The C-terminal Sequence of LMADS1 Is Essential for the Formation of Homodimers for B Function Proteins*

Received for publication, October 23, 2003, and in revised form, December 15, 2003
Published, JBC Papers in Press, December 15, 2003, DOI 10.1074/jbc.M311646200

Tsai-Yu Tseng, Hsin-Ching Liu, and Chang-Hsien Yang‡

From the Graduate Institute of Biotechnology, National Chung Hsing University, Taichung, Taiwan 40227, Republic of China

LMADS1, a lily (Lilium longiflorum) AP3 orthologue, contains the complete consensus sequence of the paleoAP3 (YGSHDLRLA) and PI-derived (YEFRVQPS-QPNLII) motifs in the C-terminal region of the protein. Interestingly, through yeast two-hybrid analysis, LMADS1 was found to be capable of forming homodimers. These results indicated that LMADS1 represents an ancestral form of the B function protein, which retains the ability to form homodimers in regulating petal and stamen development in lily. To explore the involvement of the conserved motifs in the C-terminal region of LMADS1 in forming homodimers, truncated forms of LMADS1 were generated, and their ability to form homodimers was analyzed using yeast two-hybrid and electrophoretic mobility shift assay. The ability of LMADS1 to form homodimers decreased once the C-terminal paleoAP3 motif was deleted. When both paleoAP3 and PI-derived motifs were deleted, the ability of LMADS1 to form homodimers was completely abolished. This result indicated that although the paleoAP3 motif promotes the formation of LMADS1 homodimers, the PI-derived motif is essential. Deletion analysis indicated that two amino acids, RV, of the 5 final amino acids, YEFRV, in the PI-derived motif are essential for the formation of homodimers. Further, point mutation analysis indicated that amino acid Val was absolutely necessary, whereas residue Arg played a less important role in the formation of homodimers. Furthermore, Arabidopsis AP3 was able to form homodimers once its C-terminal paleoAP3 motif was deleted. When both paleoAP3 and PI-derived motifs were deleted, the ability of LMADS1 to form homodimers was completely abolished. This result indicated that the C-terminal region of LMADS1 is responsible and essential for homodimer formation of the ancestral form of the B function protein.

MADS box genes have been thought to play central roles in flower development (1–4). The most representative MADS box genes in the B function group are AP3 and PI, which play major roles in specifying petal and stamen development in Arabidopsis (5–8). In Arabidopsis, ap3 and pi mutants have identical phenotypes, producing the sepal structure in the second whorl. Different from other MADS box proteins, which form homodimers in regulating flower development, in various plant species, proteins in the euAP3 and paleoAP3 lineages were stable and functional in the cell, regulating petal and stamen development only in heterodimer form with a protein in the PI lineage (9, 11, 20, 30). For example, heterodimers were formed between AP3 and PI in Arabidopsis (29), DEFICIENS (DEF) and GLOBOSA (GLO) in Antirrhinum (9, 11), and OsMADS16 and OsMADS4 in rice (20), respectively. This partnership was supported by the expression pattern for AP3 and PI orthologues. They are expressed at the same time and in the same areas, in petals and stamens, during flower development (11, 31, 32). Furthermore, expression of AP3 and PI orthologues is maintained by an autoregulatory circuit (6, 7, 33, 34). AP3/PI or DEF/GLO heterodimers autoregulated their own expression by binding to the specific CAARG sequence elements in their promoter regions. Therefore, it is interesting to explore the origin of this obligate nature of heterodimerization for AP3/PI during evolution. Because there was only one B gene before duplication, it is postulated that the ancestral B protein should function as a homodimer in regulating gene expression (17, 24). The ability of B proteins to form homodimers was gradually lost and replaced by a preference to form heterodimers with other B proteins. This assumption was supported by a recent finding that the LMADS1 of the monocot lily (Lilium longiflorum) in paleoAP3 lineage was able to form homodimers in

* This work was supported by grants (to C.-H. Y.) from the Council of Agriculture and National Science Council Grants 91AS-3.1.1-FD-72 and NSC92-2311-B-005-022. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.
‡ To whom correspondence should be addressed. Tel.: 886-4-2285-3126; Fax: 886-4-2285-3126; E-mail: chyang@dragon.nchu.edu.tw.

This paper is available on line at http://www.jbc.org
C-terminal Sequence of B Protein Essential for Homodimers

regulating flower development (24). An orchid (*Oncidium Gower Ramsey*) AP3-like protein OMADS3 is also able to strongly form homodimers (28). A similar result was observed for the lily (*Lilium regale*) PI orthologue and an ancient B protein from the gymnosperm *Gnetum gnetum*, which have the ability to bind DNA as a homodimer (17).

