BLAST program with MBD protein motifs as probe sequence. The full-length cDNA was isolated by RACE technology with SMART RACE-PCR kit (Clontech, Palo Alto, CA, USA).

**Semi-quantitative RT-PCR assay**

As an internal control, an 228-bp PCR fragment of the actin gene ACT1 (For: 5’- TGTTATGGTAGGGATGGG-GTC-3’, Rev: 5’-TTTCTTCATATTGCGTACGG-3’) in red pepper was amplified from various red pepper samples. The primers for *CaMBD1* (For: 5’-ATGGAGAAGAATTCAAGTTTTG-3’, Rev: 5’-TTATGGTCTGCAGTTCTCAGTTCCC TTAC-3’) was used for expression analysis. PCR was done with 25 cycles to amplify *CaMBD1* and ATC1. The PCR products were separated on a 1% agarose gel and then visualized under UV light.

**Nuclear localization assays**

For onion epidermis nuclear localization assays (Baumbusch et al. 2001), the entire coding region of *CaMBD1* was amplified by PCR and inserted into the SalI and Ncol sites of psGFP (S65T) vector to yield the fusion in frame with GFP. Particle bombardment was performed on a Bio-Rad Biolistic PDS 1000/He system to introduce the fusion construct into onion (*Allium cepa* L.) epidermal cells. The
initial pressure of bombardment was 1,100 psi and the traveling distance of the particles to the plant tissues was 6 cm. Bombarded tissues were placed on the water plates and incubated at 4°C for 24 h in the light, followed by monitoring the localization of GFP with a confocal microscope (AX70; Olympus, Tokyo, Japan).

**Generation of CaMBD1 transgenic Arabidopsis plants**

To generate transgenic *Arabidopsis* plants, the full length *CaMBD1* was amplified from the plasmid containing *CaMBD1* with two primers (forward: 5'-AGAGGATC-CAAGATGGAGAATCAAGTTTTG-3', reverse: 5'-ATAGAGCT CTTATGGTCTGCAGTCCCTAC-3') and inserted into the BamH I and Sac I sites of the plant binary vector pBI121. The *CaMBD1* gene driven under the cauliflower mosaic virus (CaMV) 35S promoter was then introduced into *Arabidopsis* plants via *Agrobacterium* mediated transformation (Clough and Bent 1998).

**Northern hybridization**

RNA blot analysis was performed using total RNA extracted from samples by the AGPC method (Nakano et al. 2000). Approximately 10-µg of RNA was separated by electrophoresis, transferred to N’nylon membrane, and subjected to hybridization with radioactively labeled probes as described (Yoda et al. 2002).

---

**Fig. 1.** Comparison of 12 MBD proteins. (A) Amino acid alignment of MBD region. The MBDs from human, *Arabidopsis*, maize, rice and red pepper proteins were aligned using ClustalW. This alignment was shaded using Boxshade. Above and below the alignment are schematic diagrams indicating the structural regions determined for the MeCP2 and MBD1 proteins (Ohki et al., 1999; Wakefield et al., 1999). The position of three β-sheets, an α-helix, and a loop region are shown. The positions of five amino acids thought to be critical for the ability to bind methylated DNA are indicated by the arrow. (B) phylogenetic tree. The relationship was calculated from whole sequences for each MBD protein. The branch lengths represent numbers of substituted residues per site.
RESULTS

Cloning and characterization of CaMBD1

The CaMBD1 gene was identified by database searching against red pepper dbEST division in GenBank with MBD protein motifs of Arabidopsis MBDs as a probe sequence. This EST sequences that showed high similarity to known MBD cDNAs from Arabidopsis MBDs were obtained by TBLASTX searches in the NCBI database. The full-length cDNA of CaMBD1 were isolated by RACE-PCR. CaMBD1 was 1,050 bp in length, consisting of a 186 bp 5’ untranslated region (UTR), a complete ORF of 669 bp encoding a polypeptide of 222 amino acids and a 3’-UTR of 195 bp. Except MBD conserved regions, some other structures were also found in CaMBD1, such as a CAMP factor. Like other plant MBDs, the deduced CaMBD1 proteins all lack the transcriptional repressor domain, which is typically contained in MeCP2 and MBD1 of mammals. The alignment based on the MBD domains was shown in Fig. 1-A, and the structural characteristics of MBDs were annotated. Generally, the MBDs can fold into a sandwich structure comprising a layer of twisted sheet, backed by another layer formed by one helix and a hairpin loop at the C terminus. These layers are both amphipathic, with the α1 helix and the sheet lying parallel and the hydrophobic faces tightly packed against each other. The sheet is composed of two long inner strands (β2 and β3) sandwiched by two shorter outer strands (α1 and β4). Phylogenetic tree analysis of the aligned MBD motif sequences shows that CaMBD1 was most similar to AtMBD10 and AtMBD11 of Arabidopsis MBD motifs (Fig. 1-B) suggesting functional equivalence.

