Glycoconjugate Journal (1997) **14:** 449–456

Isolation, characterization and molecular cloning of the bark lectins from *Maackia amurensis**

Els J.M. Van Damme[1‡], Fred Van Leuven[2] and Willy J. Peumans[1]

[1] *Laboratory for Phytopathology and Plant Protection, Katholieke Universiteit Leuven, Willem de Croylaan 42, 3001 Leuven, Belgium*
[2] *Center for Human Genetics, Katholieke Universiteit Leuven, Herestraat 49, 3001 Leuven, Belgium*

A detailed study was made of the bark lectins of the legume tree *Maackia amurensis* using a combination of protein purification and cDNA cloning. The lectins, which are the most abundant bark proteins, are a complex mixture of isoforms composed of two types of subunits of 32 and 37 kDa, respectively. Isolation and characterization of the homotetrameric isoforms indicated that the 32 kDa subunit exhibits a 100-fold stronger haemagglutinating activity than the 37 kDa subunit. Molecular cloning confirmed that the two lectin subunits are encoded by different genes. The 32 kDa subunit is apparently encoded by a single gene, whereas two highly homologous genes encode the 37 kDa subunit. A comparison of the deduced amino acid sequences of the bark lectin cDNAs and the previously described cDNA encoding the seed haemagglutinin demonstrated that they are encoded by different genes.

***Keywords*: bark proteins, *Maackia amurensis*, cDNA cloning, lectin**

***Abbreviations*: *LECMAHb*, cDNA clone encoding *Maackia amurensis* bark haemagglutinin; *LECMALb*, cDNA clone encoding *Maackia amurensis* bark leucoagglutinin; MALb, *Maackia amurensis* bark leucoagglutinin; MAHb, *Maackia amurensis* bark haemagglutinin**

Introduction

Many legume species belonging to different taxonomic groupings contain carbohydrate-binding proteins which are usually referred to as lectins or agglutinins. Detailed biochemical and molecular biological studies have demonstrated that all these lectins belong to a single family of evolutionary related proteins with homologous amino acid sequences and a similar three dimensional structure. Despite their close resemblance legume lectins strongly differ with respect to their carbohydrate-binding specificity [1]. At present, legume lectins have been identified with a carbohydrate-binding site complementary to fucose, galactose/*N*-acetylgalactosamine, mannose/glucose, *N*-acetylglucosamine, sialic acid or complex oligosaccharides. Due to their highly specific sugar-binding properties many legume lectins are widely used in biological and biomedical research. Of particular interest are the sialic acid binding lectins from the seeds of the legume tree *Maackia amurensis* since they are – together with a few *Sambucus* (elderberry) agglutinins – the only plant lectins that specifically bind to sialic acid. Moreover, since the *Maackia amurensis* lectins are the only carbohydrate-binding proteins that specifically recognize the Neu5Acα2,3Galβ1,4GlcNAc/Glc sequence, they are extremely useful to detect, localize and isolate glycoconjugates carrying this particular type of oligosaccharide [2].

In the first report on the *Maackia amurensis* lectins Kawaguchi *et al.* [3] described the purification and characterization of two different carbohydrate-binding proteins from the seeds. Using affinity chromatography on immobilized porcine thyroglobulin a strong haemagglutinin (called MAM) and a potent mitogenic lectin (called MAL) were isolated. Later, a reinvestigation of the mitogenic lectin revealed that it strongly agglutinated the mouse lymphoma cell line BW5147, which prompted the authors to call it *Maackia amurensis* leucoagglutinin or MAL [4]. Detailed carbohydrate-binding specificity studies indicated that MAL had a high affinity for complex-type tri- and tetra-antennary Asn-linked oligosaccharides containing terminal sialic acid residues linked α2,3 to penultimate galactose residues [4]. More recently, a detailed characterization demonstrated that MAL (but not MAH) reacts with greatest affinity with the trisaccharide sequence Neu5Acα2,3Galβ1,4GlcNAc/Glc [5]. Similar studies on the carbohydrate-binding specificity of MAH indicated that the sialylated glycans recognized by the haemagglutinin are

* The nucleotide sequences reported in this paper will appear in the Genbank™/EMBL Data library with the accession numbers U65008, U65009 and U65010.

‡ To whom correspondence should be addressed. Fax: 32 16 322976.

0282–0080

structurally different from these bound by the leucoagglutinin. Using affinity chromatography of various glycopeptides from human glycophorin A on immobilized MAH, Konami *et al.* [6] concluded that MAH reacts with the highest activity with the Ser/Thr-linked tetrasaccharide Neu5Acα2,3Galβ1,3(Neu5Acα2,6)GalNac. In an attempt to elucidate the relationship between the particular specificity of MAH and its primary structure, the amino acid sequence of the purified protein as well as the deduced amino acid sequence of its corresponding cDNA were determined [7, 8]. Thereby it was observed that one asparagine residue, which has been conserved in the primary sequence of all other legume lectins, is substituted in MAH by aspartic acid.

In this report we describe the isolation, partial characterization and molecular cloning of two different lectins from the bark of *Maackia amurensis*. Both bark lectins strongly differ with respect to their agglutination activity. Sequence analysis of their corresponding cDNAs demonstrated that the subunits of the two bark lectins are encoded by different genes. In addition, a comparison to the sequence of MAH from seeds, indicated that both bark lectins are homologous to but not identical to the seed lectin.