In addition to forming homodimers, LMADS1 was also able to form heterodimers with *Arabidopsis* PI and produced dominant negative AP3-like mutations in transgenic *Arabidopsis* plants (29). Complete consensus sequence by LMADS1 to the paleoAP3 and PI-derived motifs in the C-terminal region of the protein is an indication of its ancestry to the B function gene (24). Thus, LMADS1 has become an excellent candidate for studying the transition from homodimerization to heterodimerization of B protein functions during evolution. To explore the possible involvement of the conserved paleoAP3 and PI-derived motifs in the C-terminal region of LMADS1 in forming homodimers, various deletions and point mutations in the two motifs within the truncated LMADS1 were executed, and the subsequent ability of truncated LMADS1 to form homodimers was analyzed. The result indicated that these two motifs were absolutely essential for LMADS1 to form homodimers. This conclusion was further supported by the result that *Arabidopsis* AP3 acquired the ability to form homodimers once its C-terminal region was replaced by that of LMADS1.

**EXPERIMENTAL PROCEDURES**

**Generation of Various Truncated cDNAs for LMADS1**—The various truncated cDNA for LMADS1 was amplified by PCR using plasmid pGBK7-LMADS1 (24) as templates (see Fig. 1A). Specific 5′ primer used in PCR was L1-5′-TTCCTGGAATTTCCGAGAAGC-3′, and the specific 3′ primers for each construct are as follows: L1-C1, delete paleoAP3 motif (LMADS1-219; 5′-GGATCCCTATCGAGGGAAGAGCATG-3′); and L1-C2, delete paleoAP3 and PI-derived motifs (LMADS1-203; 5′-GGATCCCTAACGATGGAGACACCCG-3′). L1-C1, delete paleoAP3 and the last 2 amino acids of PI-derived motifs (LMADS1-214; 5′-GGATCCTTAGTTGCGGCTCGG-3′); L1-C2-2, delete paleoAP3 and the last 4 amino acids of PI-derived motifs (LMADS1-212; 5′-GGATCCTTACGTCGTTGAGCAC-3′); L1-C2-3, delete paleoAP3 and the last 6 amino acids of PI-derived motifs (LMADS1-210; 5′-GGATCCTTACGTTGGACTCCGAGA-3′); L1-C2-4, delete paleoAP3 and the last 8 amino acids of PI-derived motifs (LMADS1-208; 5′-GGATCCTTACGGTCAACCATG-3′); and L1-C2-5, delete paleoAP3 and the last 10 amino acids of PI-derived motifs (LMADS1-206; 5′-GGATCCTTAGTTGTTGCGGCTCGG-3′). Specific 5′ and 3′ primes contained either the generated EcoRI recognition site (5′-GAAATTC-3′, underlined) or BamHI recognition site (5′-GGATCC-3′, underlined) to facilitate the cloning of cDNAs. For these constructs, all of the 3′ primes contained a TAG or TAA stop codon (bold type) created right before the BamHI recognition site. PCR fragments were digested with appropriate enzymes and ligated into either plasmid pGBK7 (binding domain vector) or PGAD7 (activation domain vector) for yeast two-hybrid analysis.

**Generation of cDNAs with Point Mutation and Amino Acid Substitution for LMADS1**—The various cDNA-containing point mutations for LMADS1 were amplified by PCR using plasmid pGBK7-LMADS1 (24) as templates (see Fig. 1A). The specific 5′ primer used in PCR was L1-5′-TTCCTGGAATTTCCGAGAAGC-3′ and the specific 3′ primes for each construct are as follows: L1-R207, with Arg substituted by Ala at position 207 (LMADS1-R207A; 5′-GGATCCTTAGTTGCGGAGACACCCG-3′); L1-L208, with Val substituted by Ala at position 208 (LMADS1-L208A; 5′-GGATCCTTACGTCGTTGAGCAC-3′); L1-V208, with Val substituted by Arg at position 208 (LMADS1-V208R; 5′-GGATCCTTACGTCGTTGAGCAC-3′); and L1-L208R, with Arg substituted by Ala at position 208, and Val substituted by Ala at position 208 (LMADS1-L208R208A; 5′-GGATCCTTACGTCGTTGAGCAC-3′). Specific 5′ and 3′ primes contained either the generated EcoRI recognition site (5′-GAAATTC-3′, underlined) or BamHI recognition site (5′-GGATCC-3′, underlined) to facilitate the cloning of cDNAs. For these constructs, all of the 3′ primes contained a TAA stop codon created right before the BamHI recognition site. The DNA sequences mutated are in bold type and underlined. The PCR fragments were digested with appropriate enzymes and ligated into either plasmid pGBK7 (binding domain vector) or PGAD7 (activation domain vector) for yeast two-hybrid analysis.

