Putative Calmodulin-Binding R2R3-MYB Transcription Factors in Rice (Oryza sativa L.)

THANIN CHANTARACHOT1, TEERAPONG BUABOOCHA1,2, HONGYA GU3 & SUPACHITRA CHADCHAWAN1,*

1 Environment and Plant Physiology Research Unit, Department of Botany, Faculty of Science, Chulalongkorn University, Bangkok 10330, Thailand
2 Department of Biochemistry, Faculty of Science, Chulalongkorn University, Bangkok 10330, Thailand
3 College of Life Sciences, Peking University, Beijing 100871, China

ABSTRACT. Calmodulins (CaMs) are calcium-binding proteins playing vital roles in signal transduction regulating many aspects of plant development as well as responses to environmental stimuli by interacting with various classes of their binding partners including transcription factors (TFs). An increasing number of CaM-binding TFs have been identified and characterized in Arabidopsis. However, a limited number of CaM-targeted TFs have been reported in rice. This study aimed to investigate the existence of the interaction between CaMs and MYB TFs in rice, based on the knowledge from Arabidopsis. Putative CaM-binding R2R3-MYB TFs were obtained from rice genome by functional searching, using an Arabidopsis known CaM-binding TF as a query. A phylogenetic tree was constructed to confirm the relationship between the putative proteins and those CaM-binding MYB TFs from Arabidopsis. The predicted CaM-binding motifs were also determined. However, a test of interaction between four of the putative proteins and OsCaM1-1 by yeast two-hybrid assay showed the absence of protein interaction. More CaM and CaM-like proteins will be tested for the interaction in the future, using other plant-based approaches.

KEYWORDS: Calmodulin, MYB transcription factor, protein-protein interaction, rice

INTRODUCTION

Calmodulin (CaM), one of the eukaryotic Ca2+-sensor proteins, plays important roles in a wide variety of physiological processes ranging from the control of growth and development to plant responses to environmental stimuli. Typical CaMs contain four EF-hands, helix-loop-helix structural motifs, responsible for Ca2+ binding (Snedden & Fromm, 2001). Unlike calcium dependent protein kinases (CDPKs), CaMs lack their own catalytic activities or effector domains; thereby relaying particular signals to their target effectors by means of direct interaction.

Accumulating evidence has disclosed a diverse set of CaM-target proteins associated with almost all biological processes including
metabolism, ion transport, cytoskeleton organization, disease resistance, stress tolerance, and gene regulation (Yang & Poovaiah, 2003). The mode of action of CaMs in mediating plant response to developmental cues or environmental stimuli can be categorized into two distinct pathways. First, the Ca\(^{2+}\)/CaM complex binds to and modulates activities of its functional target proteins that directly participate in cellular responses to particular signals. The other way is that the Ca\(^{2+}\)/CaM complex indirectly activates cellular responses by regulating the expression of genes encoding downstream effectors, through a specific class of CaM-binding transcription factors (Kim et al., 2009).

In plant species examined to date, a large number of TFs have been reported to interact with CaMs. These CaM-binding TFs belong to different families including CaM-binding transcription activator (CAMTA), AtGT-2 (GT-element-binding protein/ GTL), MYB TF, zinc-finger-type WRKY TF, NAC domain-containing CBNAC/NT9 TF, bZIP (TGA) TF, Scarecrow/Scarecrow-like TF, MADS-box (AGL) protein, no apical meristem (NAM) TF, auxin-responsive protein (SAUR_B) and AUX/IAA TF families (Park et al., 2005; Yoo et al., 2005; Finkler et al., 2007; Kim et al., 2007; Popescu et al., 2007; Galon et al., 2010).