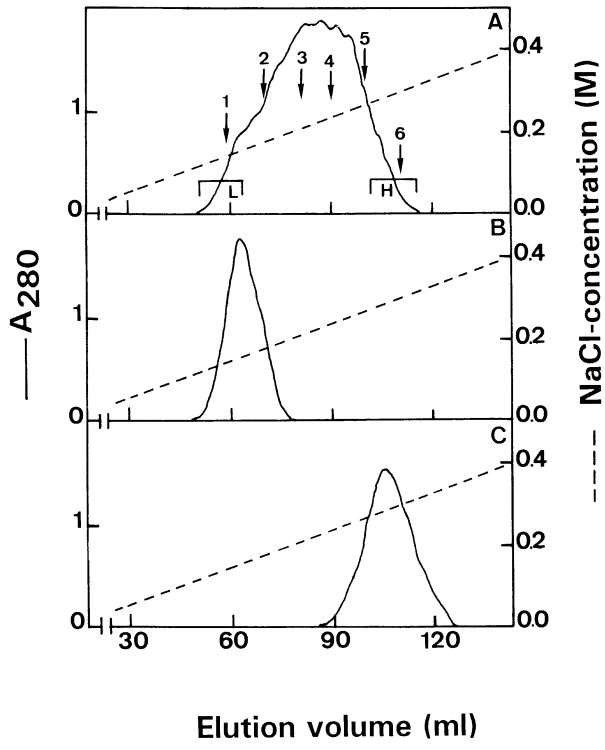

Figure 1. Ion exchange chromatography of the *Maackia amurensis* bark lectins. (A) Total affinity-purified lectin was subjected to ion exchange chromatography on a semipreparative Neobar CS15/4 column (Dynochrom, Lillestrom, Norway) using a Pharmacia FPLC system. Samples (50 mg) dissolved in formate buffer (25 mM formic acid adjusted to pH 3.8 with NaOH) were loaded on the column. After washing the column with 12 ml of buffer, proteins were eluted with a linear gradient (180 ml) of increasing NaCl concentration (0–0.5 M) (in the same buffer) at a flow rate of 3 ml min^{-1}. Fractions (3 ml each) were collected and analysed by SDS-PAGE (see Fig. 2A). Fractions eluting at the beginning (L) and at the end (H) of the peak were pooled as indicated and rechromatographed twice. The elution patterns of the final preparations are shown in B and C, respectively.

Materials and methods

Plant material

Branches were collected from a *Maackia amurensis* tree at the arboretum of Bokrijk (Belgium). After removal of the sclerotized epidermis the inner bark was stripped with a knife, cut into small pieces and stored at $-70\,^{\circ}C$ or $-20\,^{\circ}C$. Bark samples used for the isolation of RNA were collected at the end of September since legume bark lectins usually accumulate during the autumn. The bark material destined for the isolation of proteins was collected at the end of December.

Isolation of the *Maackia amurensis* bark lectins

The bark of *Maackia amurensis* was lyophilized and powdered in a coffee mill. Twenty grams of the bark meal were extracted in 200 ml of 0.2 M NaCl containing 0.2 g l^{-1} ascorbic acid (adjusted to pH 6.5) by stirring for 1 h at room temperature and the resulting homogenate centrifuged at $9000 \times \mathbf{g}$ for 10 min. The supernatant was taken off, filtered through filter paper (Whatman 3 MM) and loaded on a column (2.6 cm $\times$ 10 cm; 50 ml bed volume) of fetuin-Sepharose 4B equilibrated with PBS (1.5 mM KH_2PO_4, 10 mM $NaHPO_4$ (pH 7.4), 3 mM KCl, 140 mM NaCl). After passing the extract, the column was washed with PBS until the A_{280} fell below 0.01 and the lectins desorbed with 20 mM acetic acid. Using this procedure, all the agglutinating activity present in the crude extract was retained on the column and could be eluted with acetic acid. The yield of affinity purified lectin was about 12 mg g^{-1} of bark meal.

Purification of the homotetrameric *Maackia amurensis* bark lectins

To isolate the individual lectins, the affinity-purified lectin fraction was dialysed against water, lyophilized and redissolved in 20 ml of Na-formate buffer. Aliquots of about 10 mg were loaded onto a semi-preparative FPLC ion exchange column (Neobar CS15/4, Dynochrom, Lillestrom, Norway) and eluted with a linear gradient of increasing NaCl concentration in Na-formate buffer. As shown in Figure 1A, the total affinity-purified lectin fraction eluted in a broad peak. SDS-PAGE analysis of the different fractions indicated that the polypeptide pattern changed as a function of the NaCl concentration. The first fractions contained only the 37 kDa polypeptide whereas the last fractions yielded a single polypeptide of 32 kDa. In the fractions inbetween, the intensity of the 32 kDa polypeptide band gradually increased as a function of the elution time