**Generation of Chimeric AP3 and PI cDNAs Containing C-terminal Portion of LMADS1**—PCR was used to generate the chimeric cDNAs, AP3-LIC and PI-LIC in which the C-terminal region of AP3 or PI were replaced by that of LMADS1 (see Fig. 1B). For generation of AP3-LIC, two cDNAs (AP3-C and P3-LIC) were first produced by PCR. Primers P3-EI (5′-GGATCCTAGTTGCGGAGACACCCG-3′) and P3-EI (5′-GGATCCTTACGTCGTTGAGCAC-3′) were used in PCR using pGBK7-PI-LIC (24) as templates to produce cDNA PI-LIC that encoded 154 amino acids of the C-terminal domain of the LMADS1 protein. For generation of AP3-LIC, two cDNAs (AP3-C and P3-LIC) were first produced by PCR. Primers P3-EI (5′-GGATCCTAGTTGCGGAGACACCCG-3′) and P3-EI (5′-GGATCCTTACGTCGTTGAGCAC-3′) were used in PCR using pGBK7-PI-LIC (24) as templates to produce cDNA PI-LIC that encoded 154 amino acids of the C-terminal domain of the LMADS1 protein. A second round of PCR using I-EI and the LI-BI as primers and cDNAs AP3-C and PI-LIC as templates was performed to generate chimeric cDNA AP3-LIC that produced a protein fused with the N-terminal I and K domains of AP3 and the C-terminal domain of LMADS1 (see Fig. 1B).

A similar strategy was used to generate PI-LIC. For generation of PI-LIC, two cDNAs (PI-C and I-LIC) were first produced by PCR. Primers I-EI (5′-GGATCCTAGTTGCGGAGACACCCG-3′) and I-C (5′-GTATTTGTTCTCTCCTGGCCCATCATCTT-3′) were used in PCR using pGBK7-PI (24) as templates to produce cDNA PI-C that encoded 154 amino acids of the N-terminal portion (M1, and K domains) of PI protein. Primers I-LIC (5′-AGATGCGGAAAGACGCAACCAAGAAC-3′) and I-LIC (5′-GGATCCTTAGTTGCGGAGACACCCG-3′) were used in PCR using pGBK7-PI-LIC (24) as templates to produce cDNA I-LIC that encoded 71 amino acids of the C-terminal domain of the LMADS1 protein. A second round of PCR using I-EI and the L1-C2 as primers and cDNAs PI-C and L1-C2 as templates was performed to generate chimeric cDNA PI-LIC that produced a protein fused with the N-terminal I and C-terminal domain of LMADS1.

For generation of cDNAs truncated with paleoAP3 or PI-derived motifs in the 3′ end of AP3-LIC, combinations of specific primers were used in PCR. The 5′ specific primer for deletion of paleoAP3 motifs is P3-LIC (5′-GGATCCTAGTTGCGGAGACACCCG-3′); the 3′ specific primer for deletion of paleoAP3 and PI-derived motifs (AP3-LIC-219) is L1-C2. The 3′ specific primer for deletion of paleoAP3 and PI-derived motifs (AP3-LIC-203) is L1-C2. Specific 5′ and 3′ primes contained the generated EcoRI or BamHI recognition site to facilitate the cloning of cDNAs. PCR fragments were digested with appropriate enzymes and ligated into either plasmid pGBK7 (binding domain vector) or PGAD7 (activation domain vector) for yeast two-hybrid analysis.

**Yeast Two-hybrid Analysis—**Yeast two-hybrid analysis was performed using the MATCHMAKER yeast two-hybrid system 3 (Clontech). In this system, strain Y187 was used for transformation, and lacZ was used as the reporter gene. Yeast transformation was performed by using the lithium acetate method (24, 35). The transformants co-transformed with the corresponding domain and activation domain plasmid were selected on selection medium lacking tryptophan and leucine (Trp-Leu) according to the manufacturer’s instructions. For the analysis of β-galactosidase activity, positive transformants grown on selection medium were further grown and suspended into Z buffer (100 mm NaPO4, 10 mm KCl, 1 mm MgSO4, 50 mm β-mercapto-ethanol, pH 7.0) containing o-nitrophenyl-β-d-galactopyranoside (4 mg/ml in Z buffer) as a substrate. β-Galactosidase activity was calculated according to Miller (36).