CaMBD1 is localized to the nucleus

To investigate the cellular localization of the CaMBD1 protein, the CaMBD1 gene was cloned into a pBIN 35S-m-GFP vector downstream of a constitutive CaMV 35S promoter and upstream of a GFP gene to create a CaMBD1-GFP fusion. These constructs were introduced into the inner epidermis of onion cells using particle bombardment. After 24 hours, free GFP protein was distributed throughout the cells (nucleus, cytoplasm, plasmalemma) of the control whereas the CaMBD1-GFP fusion protein was limited to nuclei (Fig. 2).

Expression of CaMBD1 in different organs

The spatial expression pattern of the CaMBD1 was characterized via quantitative real-time RT-PCR and it was found that CaMBD1 was differentially expressed in detected tissues (root, stem, leaf, inflorescence, dry seeds, embryo,

---

![Fig. 2](image-url) Nuclear localization of GFP-CaMBD1 fusion protein. Onion bulbs were bombarded with gold particle coated with GFP-CaMBD1.
endosperm, radicle and plantule). The expression level of CaMBD1 was higher in dry seeds and endosperm tissues than that in other tissues (Fig. 3).

**Expression of CaMBD1 under water-stress**

To determine whether the CaMBD1 gene are responsive to water-stress in red pepper, the expression of CaMBD1 in roots and leaves under the PEG induced water-stress was monitored by quantitative real-time RT-PCR under water-stress. Leaves and roots from PEG stressed red pepper seedlings were used. CK, un-stressed control; 1–7, indicated the 0.5 h, 1 h, 2 h, 6 h, 12 h, 24 h and 48 h time points after water-stress respectively.
detected via quantitative real-time RT-PCR. The results showed that the CaMBD1 gene was differentially expressed under water-stress, while expression patterns in roots were different from that in leaves (Fig. 4). In roots, the CaMBD1 was down-regulated in general, and its expression level was significantly lower at the 0.5 h, 1 h, 2 h, 12 h and 48 h time points after water-stress than that of control. These results indicated that CaMBD1 might be involved in the water-stress response in red pepper.

**Transgenic Arabidopsis with CaMBD1 expression**

To investigate the function of CaMBD1 in plant, Arabidopsis was transformed with a 35S::CaMBD1 construct. Thirty-six transformants were obtained and confirmed through PCR using construct-specific primers (data not shown). A number of aberrant phenotypic traits were observed (Fig. 5). Wild-type plants did not show such phenotypes. A third of the transgenic plants displayed late flowering compared to wild-type (Table 1), i.e. they flowered from 20 to 31 days after transfer to soil, with an average of 22.1 ± 4.1 days. Wild-type plants, on the other hand, flowered after 16.2 ± 3.2 days. Eleven transgenic plants developed a majority of siliques that were abnormally small. In two of these plants the siliques hardly developed

| Phenotype                          | No. of plants |
|------------------------------------|---------------|
| Reduced seed production            | 9             |
| Many small siliques                | 6             |
| Serrated leaves                    | 6             |
| Late flowering                     | 25            |
| Extra rosettes                     | 3             |
| Reduced apical dominance           | 3             |
| Combined phenotypes                | 5             |
| Normal                             | 8             |

**Fig. 5.** Phenotypes of Arabidopsis plants transformed with a 35S::CaMBD1 construct.
Isolation and Expression Analysis of CaMBD1 Gene Encoding Methyl-CpG-binding Domain Proteins in Red Pepper (Capsicum annum L.)

Fig. 6. Northern hybridization using mRNA from rosette leaves of wildtype (wt) and four independent 35S::CaMBD1 primary transformants, cf. numbering and phenotypes in Figure 5.

at all (Fig. 5A-D). The infertility was most likely a result of mechanical male sterility (Fig. 5E-H). Six independent transgenic lines (T1-T6) that were analyzed showed the presence of the CaMBD1 specific transcript (Fig. 6). The signal was variable among different transformants indicating possible position effect due to random insertions of the foreign gene into the host plant genome.