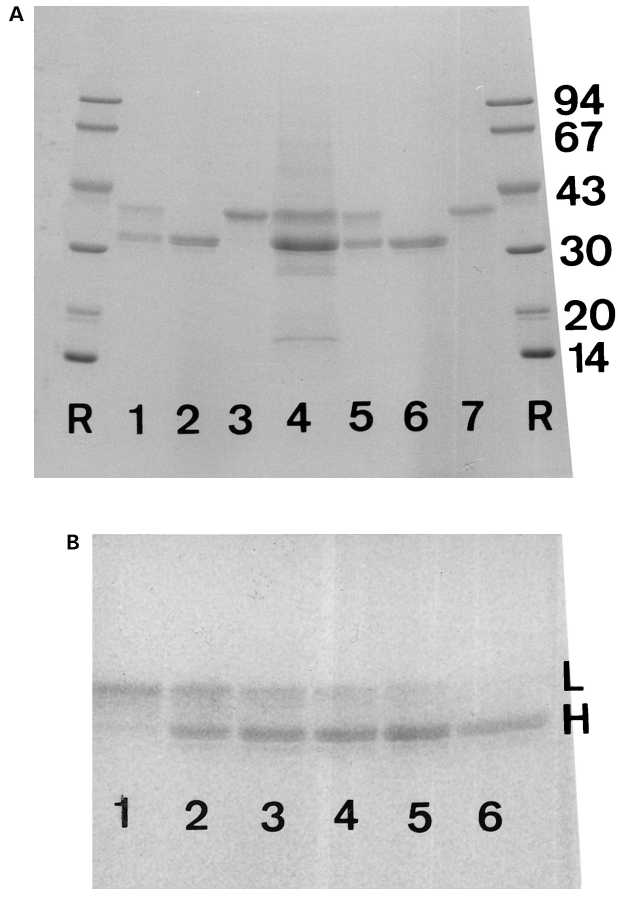

Figure 2. A. SDS-PAGE of the crude extract and the purified lectins from *Maackia amurensis* bark. Lanes 1, 2 and 3 were loaded with 25 µg of reduced (with 2-mercaptoethanol) total bark lectin, and purified MAHb and MALb, respectively. Crude extract (50 µl; reduced) was loaded in lane 4. Lanes 5, 6 and 7 were loaded with 25 µg of unreduced total lectin, and purified MAHb and MALb, respectively. Molecular mass reference proteins (lysozyme (14 kDa), soybean trypsin inhibitor (20 kDa), carbonic anhydrase (30 kDa), ovalbumin (43 kDa), bovine serum albumin (67 kDa) and phosphorylase b (94 kDa)) were run in lane R. B. SDS-PAGE of fractions 1–6 after ion exchange chromatography of total affinity purified lectin (see Fig. 1). Lanes 1–6 contain fractions 1–6 after reduction with 2-mercaptoethanol. H and L refer to the subunits of MAHb and MALb, respectively.

(Figure 2B). The first and last fractions containing the lectin were pooled as indicated in Figure 1A, dialysed against an Na-formate buffer and rechromatographed twice. As shown in Figure 2A, the final preparations (shown in Figure 1B and C, respectively) contained exclusively 32 and 37 kDa polypeptides, respectively. These preparations were dialysed against water, lyophilized and used for further analyses.

Gel filtration

Gel filtration of purified *Maackia* lectins was done on a Pharmacia (Uppsala, Sweden) Superose 12 column equilibrated with PBS containing 0.1 M galactose. Lectin samples (200 µl containing about 500 µg pure protein) were loaded on the column and chromatographed at a flow rate of 20 ml h^{-1}. Since some lectins behave anomalously upon gel filtration we used the lectins from *Phaseolus vulgaris* (M_r 130 kDa) and pea lectin (M_r 50 kDa) as molecular mass reference markers.

Electrophoresis

Lectin preparations were analysed by SDS-PAGE using 12.5–25% (w/v) acrylamide gradient gels as described by Laemmli [9].

Agglutination assays

Haemagglutination assays were carried out in small glass tubes in a final volume of 0.1 ml containing 80 µl of a 1% suspension of human (type A) red blood cells and 20 µl of lectin solutions (serially diluted with two-fold increments). Agglutination was controlled visually after 1 h at room temperature. The leucoagglutinating activity was determined with a suspension (400 000 cells ml^{-1}) of murine leukaemia L1210 cells (derived from a lymphomic leukaemia). The latter cells were kindly provided by Dr J. Balzarini (Rega Institute, Leuven, Belgium).

Carbohydrate analysis

The total carbohydrate content of the lectins was estimated by the phenol sulphuric acid method of Dubois *et al.* [10] using D-glucose as a reference.

Amino acid sequence analysis

Samples of purified *Maackia* lectins were separated by SDS-PAGE and electroblotted on a PVDF membrane. Polypeptides were excised from the blots and sequenced on an Applied Biosystems (Foster City, California, USA) Model 477A protein sequencer interfaced with an Applied Biosystems model 120A on-line analyser.

RNA isolation

Maackia bark was powdered in liquid nitrogen with mortar and pestle and the total cellular RNA prepared as described by Van Damme and Peumans [11]. Poly(A)-rich RNA was prepared from the total RNA by chromatography on oligodeoxythymidine cellulose.

Construction and screening of cDNA library

A cDNA library was contructed from poly(A)-rich RNA isolated from *Maackia* bark using the cDNA synthesis kit from Pharmacia (Uppsala, Sweden). cDNA fragments were inserted into the *Eco*RI site of the multifunctional phagemid pT_7T_3 18U (Pharmacia, Uppsala, Sweden). The library was propagated in *Escherichia coli* XL1 Blue (Stratagene, La Jolla, USA).