**Electrophoretic Mobility Shift Assay (EMSA) —**The sequence of the CArG1 oligonucleotide (GCAACTCTTTTCCTTTTTAGGTCGACAT) was used in the EMSA with the Antirrhinum DEF promoter (17). The oligonucleotides were labeled with 32P at their 5′ ends and gel purified prior to EMSA. The GST fusion proteins were all generated from recombinant pGEX-4T plasmid (Amersham Biosciences) transformed into *Escherichia coli* BL21 cells. The various truncated cDNA for LMADS1 were used to generate GST fusion proteins which were amplified by PCR using plasmid pGBK7-LMADS1 (24) as templates (see Fig. 1A). The specific primer used in PCR was L1-5′-GGATCCTAGTTGCGGAGACACCCG-3′ and the 3′ specific primers for each construct are as follows: L1-SF1 (5′-GGATCCTAGTTGCGGAGACACCCG-3′); L1-SF2, delete paleoAP3 motif (LMADS1-219E; 5′-GGATCCTAGTTGCGGAGACACCCG-3′); L1-E1C, delete paleoAP3 and PI-derived motifs (LMADS1-203E, 203E).
RESULTS

paleoAP3 and PI-derived Motifs Are Required for The Formation of LMADS1 Homodimers—To explore the possible involvement of the paleoAP3 and PI-derived motifs in the C terminus of LMADS1 in the formation of homodimers, constructs containing cDNAs with deletions of the paleoAP3 motif (LMADS1–219) or both the paleoAP3 and PI-derived motifs (LMADS1–203) (Fig. 1A) were transformed into yeast followed by two-hybrid analysis. As shown in Figs. 1A and 2, the ability of LMADS1–219 to form homodimers was decreased to approximately 75% of that for full-length LMADS1. This suggested a positive role for the paleoAP3 motif in the formation of homodimers in LMADS1. Interestingly, the ability of LMADS1 to form homodimers was completely abolished once both the paleoAP3 and PI-derived motifs were deleted as seen in LMADS1–203 (Figs. 1A and 2). This result indicated that the PI-derived motif was essential for the formation of homodimers in LMADS1 once the paleoAP3 motif was absent. As controls, neither PI nor AP3 formed homodimers (Fig. 2).

To further confirm the result obtained from yeast two-hybrid analysis, EMSA, a technique established for the investigation of dimerization and DNA binding of MIKC-type MADS domain proteins (9, 30, 17), was employed. A stretch of DNA sequence (probe CarG1 from the Antirrhinum DEF promoter, including CarG box (CC(A/T)GG), which has been thought to be the region bound by the MADS proteins in regulating gene expression (34, 37c–39), was used to investigate the sequence-specific DNA binding of various forms of LMADS1 protein homodimers to this sequence. As shown in Fig. 3, LMADS1 proteins bound efficiently to probe CarG1. The binding of LMADS1 proteins to CarG1 was confirmed by the supershift assay in which the signal of protein–DNA complexes was shifted (from position of arrow 1 to that of arrow 2 in Fig. 3) on the GST antiserum were added (Fig. 3). When the binding of CarG1 to LMADS1–219E (with paleoAP3 motif deletion) was analyzed, the signal of protein–DNA complexes was also observed and seen to shift in the supershift assay (Fig. 3). The ability of LMADS1–219E homodimers to bind CarG1 was clearly weaker than that for full-length LMADS1 (Fig. 3). This indicated a positive role for the paleoAP3 motif in the formation of homodimers in LMADS1 as was seen in yeast two-hybrid analysis (Figs. 1A and 2).

Interestingly, the ability of LMADS1–203E (with both paleoAP3 and PI-derived motifs deletions) proteins to bind to CarG1 was completely abolished as seen for LMADS1–203 in yeast two-hybrid analysis (Figs. 1A and 2). This result confirmed that the PI-derived motif was essential for the formation of LMADS1 homodimers once the paleoAP3 motif was absent.

