Full Length Research Paper

Characterization of a mannose-binding lectin gene from *Typhonium divaricatum* (L.) Decne

Weiwen Kong\(^2,1\), Zhongxiang Deng\(^2\), Jiong Fei\(^1\), Qian Wang\(^2\), Xiaofen Sun\(^2\) and Kexuan Tang\(^1,2\)*

\(^1\)Shanghai Key Laboratory of Agrobiotechnology, Plant Biotechnology Research Center, School of Agriculture and Biology, Fudan-SJTU-Nottingham Plant Biotechnology R&D Center, Shanghai Jiao Tong University, Shanghai 200030, China.

\(^2\)State Key Laboratory of Genetic Engineering, School of Life Sciences, Fudan-SJTU-Nottingham Plant Biotechnology R&D Center, Morgan-Tan International Center for Life Sciences, Fudan University, Shanghai 200433, China.

Accepted 18 April, 2006

Monocot mannose-binding lectins (MMBLs) or *agglutinins* are an extended superfamily of structurally and evolutionarily related proteins and they play important roles in plant defenses. Here the full-length cDNA of monocot mannose-binding *agglutinin* (designated as TDA, GenBank accession no.: AY347940) was isolated from *Typhonium divaricatum*, a traditional Chinese medicinal herb. Sequence analysis revealed that the full-length cDNA of TDA was 870 bp, and had a 594 bp open reading frame (ORF) encoding a putative 197-aa *agglutinin* precursor with a C-terminal domain. Multiple alignments of TDA amino acids with those of seven other MMBLs revealed three highly conserved domains among them, indicating TDA belongs to a member of the MMBL superfamily. Predicted tertiary structure analysis showed that TDA had three potential equal mannose-binding sites. Phylogenetic analysis indicated that 20 MMBLs including TDA belonged to an extended superfamily. Northern blot analysis showed that TDA expressed in *T. divaricatum* in a tissue-specific manner, with highest expression in tuber and almost no expression in petiole and leaf. The cloning and characterization of TDA will enable us to study its potential insect resistance function in the future.

Key words: cDNA cloning, expression, *Typhonium divaricatum* (L.) Decne, mannose-binding lectin.

INTRODUCTION

Lectins or agglutinins are kinds of proteins with at least one non-catalytic domain that binds reversibly to specific *monosaccharides* or oligosaccharides (Peumans and Van Damme, 1995), and mediate a wide variety of biological processes, such as cell to cell and host to pathogen interactions, serum glycoprotein turnover, and innate immune responses (Vijayan and Chandra, 1999). In flowering plants, all known lectins can be classified into seven families including monocot mannose-binding lectins (MMBLs), which are structurally and evolutionarily related proteins (Van Damme et al., 1998). Each member of these families possesses typical sugar-binding motifs. Mannose-binding lectins are widely distributed in higher plants and are believed to play a role in recognition of high-mannose type glycans of foreign microorganisms or plant predators. In well-known plant agglutinins, MMBL is believed to be the safest agglutinin for diet (Peumans and Van Damme, 1996).

At present, many MMBLs have been isolated and analyzed from monocotyledonous plant species including those from Amaryllidaceae (Van Damme et al., 1987a, 1991), Liliaceae (Oda and Minami, 1986; Koike et al., 1995; Van Damme et al., 1996), Orchidaceae (Van Damme et al., 1987b, 1994), Araceae (Kaku et al., 1990; Van Damme et al., 1993), Alliaceae (Smeets et al., 1997), Iridaceae (Van Damme et al., 2000), Bromeliaceae (Neuteboom et al., 2002) and Dioscoreaceae (Gaidamashvili et al., 2004). They constitute a structurally and evolutionarily related superfamily and react exclusively with mannose and mannose-binding N-glycans.

There are obvious bioactivity and resistance as plant defense proteins to insects and/or nematodes in different MMBLs. Agglutinins such as *Galanthus nivalis* agglutinin...
(GNA) can confer strong resistance to Homoptera insects (Sauvion et al., 1996). GNA is the first MMBL isolated and characterized in details (Van Damme et al., 1987). Expressions of GNA in transgenic plants conferred resistance to sap-sucking insects such as aphids, Lepidopteran, nematodes or Lacanobia oleracea (tomato moth) (Hilder et al., 1995; Gatehouse et al., 1996; Fitches et al., 1997; Ripol et al., 2003). In transgenic rice plants, expression of GNA decreased survival and overall fecundity of brown planthopper (Nilaparvata lugens), retarded its development, and had a deterrent effect on insect feeding (Rao et al., 1998).