Recombinant lectin clones were screened using the random-primer-labelled bark lectin cDNA clones from *Cladrastis lutea* [12] and *Robinia pseudoacacia* [13] as probes. Hybridization was done overnight at 50–65 °C as reported previously [14]. Colonies that produced positive signals were selected and rescreened at low density using the same conditions. Plasmids were isolated from purified single colonies on a miniprep scale using the alkaline lysis method as described by Mierendorf and Pfeffer [15] and sequenced by the dideoxy method [16]. DNA sequences were analysed using programs from PC Gene (Intelligenetics, Mountain View, USA) and Genepro (Riverside Scientific, Seattle, USA).

Northern blot analysis

RNA electrophoresis was performed according to Maniatis *et al.* [17]. Approximately 3 μg of poly(A)-rich RNA was denatured in glyoxal and Me_2SO and separated in a 1.2% (w/v) agarose gel. Following electrophoresis the RNA was transferred to Immobilon N membranes (Millipore, Bedford, USA) and the blot hybridized using a random-primer-labelled lectin cDNA insert. Hybridization was performed as reported by Van Damme *et al.* [18]. An RNA ladder (0.16–1.77 kb) was used as a marker.

Genomic DNA analysis

Total DNA from *Maackia* bark was isolated according to the procedures described by Dellaporta *et al.* [19] and de Kochko and Hamon [20]. The DNA preparation was treated with RNAse in order to remove any contaminating RNA. Approximately 50 μg of DNA was digested with restriction endonucleases and subjected to electrophoresis in a 0.8% (w/v) agarose gel. DNA was transferred to Immobilon N membranes (Millipore, Bedford, USA) and hybridized at 68 °C using the ^{32}P-labelled cDNA insert as a probe [14].

Results

Nomenclature of the *Maackia amurensis* lectins and lectin genes

Seeds of *Maackia amurensis* contain two lectins which exhibit different agglutination and mitogenic properties. One of them is a strong haemagglutinin whereas the other is a strong mitogen. According to these activities the lectins were originally referred to as *Maackia amurensis* haemagglutinin (MAH) and *Maackia amurensis* mitogen (MAM), respectively [3]. Later, the mitogenic lectin has been designated as the *Maackia amurensis* leucoagglutinin (MAL) because of its ability to agglutinate the mouse lymphoma cell line BW5147 [4]. As will be demonstrated below *Maackia* contains also two bark lectins, which like the seed lectins clearly differ with respect to their agglutination properties. To distinguish these bark lectins from the previously described seed lectins, a new nomenclature is proposed. MAHs and MALs refer to the haemagglutinating and leucoagglutinating seed lectins, respectively. Their corresponding genes are *LECMAHs* and *LECMALs*. Similarly, the haemagglutinating and leucoagglutinating bark lectins are designated as MAHb and MALb, respectively, and their corresponding genes are referred to as *LECMAHb* and *LECMALb*

Lectins are the most prominent proteins in the bark of *Maackia amurensis*

Crude extracts from *Maackia* bark readily agglutinate human erythrocytes. The agglutination activity was not affected by any simple sugar but was completely inhibited by fetuin. To isolate the bark lectins, a procedure was developed based on affinity chromatography on immobilized fetuin. SDS-PAGE of the affinity-purified lectin indicated that it contained two polypeptides of 32 and 37 kDa, respectively (Figure 2A). A comparison of the polypeptide pattern of the crude bark extract to that of the affinity-purified lectin revealed that the most prominent protein bands correspond to the lectin subunits. An estimation based on determinations of the total protein content (about 20 mg g^{-1} bark meal) and the yield of affinity-purified lectin (12 mg g^{-1} bark meal) indicated that the lectins represent about 60% of the soluble bark protein. The latter value is in good agreement with the relative intensity of the polypeptide bands after SDS-PAGE of the crude extract.

Isolation and characterization of the *Maackia amurensis* bark lectins

Total affinity purified bark lectin yielded two polypeptide bands of 32 and 37 kDa (Figure 2A) upon SDS-PAGE and eluted with an apparent M_r of about 125 kDa upon gel filtration on a Superose 12 column (results not shown), which indicated that it is a tetrameric protein composed of two different subunits. To investigate the possible occurrence of isolectins, the affinity-purified lectin was analysed by ion exchange chromatography. Unlike the *Maackia* seed lectin, which could easily be resolved in MALs and MAHs by cation exchange chromatography [3], total affinity-purified bark lectin eluted in a broad peak (Figure 1A). However, SDS-PAGE analysis indicated that the polypeptide pattern of the lectin changed as a function of elution time (Figure 2B). The first eluting lectin contained exclusively the 37 kDa polypeptide. Upon increasing the NaCl concentration, the 32 kDa polypeptide appeared and gradually increased. At the end of the peak, the lectin fractions contained exclusively 32 kDa subunits. This particular pattern is reminiscent of the isolectin pattern of the lectins from the seeds of *Phaseolus vulgaris* and the bark of *Robinia pseudoacacia* and *Sophora japonica*. In the latter three cases two different lectin subunits associate in all possible

combinations into tetramers whereby two homotetrameric and three heterotetrameric isolectins are formed [13, 21]. By repeating the ion exchange chromatography step with the *Maackia* bark lectins, preparations containing exclusively 37 or 32 kDa subunits were obtained (shown in Figures 1B and C, respectively). These preparations were used for further analysis and characterization of the bark lectins. It should be emphasized that the homotetrameric isolectins represent only a small fraction of the total lectin. The heterotetrameric forms, which represent the bulk of the lectin were not separated since they did not yield separate peaks upon ion exchange chromatography. However, since the lectins in the top fractions of the peak shown in Figure 1A eluted consistently (in a symmetrical peak) at the same position upon repeated rechromatography (results not shown) there is not doubt that they are heterotetramers rather than mixtures of uncompletely resolved homotetramers of 37 or 32 kDa subunits.