Amino Acids in the PI-derived Motif Play Different Roles in the Formation of LMADS1 Homodimers—To further explore the role of the PI-derived motif in the formation of LMADS1 homodimers, constructs containing cDNAs with a series of deletions of amino acids in the PI-derived motif (LMADS1–214 to LMADS1–206) (Fig. 1A) were transformed into yeast followed by two-hybrid analysis. As shown in Fig. 4, the ability of LMADS1–214 (deletion of 2 amino acids, LH, in the PI-derived motif) to form homodimers decreased to approximately 40% of that for full-length LMADS1 (Fig. 1A). This suggested a positive role for amino acids LH in the PI-derived motif in the formation of LMADS1 homodimers (Fig. 1A). Interestingly, when 2 more amino acids (PN) were deleted in the C terminus of the PI-derived motif (LMADS1–212), the ability of LMADS1–212 to form homodimers increased approximately 2-fold of that for LMADS1–214 and was similar to that for LMADS1–219, which contained the entire PI-derived motif (Figs. 1A and 4). The ability to form homodimers for LMADS1–210 (6 amino acids, SQPNLH, in PI-derived motif deleted) was even slightly higher than that observed in LMADS1–212 (Figs. 1A and 4). This suggested a possibly negative role for amino acids SQPN in the PI-derived motif during formation of LMADS1 homodimers (Fig. 1A). When 8 amino acids (QPSPQNLH) were deleted in the C terminus of the PI-derived motif (LMADS1–208), the ability to form homodimers decreased approximately 40% of that for full-length LMADS1 and was similar to that for LMADS1–214 (Figs. 1A and 4). This suggested a positive role for amino acids QP in the PI-derived motif during formation of LMADS1 homodimers (Fig. 1A). The ability of LMADS1 to form homodimers was completely abolished once 10 amino acids (RQPVSPQNLH) in the PI-derived motif were deleted in LMADS1–206 (leaving only three amino acids YEF remaining) and was similar to that for LMADS1–203 (Figs. 1A and 4). This result indicated that amino acids RV in the PI-derived motif are particularly essential for the formation of homodimers for LMADS1 (Fig. 1A).

To further examine the role of amino acids RV in the formation of LMADS1–208 homodimers, amino acid substitution for these two amino acid residues through site-specific mutagenesis was employed. When the arginine residue at position 207 was substituted by alanine in LMADS1–R207A, the ability to form homodimers decreased to approximately 50% of that for LMADS1–208 (Figs. 1A and 4). This suggested that amino acids YEF remaining was similar to that for LMADS1–203 (Figs. 1A and 4). This result indicated that amino acids RV in the PI-derived motif are particularly essential for the formation of homodimers for LMADS1 (Fig. 1A).
**FIG. 1.** Truncated and chimeric constructs used for yeast two-hybrid analyses. A, a series of deletions in the 3' end of the *LMADS1* generated *LMADS1* proteins (contained 228 amino acids) truncated with a variable C terminus in the paleoAP3 and PI-derived motifs. Constructs *LMADS1*–219 to *LMADS1*–203 encoded proteins with 8- and 25-amino acid deletions in paleoAP3 and PI-derived motifs of *LMADS1*, respectively. Constructs *LMADS1-R207A*, *LMADS1-V208A*, *LMADS1-V208R*, and *LMADS1-RV208AA* encoded proteins with various amino acid substitutions (boxed) in Arg207 and Val208 of PI-derived motifs. Construct AP3-LIC encoding a chimeric AP3 protein with the C terminus of *LMADS1*. Constructs AP3-LIC-219 and AP3-LIC-203 encoding proteins with 8- and 25-amino acid deletions in paleoAP3 and PI-derived motifs of AP3-LIC, respectively. Construct PI-LIC encoding a chimeric PI protein with the C terminus of *LMADS1*. The *first large boxed region* represents the conserved PI-derived motif sequence (YEFRVQPSQPNLH). The *second large boxed region* (YGSHDLRLA) contains sequences for the paleoAP3 motif. The *horizontal lines* above the PI-derived and paleoAP3 motifs indicate the amino acids that possibly played either positive (+) or negative (−) roles in the homodimerization of *LMADS1*. The β-galactosidase activity shown at the right indicates the relative ability of homodimerization for each truncated or chimeric protein. B, strategy used to generate AP3-LIC cDNA encoding a chimeric AP3 protein with the C terminus of *LMADS1* (see "Experimental Procedures" for detail). The *two filled boxes* in the C terminus of *LMADS1* represent the PI-derived and paleoAP3 motifs. *I*, *K*, and *C* indicate the I, K, and C domains, respectively.
form homodimers was as strong as that observed for LMDAS1. By contrast, PI-LIC was not able to form homodimers as seen for P1 (Figs. 1A and 5).

Furthermore, constructs containing AP3-LIC with paleoAP3 motif (AP3-LIC-219) deletion or deletion of both the paleoAP3 and PI-derived motifs (AP3-LIC-203) (Fig. 1A) were transformed into yeast and two-hybrid analysis performed. As shown in Figs. 1A and 5, the ability of AP3-LIC-219 to form homodimers decreased to approximately 80% of that for full-length AP3-LIC. The ability of AP3-LIC to form homodimers was completely abolished once both paleoAP3 and PI-derived motifs were deleted as seen in AP3-LIC-203 (Figs. 1A and 5). This result provides evidence to support the idea that the paleoAP3 and PI-derived motifs in LMDAS1 are essential for the formation of homodimers.