Some other MMBLs confer resistance to pathogenic fungi in plants. For example, it was proven that agglutinins such as Gastrodia elata anti-fungal protein (GAFP-1) conferred resistance to Botrytis cinerea, Gibberella zeae, Ganoderma lucidum, Rhizoctonia solani and Valsa ambients (Xu et al., 1998; Wang et al., 2001).

Previous studies have demonstrated that the function of lectins depends on their tertiary structures characteristic of specific carbohydrate binding activities. The ways for lectins possessing specific carbohydrate-binding capacity are the extensive use of water bridges, post-translational modification and oligomerisation (Vijayan and Chandra, 1999). Small alterations in tertiary structure, consequent to variation in sequence, however, leads to large changes in quaternary association, a result that is of considerable general interest in relation to protein folding (Vijayan and Chandra, 1999). Exogenous saccharides including glycoproteins on surface of fungi, plant viruses, the midgut epithelial membrane of insects and plant predators are the most possible receptors of lectins. All MMBLs are very similar at the level of protein, whereas there are significant differences in their processing and posttranslational modifications of the primary translation products corresponding to their genes (Smeets et al., 1994).

Homopteran insects often cause severe damage to crop plants like rice, mustards, peas, cabbages, and many others. One of the most devastating are Nephotettix sp. (Green Leaf Hopper of rice) and Lipaphis erysimi (mustard aphid). They not only destroy the crops by sucking the crop’s sap but also transmit various plant viruses into the crop as pathogen vectors while feeding. Furthermore, due to their unique feeding habits and fast multiplicity, they are difficult to control by traditional approaches. Genetic engineering, by the use of insect-resistant genes, has been demonstrated to be an efficient way for the control of insect pests. However, the cloning and characterization of insect-resistant genes are the prerequisite to deploy the strategy in controlling insect pests.

As MMBLs can confer resistance to sap-sucking or Homopteran insect pests and nematodes in crops and also are mostly safe for diet, increasing studies have been performed on this group of lectins. The cloning and characterization of more genes in this group will be helpful for extending gene resources used in genetic engineering for insect resistance and for more understanding plant agglutinins with diverse function.

Typhonium divaricatum (L.) Decne is a traditional Chinese medicinal herb and no insects and disease damages are visible during its growth and development in open fields. Whether its natural resistance to pests and diseases are related to its inner lectin activities is unknown. Here we report, for the first time, the cloning and characterization of agglutinin gene from T. divaricatum by rapid amplification of cDNA ends (RACE). The expression of TDA in different tissues of T. divaricatum was also studied.

MATERIALS AND METHODS

Plant materials and RNA isolation

T. divaricatum plants were collected from Southwest Normal University campus, Chongqing, China. Plant tissues were immersed in liquid nitrogen immediately after excision and stored in -70°C before RNA extraction. CTAB extraction procedure was used to isolate total RNA (Jaakola et al., 2001).

3′ RACE of TDA

MQDCCNL is found to be a conserved motif of MMBLs based on the comparison study. Based on this sequence, primer T30 (5′-ATGAGGATGACTGAACCT-3′) was designed to amplify the 3′ cDNA end sequence of TDA. According to the protocol of the 3′ RACE System for Rapid Amplification of cDNA Ends (GIBCO BRL, Life Technologies, USA), an aliquot of isolated 100 ng RNA was reversely transcribed using the cDNA synthesis primer AP (5′-GGCCACGCGTCGACTAGTAC(T17)-3′), provided with the kit (GIBCO BRL). The gene-specific primer T30 was used as the forward primer and Universal Amplification Primer (UAP, 5′-CUACUAUCUGAGCGCGGCCTGACTGAGTAC-3′) was used as the reverse primer to amplify the 3′ end cDNA. PCR was performed in a total volume of 50 μl containing 2 μl cDNA, 2 μl 10 pmol primer T30, 2 μl 10 pmol primer UAP, 1 μl 10 mmol dNTPs, 5 μl 10 × cDNA reaction buffer and 2.5 U Taq polymerase. The amplification was performed in a GeneAmp PCR System 2400 for 3 min at 94°C, 35 cycles with 1 min at 94°C, 1 min at 58°C, 2 min at 72°C. After the final cycle, the amplification was extended for 10 min at 72°C. The primary PCR product was then used in the nested PCR amplification using Abridged Universal Amplification Primer (AUAP, 5′-GGCCACGCGTCGACTAGTAC-3′) under the same PCR condition.