*Arabidopsis* MYB2 (AtMYB2), a member of R2R3-type of MYB protein family, is one of the CaM-binding TFs that have been well characterized. In past decades, its roles in such abiotic stress responses and tolerance as dehydration, salinity, and hypoxia as well as in ABA induction have been revealed in great details (Urao et al., 1993; Abe et al., 1997; Hoeren et al., 1998). Currently, the mechanisms of Ca\(^{2+}\)/CaM-mediated transcriptional regulation of stress-responsive genes in which AtMYB2 takes part have also been proposed (Galon et al., 2010). Yoo and colleagues (2005) demonstrated that GmCaM1 and GmCaM4, two divergent salt-stress-responsive CaM isoforms in soybean, were able to bind the 1-5-10 CaM-binding motif of AtMYB2 in a Ca\(^{2+}\)-dependent manner. Moreover, the two CaM isoforms differentially regulated the DNA-binding activity of AtMYB2: GmCaM4 enhanced MYB2 target gene activation (*P5CS1*, *ADH1*, and *RD22*), whereas GmCaM1 inhibited this activity. Apart from AtMYB2, other R2R3-MYB TFs were reported to interact with CaMs as well. For example, *Arabidopsis* MYB62 and MYB78 interacted with both GmCaM1 and GmCaM4 *in vitro* (Yoo et al., 2005). In addition, AtMYB14, AtMYB70 and TRF-like7 proteins were also identified as CaM-binding TFs by an *Arabidopsis* protein microarray assay with three CaMs (AtCaM1, AtCaM6, and AtCaM7) as probes (Popescu et al., 2007).

While CaM-MYB interaction in *Arabidopsis* is being progressively discovered and characterized, the existence of this kind of protein interaction in rice, which is a model plant for monocots and a staple crop worldwide, remains unstudied; calling for plant biologists to uncover Ca\(^{2+}\)/CaM-regulated of MYB-mediated gene expression.
There are five loci of Cam genes in the rice genome, namely OsCam1-1, OsCam1-2, OsCam1-3, OsCam2, and OsCam3; three of which (OsCam1-1, OsCam1-2, and OsCam1-3) encode identical proteins (Boonburapong & Buabooha, 2007). Biological functions of these proteins are being unraveled, especially for OsCam1-1. Expression of OsCam1-1 is known to be strongly responsive to dehydration and salt stress (Phean-o-pas et al., 2005). Furthermore, the role of OsCam1-1 in salt tolerance associated with a plant hormone abscisic acid has also been disclosed (Saeng-ngam et al., 2012). However, its interacting partners or downstream elements in salt–stress signaling pathway are still mysterious and need to be identified.

Here, we combined information from Arabidopsis with bioinformatics analysis and prediction to identify potential CaM-binding R2R3-MYB TFs in rice. Yeast two-hybrid assay was carried out to test interaction of these TFs and OsCam1-1 protein.

**Materials and Methods**

**Database search, sequence alignment, and helical wheel prediction**

Functional search of the rice genome for putative CaM-binding MYB TFs was carried out by blastp program. The protein sequence of AtMYB2 (accession no. NP_182241.1) was used as a query to search against the MSU Rice Genome Annotation Database (Ouyang et al., 2007). Proteins with the highest sequence similarities (>40%) to AtMYB2, together with their nucleotide and deduced amino acid sequences as well as related information, were collected.

Amino acid sequences of the putative proteins were submitted to the Pfam protein families database (Finn et al., 2010) to confirm the presence of MYB-DNA binding domains. These sequences were also aligned with AtMYB2 for sequence comparison using ClustalW2 program (http://www.ebi.ac.uk/Tools/msa/clustalo/). An 18-amino acid region of the CaM-binding motif in putative proteins was located according to the information from AtMYB2, AtMYB62 and AtMYB78. The helical wheel projection of the predicted motif was obtained by submitting the 18-amino acid sequence region to a program on the website http://cti.itc.virginia.edu/~cmg/Demo/wheel/wheelApp.html. The putative roles in abiotic stress response of the selected MYB TFs were evaluated by their expression profiles from microarray experiments in the Rice Oligonucleotide Array Database (Jung et al., 2008).

**Phylogenetic reconstruction**

The amino acid sequences of MYB genes were obtained by a keyword search from the MSU Rice Genome Annotation Database (Ouyang et al., 2007) and submitted to the Pfam protein families database (Finn et al., 2010) for domain prediction. The MYB proteins containing two MYB DNA binding domains (R2R3) were finally collected for further analysis. The complete amino acid sequences of 125 R2R3-MYB transcription factors, which were previously identified by Stracke et al. (2001), were
downloaded from The Arabidopsis Information Resource (TAIR) (http://www.arabidopsis.org). An alignment of rice and Arabidopsis R2R3-MYB protein sequences was performed with ClustalW and a phylogenetic tree was reconstructed by the Maximum Likelihood method, using MEGA5.05 software (Tamura et al., 2011). Tree nodes were evaluated with 1,000 bootstrap replicates.