The specific agglutination activity (defined as the lowest concentration which still yielded a visible agglutination) of the lectin composed of 32 kDa subunits was 4 $\mu g\,ml^{-1}$ and 0.2 $\mu g\,ml^{-1}$ with trypsin-treated rabbit and human (type A) erythrocytes, respectively. In the same test, the lectin composed of 37 kDa subunits yielded values of $>1000\ \mu g\,ml^{-1}$ and 20 $\mu g\,ml^{-1}$, respectively. When (untreated) murine leukaemia cells (L1210) were used the specific agglutination activity of the lectins composed at 32 and 37 kDa subunits was 25 $\mu g\,ml^{-1}$ and 250 $\mu g\,ml^{-1}$, respectively. Using the analogy of the seed lectins, the former lectin is called MAHb and the latter MALb. Accordingly, the 32 and 37 kDa subunits are indicated as the erythro- and leucoagglutinating subunits, respectively.

The molecular structure of the homotetrameric bark lectins was determined by SDS-PAGE and gel filtration. MAHb and MALb yielded a single polypeptide band of 32 and 37 kDa, respectively (Figure 2A). Since reduced and unreduced samples yielded identical results, the subunits are not held together by disulphide bridges. Gel filtration of the native lectins indicated that MAHb and MALb eluted with an apparent M_r of 125 kDa and 135 kDa, respectively (results not shown). It appears, therefore that MAHb and MALb are tetramers composed of four identical subunits of 32 and 37 kDa, respectively.

Estimations of the total sugar content indicated that MAHb and MALb contained 4.2% and 12.1% (w/w) covalently bound carbohydrate, which corresponds to eight and 26 monosaccharide units, respectively, per polypeptide chain. Taking into consideration the fact that the glycan chains of legume lectins are hexa, hepta or octosaccharides, it can be concluded that MAHb and MALb contain on average one and three oligosaccharide side chains, respectively, per subunit. The differences in glycosylation between MAHb and MALb are in good agreement with the apparent difference in M_r between their respecive subunits.

SDELS FPINN FVPNE ADLHF: MAHb subunit

SDELS FTINN FVPNE ADLLF: MALb subunit

SDELS FTINN FMPNQ GDLLF: MAHs subunit

Figure 3. N-terminal amino acid sequences of the polypeptides of the *Maackia amurensis* bark and seed lectins. The results for MAHs are taken from [7].

N-terminal amino acid sequencing of the MALb and MAHb polypeptides revealed that both subunits are highly homologous (18 amino acids out of 20 are identical) to but not identical to each other (Figure 3). In addition, a comparison with the previously reported N-terminal sequence of MAHs [7, 8] indicated that both bark lectins differ from the seed lectin.

Molecular cloning of the *Maackia amurensis* bark lectins

Screening of a cDNA library constructed from poly(A)-rich RNA from *Maackia* bark tissue yielded three groups of positive clones. A first group of clones (*LECMALbI*) has a deduced N-terminal amino acid sequence corresponding to that of the MAHb subunit. Sequence analysis of these clones revealed that they contain an open reading frame of 848 bp encoding a 282 amino acid protein of 30.8 kDa (Figure 4). According to the rules for protein processing of von Heijne [22] this cDNA clone contains a partial signal peptide of 24 amino acids. Cleavage of this signal peptide will result in a lectin polypeptide of 28.1 kDa with an isoelectric point of 4.27. The deduced amino acid sequence of the cDNA clone *LECMALbI* contains four putative *N*-glycosylation sites at positions 61, 113, 114 and 191, respectively, of the mature lectin polypeptide. According to the carbohydrate content of the MALb subunit only three of these putative glycosylation sites are occupied.

A second group of cDNa clones (*LECMALbII*) yields an amino acid sequence which shows a high degree of sequence homology to *LECMAHbI* but differs from the first group at some positions of the sequence. *LECMALbII* encodes a 31.2 kDa protein which after cleavage of the (partial) signal peptide will yield a lectin polypeptide of 28.1 kDa with an N-terminal sequence identical to that of *LECMALbI* and an isoelectric point of 4.18. Similar to the clone *LECMALbI* the deduced amino acid sequence of the cDNA clone *LECMALbII* also contain four putative *N*-glycosylation sites at positions 61, 113, 179 and 191, respectively, of the mature lectin polypeptide.