DISCUSSION

DNA methylation of cytosine residues is suggested to play important roles in regulating gene expression and plant development (Finnegan et al. 2000). The structural and functional analysis of the key proteins, which specifically bind to methylated DNA regions, is a critical step toward understanding such epigenetic phenomenon (Grossniklaus et al. 2001; Gehring et al. 2004; Horike et al. 2005). Until now, genome sequencing and annotation indicated that 13 and 17 MBD genes exist in Arabidopsis and rice, respectively. Also 16 MBD genes isolated from maize showed spatial expression patterns (Springer and Kaeppler 2005; Nikitina et al. 2007b). Consider revising this line the analysis of the cloned CaMBD1 showed the conserved motifs of MBD protein. Additionally, it also contain a Zinc finger CW-type, the function of this structure is not yet known, but its taxonomic distribution and its association with a variety of domains involved in DNA-binding, methylation, chromatin structure or ubiquitination suggested that the Zinc finger CW-type structure might play a certain important role in epigenetic mechanisms (Klose et al. 2005; Kim et al. 2007; Ho et al. 2008). The study of MBD’s expression pattern is helpful in understanding epigenetic regulation in plant development. Our RT-PCR analysis does not allow a rigid quantitative comparison of the expression levels between tissues, but demonstrates that all the active genes are expressed in flowers and buds. In other parts of the plant, the genes appear to be differentially expressed. Genes within the same subgroup have similar, but not identical, expression patterns, suggesting that each gene has a specialised function (Springer and Kaeppler 2005; Schule et al. 2007). We also found that CaMBD1 was abundantly expressed in dry seeds and endosperm, this spatial expression pattern might be related to the selectively gene silence in endosperm tissue, this might agree to the view point that the endosperm, a seed tissue that mediates the transfer of nutrients from the maternal parent to the embryo, is an important site of imprinting in flowering plant (Gehring et al. 2004; Horike et al. 2005). Grossniklaus et al. (2001) also studied the imprinted loci and found that the allelic copy with reduced or eliminated expression had higher levels of DNA methylation in endosperm. Water-stress is one of the major environmental conditions that adversely affect plant growth and crop yield (Bartels and Nelson 1994; Ozturk et al. 2002). Various genes are involved in water deficit response and their differential expression is suggested to contribute to the water stress tolerance in plants (Xiong et al. 2002; Zheng et al. 2004; Shinozaki and Yamaguchi-Shinozaki 2007). However, the regulatory mechanism of gene expression under water-stress, especially those involved in the DNA methylation, is still an area to be elucidated. We have found that the patterns of cytosine methylation were different under water-stress in red pepper (our unpublished data). Also we have investigated the effect of transforming Arabidopsis with a 35S::CaMBD1 construct to explore CaMBD biological function. The majority of the primary
transformants (78%) showed one or more aberrant phenotypes and a third of the plants flowered significantly later than wild-type plants. There are several lines of evidence for epigenetic control of flowering time. Repression of the FLC gene by cold treatment (vernalisation) accelerates flowering in late flowering ecotypes and a nuclear localised zinc finger protein (VRN2) with similarity to polycomb group proteins is involved in maintenance of the repressed state (Anita et al. 2003). In this study, we also found that four CaMBD1 was differentially expressed after water-stress in wheat leaves, while only three of them responded to dehydration in roots. In addition, the expression profiles in roots were different from that in leaves. The results agreed with other reports that there were different water-stress responses between leaves and roots in maize and barley (Zheng et al, 2004; Ozturk et al. 2002). The expression patterns can be observed in leaves than that in roots might indicate that CaMBD1 take part in more than one regulatory pathway in water-stress response in leaf tissues. The cloning and characterization of CaMBD1 would provide a new starting point to understand regulatory mechanisms of red pepper development and water-stress response.

REFERENCES

Adams VH, McBryant SJ, Wade PA, Woodcock CL and Hansen JC. 2007. Intrinsic disorder and autonomous domain function in the multifunctional nuclear protein, MeCP2, J. Biol. Chem. 282:15057-15064.

Anita B, Trine JM, Mirela M, Tage T, Kjetil K and Reidunn BA. 2003. Ten members of the Arabidopsis gene family encoding methyl-CpG-binding domain proteins are transcriptionally active and at least one, AtMBD11, is crucial for normal development. Nucleic Acids Research Vol. 31 18:5291-5304.

Ballestar E and Wolffe AP. 2001. Review article: Methyl-CpG-binding proteins: Targeting specific gene repression. Eur J Biochem. 268:1-6.

Bartels D and Nelson DE. 1994. Approaches to improve stress tolerance using molecular genetics. Plant Cell Environ. 17:659-667.

Baumbusch LO, Thorstensen T, Krauss V, Fischer A, Naumann K, Assalkhou R, Schulz I, Reuter G and Aalen RB. 2001. The Arabidopsis thaliana genome contains at least 29 active genes encoding SET domain proteins that can be assigned to four evolutionarily conserved classes. Nucleic Acids Res. 29:4319-4333.

Clough SJ and Bent AF. (1998) Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. Plant J. 16:735-743.

Eden S and Cedar H. 1994. Role of DNA methylation in the regulation of transcription. Curr Opin Genet Dev 4:225-259.