Yeast two-hybrid assay

OsCaM1-1 was expressed as bait protein by fusing with the GAL4 DNA-Binding Domain (DBD). The OsCaM1-1 coding region was amplified by polymerase chain reaction (PCR) with primers and condition as indicated in Table 1 using the KOD-Plus-Neo DNA polymerase (TOYOBO, Japan) and cloned into pENTR/D-TOPO vector using the pENTR Directional TOPO Cloning Kit (Invitrogen, USA), resulting in an entry clone. The coding region in the entry clone was subsequently transferred to pDEST32 by means of recombination using Gateway LR Clonase (Invitrogen, USA).

Coding sequences of the putative CaM-binding R2R3-MYB TFs were each fused to the GAL4 Activation Domain (AD) in pDEST22 to serve as preys. Coding regions of the putative TFs were amplified from their cDNA clones by PCR using the KOD-Plus-Neo DNA polymerase (TOYOBO, Japan) with specific primers and conditions indicated in Table 1 and cloned into pENTR/D-TOPO vector using the pENTR Directional TOPO Cloning Kit (Invitrogen, USA), then were subsequently transferred to pDEST22 destination vector using Gateway LR clonase enzyme (Invitrogen, USA).

Bait and prey constructs were co-transformed into yeast Saccharomyces cerevisiae MaV203 strain by lithium acetate method (Ito et al., 1983). Yeast cells co-transformed with pEXT32-Krev1 and pEXT22-RalGST-wt and with pEXT32-Krev1 and pEXT22-RalGSD-m1 vectors were used as strong and weak interaction controls, respectively. Yeast cells co-transformed with pEXT32-Krev1 and pEXT22-RalGSD-m2 vectors served as a negative control (Invitrogen, USA).

Co-transformed cells were selected on yeast synthetic dropout (SD) medium lacking Leu and Trp (SD-Leu/-Trp) (Clontech, USA). To analyze the interaction, single colonies were transferred to SD medium lacking Leu, Trp and His (SD-Leu/-Trp/-His) (Clontech, USA) supplemented with various concentrations of 3-Amino-1, 2, 4-Triazole (3AT) and to SD medium lacking Leu, Trp, and Ura (SD-Leu/-Trp/-Ura) (Clontech, USA). β-galactosidase activity, resulting from the protein interaction, was also determined by X-gal assay following standard protocol (Invitrogen, USA). Two independent experiments were conducted.

RESULTS AND DISCUSSION

Identification of putative CaM-binding MYB TFs in the rice genome

To identify rice MYB TFs with potential capability to bind with CaMs, the amino acid sequence of AtMYB2 was used as a query
to search against the rice genome. As a result, 9 loci encoding putative MYB proteins with the highest sequence similarity to AtMYB2 (E value from $4.8 \times 10^{-62}$ to $1.1 \times 10^{-41}$) were selected for further analyses. Their coding sequences were subsequently obtained and translated. As shown in Table 2, the open reading frames of these genes were between 687-1,053 bp in length, yielding the proteins of 228-350 amino acids which shared 44.8-52.4% of sequence similarities with AtMYB2. The deduced amino acids were also screened for functional domains in Pfam database. The results from the Pfam searches established all of the proteins as R2R3-MYB TFs as they contained the same two MYB DNA-binding domains, R2 and R3, as AtMYB2.

To verify the existence and function of these MYB genes, we searched for their corresponding EST and cDNA clones. As expected, each of the selected genes has reported EST or cDNA clones, confirming its in vivo gene expression.