Screening of the cDNA library from *Maackia* also yielded a third group of cDNA clones (*LECMAHb*) which definitely differed from the first two groups. Unfortunately these clones contained only part of the mature coding sequence of the *Maackia* lectin. Therefore the missing part of these

```
LECMALbI        SKPTQVLLATFLTFFFLQLNNVNSSDELSFTINNFL   12
LECMALbII   ATSNSKPTQVLLATFLTFFFLLLNNVNSSDELSFTINNFV   12
LECMAHb                                 SDELSFPINNFV   12
LECMAHs    MATSNSKPTQVLLATFLTFFLLLLNNVNSSDELSFTINNFM   12
                                        ******.****.

LECMALbI    PNEADLLFQGEASVSSTGVLQLTRVENGQPQKYSVGRALY   52
LECMALbII   PNEADLLFQGEASVSSTGVLQLTRVENGQPQQYSVGRALY   52
LECMAHb     PNEADLHFQGVASVSPTGVLQLTSQKNGQPLEYSVGRALY   52
LECMAHs     PNQGDLLFQGVATVSPTGVLQLTSEENGQPLEYSVGRALY   52
            **..** *** *.**.*******. .**** .********

LECMALbI    AAPVRIWDNTTGSVASFSTSFTFVVKAPNPSITSNGLAFF   92
LECMALbII   AAPVRIWDNTTGSVASFSTSFTFVVKAPNPTITSDGLAFF   92
LECMAHb     SAPVRIWDSTTGRVASFSTSFTFVVQKA-ARLTSDGLAFF   91
LECMAHs     TAPVRIWDSTTGAVASFSTSFTFVVKA--ARGASDGLAFF   90
            .*******.*** ************. .     *.*****

LECMALbI    LAPPDSQIPTGSVTKYLGLFNNTSDDSSNQIVAVEFDTYF  132
LECMALbII   LAPPDSQIPSGRVSKYLGLFNNSNSDSSNQIVAVEFDTYF  132
LECMAHb     LAPPDSQIPSGDVSKYLGLFNNSNSESSNQIVAVEFDTFF  131
LECMAHs     LAPPDSQIPSGSVSKYLGLFNNSNSDSSNQIVAVEFDTYF  130
            *********.* *.********....************.*

LECMALbI    LHKYNPWDPNYRHIGIDVNGIDSIKTVQWDWINGGVAFAT  172
LECMALbII   GHSYDPWDPNYRHIGIDVNGIESIKTVQWDWINGGVAFAT  172
LECMAHb     NHNYDPWDPNYRHIGIDVNGIDSIKTVQWDYINGGVAFAT  171
LECMAHs     GHSYDPWDPNYRHIGIDVNGIESIKTVQWDWINGGVAFAT  170
             *.*.****************.********.*********

LECMALbI    ITYLAPSKTLIASLVYPSNQTSFIVAASVDLKEILPEWVR  212
LECMALbII   ITYLAPNKTLIASLVYPSNQTSFIVAASVDLKEILPEWVR  212
LECMAHb     ITYLAPNKTLIASLVYPSSETSFIVAASVDLKEILPEWVR  211
LECMAHs     ITYLAPNKTLIASLVYPSNQTSFIVAASVDLKGILPEWVR  210
            ******.***********..************.*******

LECMALbI    VGFSAATGYPTEVETHDVLSWSFTSTLEANSDAATENNVH  252
LECMALbII   VGFSAATGYPTQVETHDVLSWSFTSTLEANSDAATENNVH  252
LECMAHb     VGFSAATGAPAAAETHDVRSWSFTSTFEANSPA-VDHNVH  250
LECMAHs     VGFSAATGAPKAVETHDVRSWSFTSTLEANSPADVDNNVH  250
            ******** * ..***** *******.**** * ...***

LECMALbI    IARYTA  258
LECMALbII   IARYTA  258
LECMAHb     IARYTA  256
LECMAHs     IARYTA  256
            ******
```

Figure 4. Sequence alignment of the deduced amino acid sequences of the cDNA clones *LECMALbI*, *LECMALbII*, *LECMAHb*, and the cDNA encoding the seed lectin MAHs. (·) represents gaps introduced for maximal alignment. Putative *N*-glycosylation sites are shown in bold. The arrowhead indicates the cleavage site for the signal peptide. Amino acids conserved among all *Maackia* sequences are indicated by asterisks and chemically similar residues are denoted by dots. The results for *LECMAH*s are taken from [8].

clones was amplified by PCR. Sequence analysis of these clones revealed that they have an N-terminal amino acid sequence which matches the N-terminal sequence of the *MAHb* subunit. *LECMAHb* encodes a 256 amino acid lectin polypeptide of 27.8 kDa with an isoelectric point of 4.52 (Figure 4). The sequence of MAHb contains two putative *N*-glycosylation sites at positions 112 and 178 of the mature lectin polypeptide. Since the MAHb subunit contains a single glycan chain, only one putative glycosylation site is actually occupied.

Sequence homology between the bark lectin cDNA clones from *Maackia amurensis*

The lectin polypeptides encoded by *LECMAHb*, *LECMALbI* and *LECMALbII* show a high degree of sequence homology both at the nucleotide and amino acid level. Sequence alignment of the deduced amino acid sequences of the cDNAs encoding *MALbI* and *MALbII* revealed 93.8% sequence identity (96.1% sequence similarity). MALbI and *MALbII* show 85.6 and 89.4% sequence identity to the

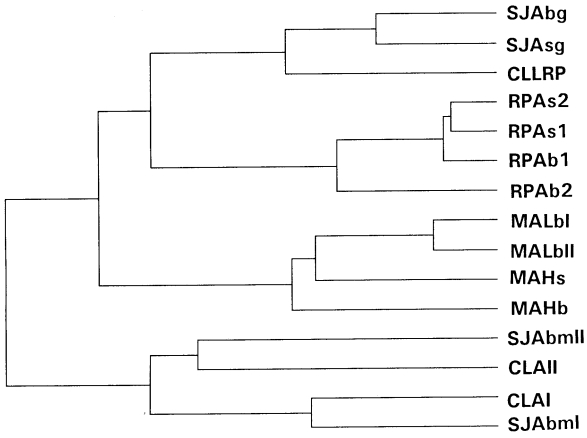

Figure 5. Phylogenetic tree of the *Maackia amurensis* lectins and some other related legume lectins. The dendrogram was constructed using the simultaneous alignment and phylogeny program (CLUSTAL) from the PC Gene software package (Intelligenetics). The different sequences encode the *Cladrastis lutea* bark lectins (CLAI, CLAII, CLLRP), *Robinia pseudoacacia* bark (RPAb1, RPAb2) and seed (RPAs1, RPAs2) lectins, *Maackia amurensis* bark (MAHb, MALbI, MALbII) and seed (MAHs) lectins, *Sophora japonica* galactose-binding bark lectin (SJAbg), mannose-binding bark lectins (SJAbmI, SJAbmII) and galactose-binding seed lectin (SJAsg).