DISCUSSION

Obligate heterodimerization is a unique characteristic for B function MADS proteins in regulating petal and stamen development (9, 11, 20, 30). Discovery of the ability to form homodimers for B function proteins in monocot lily (17, 24), orchid (28), and gymnosperm (17) is extremely interesting. These results support that an ancestral B protein functioned as a homodimer in regulating gene expression (17, 24). The ability to form homodimers for B function proteins was, however, replaced by the formation of heterodimers in eudicots after gene duplications.

LMDAS1 of lily (L. longiflorum), characterized previously in our laboratory, contains both paleoAP3 and PI-derived motifs in the C-terminal region of the protein (Fig. 6A) (24). LMDAS1 can not only form homodimers but can also form heterodimers with Arabidopsis PI efficiently (24). This suggested B function gene ancestry for LMDAS1 as a possibly transitional role from homodimerization to heterodimerization. One interesting question raised in this study was, what sequence or structure specificity allowed LMDAS1 to retain the ability to form homodimers? As shown in Fig. 6A, LMDAS1 contains a PI-derived motif (YEFRVQPSQPNLH) which showed 85% (11/13) identity to the consensus PI-derived motif (FXFR-LQPSQPNLH) found in AP3 family genes (8, 20). The 11-amino acid core (EFRVQPSQPNL) in this sequence is completely identical to the core consensus sequence of the PI motif of the P1 lineage and is only 1 amino acid different from the core consensus sequence of the PI-derived motif of the paleoAP3 lineage (Fig. 6A). In addition, 100% (9/9) identity was found between the paleoAP3 motif of LMDAS1 (YSQHDLRLA) (Fig. 6A) and consensus of the paleoAP3 motif (YQXHDLRLA) found in AP3 family genes of low eudicot, magnolid dicot, and monocot species (8, 20). This high sequence conservation indicated a great possibility that these two motifs play an important role for homodimerization.

To seek evidence for the involvement of the paleoAP3 and PI-derived motifs of LMDAS1 in forming homodimers, homodimerization for deletion mutations was analyzed. Our recent results clearly indicated that the paleoAP3 motif in the C terminus of LMDAS1 is responsible for homodimerization because the ability of LMDAS1 to form homodimers decreased once this motif was deleted as seen in LMDAS1–219. Because approximately 75% of the β-galactosidase activity still remained in LMDAS1–219 in yeast two-hybrid analysis, this paleoAP3 motif was thus useful but not obligate for homodimerization of LMDAS1. Interestingly, when the PI-derived motif was further deleted as seen in LMDAS1–203, the ability of LMDAS1 to form homodimers was completely eliminated. This result indicated that the PI-derived motif was not only necessary but also required for the formation of homodimers for LMDAS1. This assumption was supported by the fact that LRGLOA was able to form homodimers although the paleoAP3 motif was absent and only the PI motif was observed in its C terminus (17). The result obtained from yeast two-hybrid analysis was further confirmed by an independent method EMSA that test for DNA-binding depending on protein dimerization. In EMSA, LMDAS1 proteins bound efficiently to probe CAR1 whereas binding ability slightly decreased for LMDAS1–219 (with paleoAP3 motif deletion) and was abolished for LMDAS1–203E (with both paleoAP3 and PI-derived motif deletions). The result also indirectly supported that the different proteins tested in yeast two-hybrid analysis were likely expressed in yeast at the same level or were equally stable. The decrease in β-galacto-
activity. By contrast, amino acid residues 211 to 214 (SQPN) seem not to be required for homodimerization since \(\beta\)-galactosidase activity was not influenced once these four amino acids were deleted. The minimum number of amino acids required for the PI-derived motif to retain its ability for homodimerization is five (YFRV). The ability for homodimerization was completely lost by deletion of amino acids \(R_{207}V_{208}\) from the PI-derived motif in LMADS-208. This suggested a mandatory role for these two amino acids, RV, in maintaining the function of the PI-derived motif to form homodimers for LMADS1.