Ehrlich KC. 1993. Partial purification of a pea seed DNA-binding protein that specifically recognizes 5-methycytosine. Prep Biochem 23:423-438.

Finnegan EJ and Kovac KA. 2000. Plant DNA methyltransferases. Plant Molecular Biology 43:189-201.

Gehring M, Choi Y and Fischer R. 2004. Imprinting and seed development. Plant Cell 16:S203-S213.

Grossniklaus U, Spillane C, Page DR and Kohler C. 2001. Genomic imprinting and seed development: endosperm formation with and without sex. Curr. Opin. Plant Biol. 4:21-27.

Horike S, Cai S, Miyano M, Cheng JF and Kohwi-Shigematsu T. 2005. Loss of silentchromatin looping and impaired imprinting of DLX5 in Rett syndrome. Nat. Genet. 37: 31-40.

Ho KL, McNae IW, Schmiedeberg L, Klose RJ, Bird AP and Walkinshaw MD. 2008. MeCP2 binding to DNA depends upon hydration at methyl-CpG. Mol. Cell 29:525-531.

Kass SU, Pruss D and Wolffe AP. 1997. How does DNA methylation repress transcription? Trends Genet 13:444-449.

Kimbrough D, Barrera LO and Raben B. 2007. ChiP-chip for genome-wide analysis of protein binding in mammalian cells. Curr. Protoc. Mol. Biol. 21: (Chapter 21, Unit 13).

Klose RJ, Sarraf SA, Schmiedeberg L, McDermott SM, Stancheva I and Bird AP. 2005. DNA binding selectivity of MeCP2 due to a requirement for A/T sequences adjacent to methyl-CpG. Mol. Cell 19:667-678.

Nakano Y, Steward N, Koizumi N and Sano H. 2000. A tobacco NtMET1 encoding a DNA methyltransferase: molecular characterization and abnormal phenotypes of antisense transgenic tobacco plants. Plant Cell Physiol. 41:448-457.

Nikitina T, Shi X, Ghosh RP, Horowitz-Scherer RA, Hansen JC and Woodcock CL. 2007a. Multiple modes of interaction between the methylated DNA binding protein MeCP2 and chromatin. Mol. Cell. Biol. 27:864-877.

Nikitina T, Ghosh RP, Horowitz-Scherer RA, Hansen JC,
Grigoryev SA and Woodcock CL. 2007b. MeCP2-chromatin interactions include the formation of chromatosome-like structures and are altered in mutations causing Rett syndrome. J. Biol. Chem. 282:28237-28245.

Ohki I, Shimotake N, Fujita N, Nakao M and Shirakawa M. 1999. Solution structure of the methyl-CpG-binding domain of the methylation dependent transcriptional repressor MBD1. EMBO J 18:6653-6661.

Ozturk ZN, Talame V, Deyholos M, Michalowski CB, Galbraith DW, Gozukirmizi N, Tuberosa R and Bohnert HJ. (2002) Monitoring large-scale changes in transcript abundance in drought- and salt-stressed barley. Plant Mol. Biol. 48:551-573.

Schulte B, Li HH, Fisch-Kohl C, Purmann C and Francke U. 2007. DLX5 and DLX6 expression is biallelic and not modulated by MeCP2 deficiency. Am. J. Hum. Genet. 81:492-506.

Shinozaki K and Yamaguchi-Shinozaki K. 2007. Gene networks involved in drought stress response and tolerance. J. Exp. Bot. 58:221-227.

Springer NM and Kaeppler SM. 2005. Evolutionary divergence of monocot and dicot methyl-CpG-binding domain proteins. Plant Physiol. 138:92-104.

Tariq M and Paszkowski J. 2004. Review: DNA and histone methylation in plants. TRENDS in Genetics 20: 244-251.

Xiong LM, Schumaker KS, Zhu JK. 2002. Cell signal during cold, drought, and salt stress. Plant Cell. S165-S183, suppl.

Wakefield RID, Smith BO, Nan X, Free A, Sotieriou A, Uhrin D, Bird AP and Barlow PN. 1999. The solution structure of the domain from MeCP2 that binds to methylated DNA. J Mol Biol 291:1055-1065.

Yoda H, Ogawa M, Yamaguchi Y, Koizumi N and Sano H. 2002. Identification of early responsive genes associated with the hypersensitive response to tobacco mosaic virus and characterization of a WRKY-type transcription factor in tobacco plants. Mol. Genet. Genomics 267:154-161.

Zheng J, Zhao JF, Tao YZ, Wang JH, Liu YJ, Fu JJ, Jin Y, Gao P, Zhang JP, Bai YF and Wang GY. 2004. Isolation and analysis of water stress induced genes in maize seedlings by subtractive PCR and cDNA macroarray. Plant Mol. Biol. 55:807-823.