Based on the conservation of the DNA-binding domain and of amino acid motifs in the carboxy-terminal domain, 125 R2R3-MYB proteins in Arabidopsis have been divided into 25 functional subgroups (Kranz et al., 1998; Stracke et al., 2001). Interestingly, three of the reported CaM-binding R2R3-MYB TFs (AtMYB2, AtMYB62 and AtMYB78) were clustered in the same subgroup, subgroup 20. To further confirm the homology between the identified rice R2R3-MYBs and the Arabidopsis CaM-binding TFs, a Maximum Likelihood tree for R2R3-MYB proteins from both the rice and Arabidopsis genomes was constructed. The phylogenetic tree grouped seven of the nine rice R2R3-MYB proteins in the same clade with the Arabidopsis R2R3-MYB subgroup 20 (Fig. 1A). Aside from two highly conserved 3-α-helix-MYB domains at the amino-terminus (N-terminus), these seven R2R3-MYB TFs share the same WxPRL motif with those Arabidopsis members at the carboxy-terminus (C-terminus) (Fig. 1B). On the other hand, the other two (LOC_Os01g45090.1 and LOC_Os02g40530.1) proteins did not contain any motif at their C-termini, hence not being classified into any subgroup.

It is suggested that the Arabidopsis R2R3-MYB subgroup 20 functions in stress responses. AtMYB2 regulates the ABA-mediated expression of salt and dehydration responsive genes (Abe et al., 2003), AtMYB62 is involved in the response to phosphate starvation (Devaiah et al., 2009), and AtMYB108 in both biotic and abiotic stress responses associated with jasmonate signaling pathway (Mengiste et al., 2003). The possible roles of the identified rice R2R3-MYB TFs in stress responses were also evaluated by their expression profiles from microarray experiments in the Rice Oligonucleotide Array Database (Jung et al., 2008). Six rice R2R3-MYB TFs encoding genes including LOC_Os03g20090.1, LOC_Os07g48870.1, LOC_Os01g19330.1, LOC_Os11g45740.1, LOC_Os12g37690.1, and LOC_Os02g40530.1 were up-regulated under drought and/or salt stresses (data not shown). Five of them belong to subgroup 20, supporting the possibility that they are involved in
abiotic stress response like their *Arabidopsis* counterparts.

The motif for CaM recognition of *Arabidopsis* MYB TFs has been experimentally identified in AtMYB2 (Yoo *et al.*, 2005). It stretches over 18 contiguous amino acids between the two domains (R2 and R3) of DNA-binding region with an amphiphilic characteristics, a segregation of basic and polar residues on one side with hydrophobic amino acids on the other side (Fig. 2A). This motif was defined as the 1-5-10 type for the presence of hydrophobic amino acids at positions 1 (Leu6), 5 (Val10), and 10 (Ile15), two of which (Leu6 and Ile15) were the key residues for CaM interaction. CaM-binding motifs of the rice R2R3-MYB TFs were also predicted and compared to those of *Arabidopsis* CaM-binding R2R3-MYBs. As presented by sequence comparison in Fig. 2B, most of the predicted motifs of the rice MYB TFs showed similar patterns to the 1-5-10 motif of AtMYB2.

**Table 1.** Characteristics of the rice putative CaM-binding MYB TF genes and their encoded proteins

<table>
<thead>
<tr>
<th>Locus Name</th>
<th>ORF length (bp)</th>
<th>Amino acids</th>
<th>%Similarity with At-MYB2</th>
<th>No. of MYB domains</th>
<th>Transcript evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>LOC_Os05g04210.1</td>
<td>933</td>
<td>310</td>
<td>51.3%</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>LOC_Os03g20090.1</td>
<td>990</td>
<td>329</td>
<td>50.3%</td>
<td>2</td>
<td>36</td>
</tr>
<tr>
<td>LOC_Os07g48870.1</td>
<td>903</td>
<td>300</td>
<td>50.6%</td>
<td>2</td>
<td>15</td>
</tr>
<tr>
<td>LOC_Os01g19330.1</td>
<td>882</td>
<td>293</td>
<td>51.7%</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>LOC_Os11g45740.1</td>
<td>858</td>
<td>285</td>
<td>52.4%</td>
<td>2</td>
<td>14</td>
</tr>
<tr>
<td>LOC_Os01g03720.1</td>
<td>1,053</td>
<td>350</td>
<td>44.8%</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>LOC_Os12g37690.1</td>
<td>903</td>
<td>300</td>
<td>46.2%</td>
<td>2</td>
<td>12</td>
</tr>
<tr>
<td>LOC_Os01g45090.1</td>
<td>687</td>
<td>228</td>
<td>47.9%</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>LOC_Os02g40530.1</td>
<td>870</td>
<td>289</td>
<td>47.1%</td>
<td>2</td>
<td>10</td>
</tr>
</tbody>
</table>