When the role for these two amino acid residues arginine (R) and valine (V) was further examined separately by site-specific mutagenesis, different effects caused by amino acid substitutions for these two residues were observed. The ability to form homodimers decreased to approximately half of that for LMADS-208, once arginine (R) was substituted by alanine (A). By contrast, the ability to form homodimers was completely abolished once valine (V) was substituted by either alanine (A) or arginine (R). The ability to form homodimers was also completely abolished once both arginine (R) and valine (V) were substituted by alanine (A). This result indicated two things. First, it indicated that the ability of LMADS1–208 to form homodimers is not due to the number of amino acids remaining in the PI-derived motif, but rather, due to the presence of the residues arginine (R) and valine (V). Second, it clearly indicated that valine (V) at position 208 played a more important role than arginine (R) at position 207 in the PI-derived motif in regulating the formation of homodimers for LMADS1. Interestingly, arginine (R) at position 207 is the consensus residue whereas valine (V) at position 208 is not the consensus residue (L) in the PI-derived motif found in AP3 family proteins (8, 20). However, when the sequence was further analyzed, valine (V) at position 208 remains highly conserved in the PI-derived motif found in AP3 family proteins of monocots such as LMADS1, OsMADS16, SILKY, LRDEF as well as in the PI motif found in PI family proteins (Fig. 6A) (8, 24). This conservation may also reveal the possibly important function for the residue valine (V).

Another interesting question raised in this study is whether the homodimerization of LMADS1 solely depends on the paleoAP3 and PI-derived motifs. LMADS1 showed high identity (70%) to other monocot AP3 orthologues such as OsMADS16 of rice (24). Similar to LMADS1, conserved sequences, (FAFRV-
VPSQPNLH) and (GGNHDRLRGL), showed 85% (11/13) identity to the consensus sequence of the PI-derived motif and 78% (7/9) identity to the paleoAP3 motif respectively, were also identified in the C-terminal region of OsMADS16 (rice), and PI (Arabidopsis). The residues not conserved for the consensus sequences in the corresponding motif are boxed. The number under each motif indicates the number of conserved residues in this motif to the consensus sequences of the corresponding motif. B, the importance for C and K domains in forming homodimers for B function proteins. Homodimerization of a B function protein required the interaction between a conserved K domain and a corresponding ancestral C-terminal domain. LMADS1 is able to form homodimers because it contained both conserved K and C-terminal domains. The chimeric protein containing the C-terminal domain of LMADS1 was able to form homodimers once its K domain of AP3 showed a high identity (53%) to that for LMADS1. By contrast, a chimeric protein was unable to form homodimers if its K domain of PI showed a low identity (25%) to that for LMADS1. Although high K domain identity is present, the ability to form homodimers for a B protein is lost once the identity of a C-terminal domain to that of LMADS1 is below a threshold as seen for OsMADS16 and AP3. The number under each domain indicates the percentage of sequence identity for this domain to that of LMADS1. The plus and minus signs shown at the right indicates the ability for homodimerization by each wild-type or chimeric protein.
To further examine this assumption, the chimeric protein AP3-L1C containing the entire C terminus of LMADS1, plus I and K domains of Arabidopsis AP3 protein was analyzed for homodimerization. Interestingly, despite only approximately 50% identity in I and K domains for AP3 and LMADS1 (Table I), this chimeric protein AP3-L1C formed homodimers as strong as that for LMADS1 (Figs. 5 and 6B). Thus, this result provided direct evidence to support that the C-terminal region, in addition to the paleoAP3 and PI-derived motifs is also responsible for the homodimerization of LMADS1. The presence of the conserved paleoAP3 and PI-derived motifs alone is clearly not sufficient to account for ability to homodimerize as seen for OsMADS16. This assumption is further supported by the fact that OMADS3, an orchid (O. Gower Ramsey) AP3-like protein, loses the ability to form homodimers once its C-terminal region is deleted (Hsu and Yang, unpublished result).