After 1% agarose gel electrophoresis, the targeted DNA band was recovered using Gel Extraction Mini Kit (Watson Biotechnologies, Inc., P. R. China), and the purified DNA was ligated to pGEM-T Easy vector (Promega, USA) and then transformed into Escherichia coli strain DH5α (Sambrook et al., 1992). Two independent PCR positive clones were sequenced.

5′ RACE of TDA

According to the protocol of the SMART™ RACE cDNA Amplification Kit (Clontech), about 100 ng of total RNA extracted from leaves of T. divaricatum was reversely transcribed with primer 5′-CDS primer coupled with (dC) tailing and SMART II A oligo, which annealed to the tail of the RNA and served as an extended template for PowerScript RT. Two gene-specific primers T50 (5′-
GACGTAGTGCGGTGGACGAGGAGT (5') and T51 (5'-GCTGTAGAT
GACGACGTTGC-3') were designed to amplify the 5' end of TDA
based on the obtained 3'end sequence. The first round PCR was
performed with T50 as the reverse primer and Universal Primer A
Mix (UPM, long: 5'-CTAATACGACTCACTATAGGGCAAGCAGTG
GTATCAACGAGGAGT-3', short: 5'-CTAATACGACTCACTATA
GGGC-3') as the forward primer. PCR was carried out in a total
volume of 50 µl containing 1 µl 10 pmol primer T50, 5 µl 10 pmol UPM, 41.5 µl Master Mix (34.5 µl PCR-Grade Water, 5 µl 10 X Advantage 2 PCR buffer, 1 µl 10 mmol dNTP Mix, 1 µl Advantage 2 Polymerase Mix) under the following condition: the template was firstly denatured at 94°C for 3 min and then subjected to 35 cycles of amplification (94°C for 1 min, 58°C for 1 min, 72°C for 2 min) followed by 10 min at 72°C. Subsequently, nested-PCR was performed under the same reaction system and condition with nested primer NUP (5'-AAGCAGTGGTATCAACGAGGAGT-3') and
T51. The PCR products were cloned to pGEM-T Easy vector
(Promega, USA) and transformed into E. coli DH5α competent
cells, and then sequenced.

Amplification of TDA full-length cDNA

The sequences of 3' and 5' cDNA ends were aligned on Vector NTI Suite 8.0 to obtain the predicted full-length cDNA. Then two primers TUP (5'-AGTGCATCGAACTGATACCACAACT-3') and TDP (5'-TATTACAAGCCTCGCCACAGTCCT-3') were designed and synthesized according to the aligned sequence to amplify the full-
length TDA cDNA. PCR was carried out in a total volume of 50 µl of
reaction solution containing 5 µl of 10 × Pfu buffer with Mg2+, 1 µl of 2.5 mM each of dNTPs, 50 ng of TUP, 50 ng of TDP, 2.5 µl of first strand cDNA using the following hot-start protocol: the cDNA was
pre-denatured for 3 min at 94°C, then added with 2.5 U Pfu DNA
polymerase (Sangon) followed by 35 cycles of amplification (1 min at 94°C, 1 min at 58°C, 2 min at 72°C) and by final extension at
72°C for 10 min. The amplified blunt-ended DNA fragment was dA-
tailed using a dA-tailing Kit (Sangon). DNA ligation with pGEM-T Easy vector, transformation of DH5α and sequencing were then
carried out as mentioned above.

Bioinformatic analysis

Prediction of TDA coding ORF was carried out via the following web
site (http://www.ncbi.nlm.nih.gov/) and the signal peptide of lectin
precursor was predicted online (http://www.cbs.dtu.dk/services/
SignalP). Homology analysis of TDA was performed online at the
website of http://www.ncbi.nlm.nih.gov and prediction of tertiary
structure of TDA was done at http://cn.expasy.org. VNTI Suite 8
was used for multiple alignment of MMBLs and WebLab ViewerLite
was used for 3-D structure displaying (homology-based modeling
by Swiss-Model). Phylogenetic tree analysis of MMBLs from 20
different plant species belonging to 8 families was performed by
MEGA2.1 based on their Pfam domains.