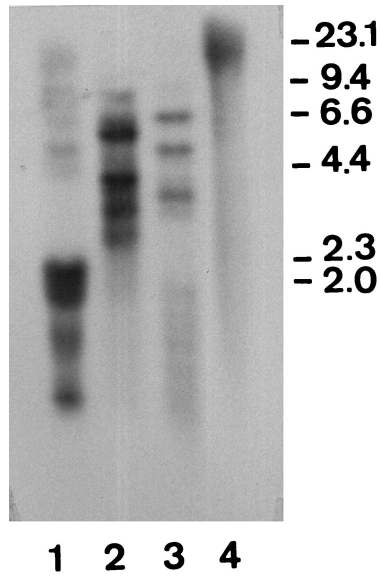

Figure 6. Southern blot of genomic DNA isolated from *Maackia amurensis* bark. DNA was digested with *Bam*HI (lane 1), *Eco*RI (lane 2), *Hind*III (lane 3) and *Pst*I (lane 4), and hybridized with the ^{32}P-labelled cDNA insert *LECMAHb* under very stringent conditions. Numbers on the right show DNA size (kb).

sequence of *Maackia* seed lectin reported by Yamamoto *et al.* [8]. Sequence alignment of MALbI and MALbII with the sequence of MAHb showed 82.4 and 85.6% sequence identity, respectively (86.7 and 88.7% sequence similarity). The sequence of MAHb reveals 89.4% sequence identity (92.5% sequence similarity to the *Maackia* seed lectin.

Sequence homology between lectin cDNA clones from *Maackia amurensis* and other legume lectins

Screening of the database for sequence homology between the *Maackia amurensis* bark lectins and previously cloned legume lectin genes revealed a reasonable homology with the mature coding sequences of all legume lectins. According to the dendrogram of alignment (Figure 5) the *Maackia* bark agglutinins are most closely related to the *Maackia amurensis* seed lectin [8], the *Robinia* seed and bark lectins [13, 23], the *Cladrastis lutea* bark lectins [12] and the galactose-binding lectins from *Sophora japonica* bark [24].

Northern blot analysis

To determine the total length of the lectin mRNAs a blot containing approximately 3 μg of poly(A)-rich RNA from *Maackia* bark was hybridized with the labelled lectin cDNA clones *LECMAHb* and *LECMALb*. Hybridization of the blot yielded a single band of approximately 1300 nucleotides for each probe (results not shown) which is in good agreement with the length of the cDNA clones isolated from the cDNA library constructed from the same RNA.

Southern blot analysis

Since the results from cDNA analysis suggest that different lectin genes are expressed in *Maackia amurensis*, Southern blot analysis was performed. Therefore, genomic DNA isolated from *Maackia* bark was digested with *Bam*HI, *Eco*RI, *Hind*III and *Pst*I. To visualize the genomic restriction fragments carrying lectin sequences the blots were hybridized using labelled cDNA inserts *LECMALbII* and *LECMAHb* as probes. The results shown in Figure 6 clearly demonstrate that the cDNA insert reacts with several restriction fragments. Both probes yielded identical results.

Discussion

Using a combination of protein purification and cDNA cloning we were able to demonstrate that the bark of the legume tree *Maackia amurensis* contains large quantities of lectins, which are very similar to but not identical with the previously described seed lectins. Analysis of the total bark lectin indicated that it is a complex mixture of isolectins, which are formed by the association of two different types of subunits of 32 and 37 kDa, respectively. Characterization of the homotetramers revealed that the haemagglutinating activity of the 32 kDa subunit is about two orders of magnitude higher than that of the 37 kDa subunit whereas the difference in leucoagglutinating activity is much less pronounced. In this respect, the 32 and 37 kDa polypeptides closely resemble the subunits of the seed haemagglutinin and leucoagglutinin, respectively. However, whereas the subunits of the MAHs and MALs apparently associate

exclusively with identical lectin polypeptides into homotetramers, the 32 and 37 kDa polypeptides of the bark lectins are capable of forming heterotetramers. At present, there is no explanation for this differential behaviour of the seed and bark lectins. It should be mentioned, however, that a similar phenomenon has been observed in *Robinia pseudoacacia*. In this legume tree also the seeds contain exclusively homotetrameric lectins, whereas the bark contains homotetrameric as well as heterotetrameric isolectins [13, 23].