The K-domain of MADS box proteins contains an amphipathic helix that has been thought to be involved in protein dimerization (40). Since the chimeric protein AP3-L1C has almost 100% of the ability to form homodimers as seen for LMADS1, it seems that when the K domain of LMADS1 is substituted no effect on homodimerization of LMADS1 is observed. However, in contrast to AP3-L1C, chimeric protein PI-L1C encodes a chimeric PI protein with the C terminus of LMADS1 which is unable to form homodimers (Figs. 5 and 6B). The difference between AP3-L1C and PI-L1C occurs at the I and K domains in the N terminus of the chimeric proteins. This revealed that I and K domains of AP3 and LMADS1 are also important for homodimerization. The I and K domains of AP3 showed 48 and 53% identity to that of LMADS1 (Table I). These two domains of PI showed 45 and 25% identity to that of LMADS1 (Table I). It is clear that the K domain is the major difference (53% versus 25% identity) between AP3-L1C and PI-L1C and likely participates in homodimerization. Therefore, we believe that a conserved K domain interacting with a corresponding ancestral C-terminal domain is required for the homodimerization of a B function protein (Fig. 6B). In our study, 53% identity for AP3 is above, whereas 25% identity of PI is below the threshold of the identity for a K domain to interact with the C-terminal domain of LMADS1 in forming homodimers (Fig. 6B). Interestingly, a similar pattern of sequence identity was also observed for LRGLOA, PI and AP3. The I and K domains of PI showed 52 and 54% identity to that of LMADS1 (Table I). These two domains of AP3 only showed 25 and 28% identity to that of LRGLOA (Table I). Therefore, the chimeric protein containing the C terminus of LRGLOA, plus I and K domains of Arabidopsis PI is expected to have great potential to form homodimers as seen for AP3-L1C. This predication remains under investigation.

Although percentage sequence identity gave some clues to protein interaction specificity, further investigation to determine which residues in the K domain might be responsible for a functional difference between AP3 and PI would be helpful. However, the amino acids are highly variable in the K domain of AP3 and PI (more than 50 among 67 amino acids are different). Thus, the identification of residues specifically responsible for this protein interaction becomes extremely difficult and will remain under investigation.

Based on our results, a model has been proposed to illustrate the possible evolution of dimerization for B function proteins. Genes such as LMADS1 were generated by duplication from an ancestral B gene containing an ancestral form of PI and paleoAP3 motifs at approximately 300 million years ago. This duplication also produced an ancestral PI gene that consequently evolved into genes such as LRGLOA of lily (17). The PI motif was conserved in LRGLOA and was slightly changed to the PI-derived motif in LMADS1 (Fig. 6A). To form homodimers, an ancestral form of PI or the PI-derived motif is absolutely required. A conserved paleoAP3 motif will enhance the ability to homodimerize. Thus, LMADS1 and LRGLOA retain the ability to form homodimers (17, 24). In addition to PI-derived and paleoAP3 motifs, a specific stretch of sequences with ancestral characteristics in the C domain and a certain level of conservation in the K domain are also required for homodimerization (Fig. 6B). Therefore, despite the high sequence conservation in other parts of the proteins, the severe sequence change in this C domain from LMADS1 to OsMADS16 (Table I) caused a loss of the ability to form homodimers for OsMADS16 (Fig. 6B). Further severe alteration of the PI-derived motif and replacement of the paleoAP3 motif by euAP3 motif in proteins such as AP3 and DEF of higher eudicots (Fig. 6A) completely transformed them into obligate heterodimers. This assumption may also be true for genes in the PI lineage. Because the PI motif was highly conserved in the PI lineage from LRGLOA and OsMADS4 to PI (Fig. 6A), variable sequences in the C and K domains should be observed for these three proteins and accounts for the difference in dimerization. As seen in Table I, in contrast to 72% identity in MADS box domain, only approximately 50% identity in I and K domains and 25% identity in C domain was observed between LRGLOA and PI. There is 83 and 73% identity in M and K domains between LRGLOA and OsMADS4. However, only 34% identity in C domain was observed between these two proteins (Table I). Interestingly, the percentage of sequence identity in the M, I, K, and C domains is almost identical for LMADS1/OsMADS16/AP3 or LRGLOA/OsMADS4/PI. This indicated that a similar evolutionary rate for AP3 and PI from their corresponding ancestral B genes has occurred.

### REFERENCES

1. Theissen, G., and Saedler, H. (1995) *Curr. Opin. Genet. Dev.* **5**, 628–639.
2. Purugganan, M. D., Rounsley, S. D., Schmidt, R. J., and Yanofsky, M. F. (1995) *Genetics* **140**, 345–356.
3. Rounsley, S. D., Ditta, G. S., and Yanofsky, M. F. (1995) *Plant Cell* **7**, 1259–1269.
4. Theissen, G., Becker, A., Di Rosa, A., Kanno, A, Kim, J. T., Munster, T.,...
The C-terminal Sequence of LMADS1 Is Essential for the Formation of Homodimers for B Function Proteins
Tsai-Yu Tzeng, Hsin-Ching Liu and Chang-Hsien Yang

J. Biol. Chem. 2004, 279:10747-10755.
doi: 10.1074/jbc.M311646200 originally published online December 15, 2003

Access the most updated version of this article at doi: 10.1074/jbc.M311646200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 39 references, 14 of which can be accessed free at
http://www.jbc.org/content/279/11/10747.full.html#ref-list-1