Northern blot analysis

Northern blot analysis was carried out to investigate the expression
pattern of TDA in different tissues of T. divaricatum (Sambrook et
al., 1992). Total RNA (10 µg/sample) extracted from leaf, petiole
tuber was denatured and separated in 1.0 % formaldehyde-
denatured (w/v) agarose gel followed by transferring onto a
Hybond-N nylon membrane (Amersham Pharmacia Biotech Ltd,
Sweden). Hybridization, detection and Fuji X-ray film exposure were
carried out according to the Gene Images CDP-Star Detection
Module Kit instruction (Amersham Pharmacia Biotech Ltd,
Sweden).

RESULTS

Cloning and sequence analysis of TDA cDNA

According to the conserved amino acid sequences of
MMBLs, primer T30 was designed and synthesized to
amplify the 3' end of T. divaricatum lectin cDNA. A 584-
bp fragment was obtained in which a 3' untranslated
region of 228-bp was found downstream from the stop
codon. Two specific primers (T50 and T51) designed and
synthesized according to the obtained 3' RACE fragment
were used for the amplification of 5' TDA cDNA fragment.
A 501-bp fragment was amplified in the nested PCR,
which overlapped 145-bp nucleotides of the 3' end
sequence. The PCR product contained a 5' untranslated
region of 81-bp upstream of the first ATG start codon.
Finally the full-length cDNA sequence of TDA was
deduced and amplified by PCR using primers TUP and
TDP, which was 870 bp (Figure 1).

Figure 1. The full-length cDNA sequence and deduced amino acid
sequence of TDA from Typhonium divaricatum. The coding
sequence is shown in capital letters and the stop codon (TGA) is
marked with *. Three putative mannose-binding site sequences
((QX DX N X Y)) are boxed. The cleavage site of the putative
signal peptide sequence is indicated by an arrowhead between
A29 and V30.
Figure 2. Multiple alignment of amino acid sequences of TDA and other MMBLs. Amino acid sequences of lectins representing different MMBL families were downloaded from Genbank and multiple alignments were performed with VNTI 8.0. Highly conserved amino acid sequences are boxed. AAQ55289 (Araceae), AAA33346 (Amaryllidaceae), AAP04617 (Araceae), AAQ18904 (Amaryllidaceae), S39489 (Alliaceae), AAM28277 (Bromeliaceae), S43462 (Orchidaceae), BAD67183 (Dioscoreaceae).

Analysis on the possible ORF of TDA revealed that it coded a putative protein precursor with 197 amino acids. SignalP analysis suggested that the first 29 amino acids of the precursor functioned as a signal peptide. The precursor contained a three-fold internal repeat (beta-prism architecture). The consensus sequence motif QXDNXVXY is involved in alpha-D-mannose recognition (Figure 1).

TDA belongs to a novel member of an extended MMBL superfamily

Alignment analysis of TDA and other 7 MMBLs from six different families was performed and the result showed that all the 8 lectins shared three common conserved domains. These conserved domains conferred on them the capacity of specifically binding to carbohydrates, bioactivities and functions as plant defense weapons (Figure 2). TDA, therefore, is a member of the MMBL superfamily.

Alignment analysis demonstrates there are obvious distinction between TDA and other MMBLs including GNA. TDA contains 197 amino acid residues with a C-terminal domain, obviously longer than GNA with 157 amino acids (Figure 2). Though there is similarity in tertiary structure between TDA and GNA, the post-translational modifications of TDA may be different from GNA.
TDA expresses strongly in *T. divaricatum* tuber

To investigate the TDA expression pattern in various tissues of *T. divaricatum*, total RNA was isolated from different tissues including leaf, petiole and tuber and subjected to Northern blot analysis using the biotin-labeled TDA cDNA fragment as the probe. The result showed that TDA expressed in a tissue-specific manner in *T. divaricatum*, with the highest expression in tuber and almost no expression in leaf and petiole (Figure 3).

**Figure 3.** Northern blot analysis for the expression of TDA in different tissues of *T. divaricatum*. Total RNA was isolated from leaf (L), petiole (P) and tuber (T) of *T. divaricatum* followed by hybridization with the biotin-labeled TDA fragment as a probe (upper panel). The rRNA was used as a loading control (lower panel).