Molecular cloning of the *Maackia* lectins confirmed the occurrence of two different lectin polypeptides in the bark. Moreover, sequence analysis of multiple cDNAs further suggested that the MALb subunit is encoded by two slightly different genes. The identification of the MALb and MAHb clones was based primarily on their homology with the N-terminal sequence of the respective polypeptides. A comparison of the isoelectric points of the polypeptides derived from *LECMALbI* and *LECMALbII* (4.27 and 4.18, respectively) and *LECMAHb* (4.52) supports this identification since MALb elutes before MAHb from the cation exchange column (as shown in Figure 1A). In addition, the presence of three oligosaccharide chains on the MALb subunits implies that they are encoded by *LECMALbI* and *LECMALbII* since the deduced sequence of *LECMAHb* contains only two putative glycosylation sites (whereas *LECMALbI* and *LECMALbII* contain four such sites) (Figure 4).

A comparison of the deduced sequences of the *Maackia amurensis* lectins indicated that MAHs is more closely related to MALb than to MAHb. This is rather surprising since one would expect that the homology is the highest between the two strong haemagglutinins MAHs and MAHb. Since no sequence data are available of MALs, no sequence comparisons can be made between the seed leucoagglutinin and the other *Maackia* lectins. A final point to discuss concerns the amino acid residues in the presumed carbohydrate-binding domain. According to Konami *et al.* [7] and Yamamoto *et al.* [8] an Asn residue, which is conserved within the carbohydrate recognition domain of all legume lectins is replaced in MAHs by an Asp residue (D_{137}). Our results confirm that such a replacement occurs also in *LECMAHb* and *LECMALbII* but not in *LECMALbI*. Therefore, the substitution of N_{137} by an aspartic acid residue can not be generalized for all *Maackia* lectins.

Acknowledgements

This work was supported in part by grants from the Catholic University of Leuven (OT/94/17) and the National Fund for Scientific Research (Belgium, Fonds voor Geneeskundig Wetenschappelijk Onderzoek grant 2.0046.93). W.P. is Research Director and E.V.D. Postdoctoral Fellow of this fund. Furthermore we want to acknowledge grant 7.0047.90 from the Nationaal Fonds voor Wetenschappelijk Onderzoek-Levenslijn fund. We thank C. Cuyx (Bokrijk) for providing the *Maackia amurensis* plant material.

References

1 Sharon N, Lis H (1990) *FASEB J* **4**: 3198–208.
2 Sata T, Lackie PM, Taatjes J, Peumans W, Roth J (1989) *J Histochem Cytochem* **37**: 1577–88.
3 Kawaguchi T, Matsumoto I, Osawa T (1974) *J Biol Chem* **249**: 2786–92.
4 Wang W-C, Cummings RD (1988) *J Biol Chem* **263**: 4576–85.
5 Knibbs R, Goldstein IJ, Ratcliff RM, Shibuya N (1991) *J Biol Chem* **266**: 83–88.
6 Konami Y, Yamamoto K, Osawa T, Irimura T (1994) *FEBS Lett* **342**: 334–38.
7 Konami Y, Ishida C, Yamamoto K, Osawa T, Irimura T (1994) *J Biochem* **115**: 767–77.
8 Yamamoto K, Ishida C, Saito M, Konami Y, Osawa T, Irimura T (1994) *Glycoconjugate J* **11**: 572–75.
9 Laemmli UK (1970) *Nature* **227**: 680–85.
10 Dubois M, Gilles KA, Hamilton JK, Rebers PA, Smith F (1956) *Anal Chem* **28**: 350–56.
11 Van Damme EJM, Peumans WJ (1993) In *Lectins and Glycobiology* (Gabius H-J, Gabius S, eds) pp. 458–68, Germany: Springer Verlag.
12 Van Damme EJM, Barre A, Bemer V, Rougé P, Van Leuven F, Peumans WJ (1995) *Plant Mol Biol* **29**: 579–98.
13 Van Damme EJM, Barre A, Smeets K, Torrekens S, Van Leuven F, Rougé P, Peumans WJ (1995) *Plant Physiol* **107**: 833–43.
14 Van Damme EJM, Goldstein IJ, Vercammen G, Vuylsteke J, Peumans WJ (1992) *Physiol Plant* **86**: 245–52.
15 Mierendorf RC, Pfeffer D (1987) *Meth Enzymol* **152**: 556–62.
16 Sanger F, Nicklen S, Coulson AR (1977) *Proc Natl Acad Sci USA* **74**: 5463–67.
17 Maniatis T, Fritsch EF, Sambrook J (1982) *Molecular cloning: a laboratory manual.* Cold Spring Harbor, New York, USA: Cold Spring Harbor Laboratory.
18 Van Damme EJM, Smeets K, Torrekens S, Van Leuven F, Goldstein IJ, Peumans WJ (1992) *Eur J Biochem* **206**: 413–20.
19 de Kochko A, Hamon S (1990) *Plant Mol Biol Rep* **8**: 3–7.
20 Dellaporta SL, Wood J, Hicks JB (1983) *Plant Mol Biol Rep* **1**: 19–21.
21 Feldsted RL, Egorin MJ, Leavitt RD, Bachur NR (1977) *J Biol Chem* **252**: 2967–71.
22 von Heijne G (1986) *Nucleic Acids Res* **11**: 4683–90.
23 Van Damme EJM, Barre A, Rougé P, Van Leuven F, Peumans WJ (1995) *Plant Mol Biol* **29**: 1197–210.
24 Van Damme EJM, Barre A, Rougé P, Peumans WJ (1997) *Plant Mol Biol* **33**: 523–36.

Received 10 September 1996, revised 21 November 1996, accepted 30 November 1996