**Table 2.** cDNA clones, primers, and PCR condition used in PCR-cloning for yeast two-hybrid experiment

<table>
<thead>
<tr>
<th>Gene</th>
<th>eDNA Clone</th>
<th>Forward</th>
<th>Reverse</th>
<th>PCR conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>LOC_Os03g20090.1</td>
<td>AK120551</td>
<td>5’CACCATGGGACATGGGCGACAGGAGAGGG3’</td>
<td>5’ACTTTTCACACCGCGCCTG3’</td>
<td>94c, 2 min; 35 cycles of 98c,10sec-62c,30sec-68c,30sec; 68c,10min</td>
</tr>
<tr>
<td>LOC_Os07g48870.1</td>
<td>AK103455</td>
<td>5’CACCATGATGATGGCGGCGAGAGGT3’</td>
<td>5’TTCACAAGGACCAA ACTTTGCGT3’</td>
<td>94c, 2 min; 35 cycles of 98c,10sec-58c,30sec-68c,30sec; 68c,10min</td>
</tr>
<tr>
<td>LOC_Os11g45740.1</td>
<td>AK069082</td>
<td>5’CACCATGGAGATGGCTCGAGAGAAGGATAG3’</td>
<td>5’TGCCCTTCCTTGATGTTCTCTA3’</td>
<td>94c, 2 min; 35 cycles of 98c,10sec-58c,30sec-68c,30sec; 68c,10min</td>
</tr>
<tr>
<td>LOC_Os02g40530.1</td>
<td>AK099283</td>
<td>5’CACCATGGGACATGGGCGACAGGAGAGGG3’</td>
<td>5’ATAGCTAGTAG AAAGCGCGG3’</td>
<td>94c, 2 min; 35 cycles of 98c,10sec-58c,30sec-68c,30sec; 68c,10min</td>
</tr>
<tr>
<td>OsCam1-1</td>
<td>AU081299</td>
<td>5’CACCATGGGACATGGGCGACAGGAGAGGG3’</td>
<td>5’TCACCTGGCCATCATGACCTTG3’</td>
<td>94c, 2 min; 35 cycles of 98c,10sec-58c,30sec-68c,20sec; 68c,10min</td>
</tr>
</tbody>
</table>
**Figure 1** Phylogenetic analysis and amino acid sequence alignment of the putative CaM-binding MYB TFs from rice and the R2R3-MYB proteins from *Arabidopsis*. A. A Maximum Likelihood tree representing the relationship between the 9 putative rice CaM-binding MYB TFs and *Arabidopsis* subgroups 19 and 20 of R2R3-MYB TFs. Bootstrap values lower than 50 were omitted. B. An amino acid sequence alignment of the 9 putative rice CaM-binding MYB TFs and the members of *Arabidopsis* R2R3-MYB TF subgroup 20. The highly conserved MYB-DNA binding domain (R2R3) is presented. Green box indicates the carboxy-terminal WxPRL motif.
Yeast two-hybrid assay

Yeast two-hybrid system was employed to test if the putative rice CaM-binding R2R3-MYB TFs identified above could interact with the salt-stress responsive OsCaM1-1. Based on sequence similarity, availability of full-length cDNA clones, and a propensity to form a CaM-binding motif, we selected four of the rice R2R3-MYB TFs for this experiment. Each of the four R2R3-MYB proteins was individually expressed as prey, a GAL4-AD fusion protein, whereas OsCaM1-1 fused to GAL4-DBD served as bait. Yeast transformants capable of growing on SD-Leu-Trp medium were analyzed for protein interaction on two selective media (SD-Leu/-Trp/-Ura and SD-Leu/-Trp/-His+3AT) and by X-gal assay. Unexpectedly as shown in Fig. 3, the yeast harboring each combination of bait and prey vectors neither grew on SD-Leu/-Trp/-Ura or SD-Leu/-Trp/-His+3AT media nor displayed β-galactosidase activity, whereas the positive controls did, indicating the absence of protein interaction.