Putative tertiary structure of TDA possesses three equal mannose-binding sites

As analyzed above, TDA had three mannose-binding domains like other members of MMBL superfamily. Analysis with ExPASy Molecular Biology Server (http://us.expasy.org/) further indicated that mimic tertiary structure modeling of TDA had three equal mannose-binding sites mainly constituted by Gln(Q), Asp(D), Asn(N), Val(V) and Tyr(Y) (Figure 4), suggesting TDA to have similar bioactive functions like other members of MMBL superfamily.

**Figure 4.** Predicted tertiary structure of TDA monomer. The strands of &beta;-sheet associate in three four-stranded bundles by turns and loops and form the &beta;-prism fold. Adjacent Q_D_N_V_Y amino acids constitute the three potential mannose-binding sites [(I), (II), (III)] respectively.

Phylogenetic analysis indicates all MMBLs belong to an extended superfamily

TDA shares a high level of homology with other MMBLs from different families. As shown in Figure 5, TDA (AAQ55289), one member of the Araceae lectins, was not closely related to other *Araceae lectins* from *Amorphophallus konjac* (AAP04617), *Arisaema heterophyllum* (AAP50524), *Pinellia ternata* (AAP20876), *Zantedeschia aethiopica* (AAQ75079), suggesting that various members of the *Araceae lectins* are distributed into different subgroups. Additionally, different members of the Amaryllidaceae lectins were also divided into two diverse subgroups. This result indicates that all MMBLs from different families are strictly and structurally related proteins; which belong to an extended superfamily (Figure 5).

DISCUSSION

In this study, the full-length cDNA encoding an agglutinin designated as TDA was successfully isolated from *T. divaricatum*. TDA was showed to be a tuber-specific expressing gene and TDA contains 197 amino acid residues with a C-terminal domain. The putative tertiary structure of TDA shared typical features of MMBLs, and TDA belonged to B-lectin.

Plant lectins play an important role in resistance to pathogens, insects and plant predators (Peumans and Van Damme 1995; Barre et al., 2001). Most lectins express in a relatively low tissue- or organ-specific manner and some lectins express with higher level in plants when plants face attack by other organisms. The expression of TDA is tuber-specific and its role in plant as defense protein remains to be elucidated.

It is suggested for a long time that MMBLs play important role as defense proteins in plant resistance to hazardous organisms because of their mannose-binding capacity and erythrocytes agglutination activity. This suggestion has been demonstrated in recent years by showing that many mannose-binding lectins have more or less resistance or inhibition to various insects, nemato-
Figure 5. Phylogenetic tree built from the amino acid sequences of MMBLs from different species. Amino acid sequences of lectins representing Amaryllidaceae, Alliaceae, Araceae, Bromeliaceae, Dioscoreaceae, Iridaceae, Liliaceae and Orchidaceae MMBL families were downloaded from Genbank and phylogenetic tree analysis was performed with MEGA2.1 based on their Pfam domains (B-lectin). AAA33346 (Galanthus nivalis); AAG10402 (Crocus vernus); AAM28277 (Ananas comosus); AAM94381 (Zephyranthes candida); AAO59506 (Crinum asiaticum); AAP04617 (Amorphophallus konjac); AAP20876 (Pinellia ternata); AAP50524 (Arisaema heterophyllum); AAQ18904 (Zephyranthes grandiflora); AAQ55289 (Typhonium divaricatum); AAQ75079 (Zantedeschia aethiopica); AAW22055 (Lycoris sp. JKB-2004); BAD18020 (Dioscorea batatas); BAD67183 (Dioscorea polystachya); S38258 (Allium ursinum); S39489 (Allium porrum); S43462 (Epipactis helleborine); S43761 (Clivia miniata); S62647 (Tulipa hybrid cultivar); 2102296A (Listera ovata). AAG10402, AAP20876, AAP50524 and S62647 hold two B-lectin domains respectively, each domain (AAG10402a, b; AAP20876a, b; AAP50524a, b; and S62647a, b) of the four lectins was used to construct the phylogenetic tree.

ACKNOWLEDGEMENTS

This research is financially supported by China National “863” High-Tech Program, China Ministry of Education and Shanghai Science and Technology Committee.

REFERENCES


Fitches E, Gatehouse AMW, Gatehouse JA (1997). Effects of snowdrop lectin (GNA) delivered via artificial diet and transgenic plants on the development of tomato moth (Lacanobia oleracea) larvae in laboratory and glasshouse trials. J. Insect Physiol. 43: 727-739.