Two hypotheses were raised to account for the lack of interaction with OsCaM1-1 by the four selected rice R2R3-MYB TFs in yeast two-hybrid assays. First, the R2R3-MYB TFs might not serve as the interacting-partners of OsCaM1-1. This hypothesis is supported by a result from Perpescu and colleagues’ experiment in Arabidopsis (2007) indicating that CaM-target proteins were specific to only one or a few CaM/CMLs. There are many CaMs and CaM-like proteins (CMLs) in the rice genome, each of which may specifically function through different targets. The binding partner of R2R3-MYB TFs in rice might not be OsCaM1-1, but rather another CaM or CML. Following this hypothesis, therefore, the CaM-R2R3-MYB TF interaction should be investigated for a larger number of both CaM/CML and R2R3-MYB proteins.

The other hypothesis is that R2R3-MYB TFs and OsCaM1-1 do interact in rice, but this interaction could not be detected by the yeast two-hybrid assay. It has been claimed by Reddy et al. (2011) that the yeast-two hybrid assay might not be an ideal system for CaM-protein interaction because the interaction of CaM with its targets is normally dependent on Ca\textsuperscript{2+}. Hence, the unexpected result may be due to the different cellular Ca\textsuperscript{2+} signatures between plant and yeast cells. It is difficult to manipulate calcium concentration in yeast cells during this process. Moreover, it was possible that plant proteins produced in yeast might not be as functionally active as those in plants. An assay of 96 Arabidopsis protein kinases produced by yeast revealed that only 3-5% of the proteins were active, while 82% were active when expressed in planta (Perpescu et al., 2007). Thereby to address this hypothesis, plant-based expression systems such as co-immunoprecipitation (CoIP) or Bimolecular Fluorescence Complementation (BiFC), which have been successfully used for studies of CaM interaction (Takahashi et al., 2011), are prospective approaches for verification of the protein interaction.
**Figure 2** The CaM-binding motif of AtMYB2 and a comparison of the putative motifs from rice R2R3-MYB TFs with *Arabidopsis* CaM-binding MYB TFs. A. a helical wheel projection of the CaM-binding motif of AtMYB2. B. an alignment of putative CaM-binding motifs of the rice R2R3-MYB TFs with *Arabidopsis* CaM-binding MYB TFs. The most hydrophobic residues, according to the Kyte-Doolittle hydrophobicity scale (Kyte & Doolittle, 1982), are colored red and the most hydrophilic ones are colored blue. Asterisks indicate the key residues for CaM-interaction.

**Figure 3** Yeast two-hybrid assay for the interaction between OsCaM1-1 and the putative CaM-binding R2R3-MYB TFs. The growth of yeast cells on SD-Leu/-Trp/-Ura and SD-Leu/-Trp/-His + 50 mM 3AT was determined after 4 and 3 days, respectively. Five independent clones of the yeast transformants are presented.
CONCLUSION

In this study we have identified a number of putative rice CaM-binding R2R3-MYB TFs based on gene expression and structural homology with their known counterparts in Arabidopsis. Initial tests of the interaction between these TFs with a salt-stress-associated OsCaM1-1 by yeast two-hybrid system showed negative results. Therefore, a larger number of rice CaMs/CMLs and R2R3-MYB TFs will need to be further determined for such interactions using other plant-based approaches.

ACKNOWLEDGEMENTS

The research was supported by the Higher Education Research Promotion and National Research University Project of Thailand, Office of the Higher Education Commission (FW656B) and Faculty of Science, Chulalongkorn University (AIBI). The facilities were supported by the Thai Government Stimulus Package 2 (TKK2555), under PERFECTA. We greatly thank Prof. Hongya Gu and Prof. Li-Jia Qu at the National Laboratory for Protein Engineering and Plant Genetic Engineering, College of Life Sciences, Peking University, P.R. China for kindly providing the facilities for Y2H experiment. TC was supported by the Graduate School, Chulalongkorn University and the Development and Promotion of Science and Technology Talents Project of Thailand.

REFERENCES

Abe, H., Ito, T., Seko, M., Shinozaki, K., & Yamaguchi-Shinozaki, K. 2003. Arabidopsis AtMYC2 (bHLH) and AtMYB2 (MYB) function as transcriptional activators in abscisic acid signaling. The Plant Cell 15: 63-78.


