

The Bark of *Robinia pseudoacacia* Contains a Complex Mixture of Lectins¹

Characterization of the Proteins and the cDNA Clones

Els J. M. Van Damme*, Annick Barre, Koen Smeets, Sophie Torrekens, Fred Van Leuven, Pierre Rougé, and Willy J. Peumans

Laboratory for Phytopathology and Plant Protection, Katholieke Universiteit Leuven, Willem de Croylaan 42, 3001 Leuven, Belgium (E.J.M.V.D., K.S., W.J.P.); Laboratory of Cellular Biology, University Paul Sabatier, 35 Chemin des Maraîchers, 31062 Toulouse Cedex, France (A.B., P.R.); and Center for Human Genetics, Katholieke Universiteit Leuven, Herestraat 49, 3001 Leuven, Belgium (S.T., F.V.L.)

Two lectins were isolated from the inner bark of *Robinia pseudoacacia* (black locust). The first (and major) lectin (called RPbAI) is composed of five isolectins that originate from the association of 31.5- and 29-kD polypeptides into tetramers. In contrast, the second (minor) lectin (called RPbAII) is a homotetramer composed of 26-kD subunits. The cDNA clones encoding the polypeptides of RPbAI and RPbAII were isolated and their sequences determined. Apparently all three polypeptides are translated from mRNAs of approximately 1.2 kb. Alignment of the deduced amino acid sequences of the different clones indicates that the 31.5- and 29-kD RPbAI polypeptides show approximately 80% sequence identity and are homologous to the previously reported legume seed lectins, whereas the 26-kD RPbAII polypeptide shows only 33% sequence identity to the previously described legume lectins. Modeling the 31.5-kD subunit of RPbAI predicts that its three-dimensional structure is strongly related to the three-dimensional models that have been determined thus far for a few legume lectins. Southern blot analysis of genomic DNA isolated from *Robinia* has revealed that the *Robinia* bark lectins are the result of the expression of a small family of lectin genes.

Plant lectins are a very heterogeneous group of (glyco)proteins classified together on the basis of a single common property, namely their ability to specifically recognize and bind carbohydrate ligands (Goldstein and Poretz, 1986). One of the most intensively studied classes of these proteins is the legume seed lectins. Because of their

relative high concentrations in readily available seeds of numerous common legumes, most of early lectin research was focused on the study of different aspects of legume seed lectins. As a result, many of these proteins have been characterized in great detail with respect to their biochemical and physicochemical properties and carbohydrate-binding specificity. Moreover, partial or complete chemical sequencing of legume lectins and, more recently, the molecular cloning of some of their genes provided much information about the amino acid sequences of lectin polypeptides from numerous species belonging to almost all taxonomic groups within the legume family (Sharon and Lis, 1990). A critical analysis of the data available at present indicates that in spite of the obvious differences in sugar binding specificity, all legume lectins closely resemble each other with respect to their biochemical/physicochemical properties and, in addition, show an important degree of sequence homology.

Although legume lectins are certainly most widespread in seeds, they are also found in different types of vegetative tissues (such as leaves, stems, and roots) of various species. It has been shown, for instance, that the most prominent protein in the inner bark of *Robinia pseudoacacia* (black locust) is an agglutinin (Nsimba-Lubaki and Peumans, 1986). As early as 1978, Horejsi et al. isolated a lectin from the bark of black locust by specific adsorption on formaldehyde-fixed human erythrocytes. Partial characterization of this lectin indicated that it was a 110-kD protein composed of two types of subunits with molecular masses of 29 and 31.5 kD that exhibited specificity toward *N*-acetylgalactosamine. More recently, Tazaki and Yoshida (1992) purified the black locust bark lectin by sequential ion-exchange chromatography. In contrast to Horejsi et al. (1978), they found that the bark lectin is a homotetrameric protein (106 kD) composed of a single type of 29-kD sub-

¹ Supported in part by grants from the Catholic University of Leuven (OT/94/17), the Land Use and Tenure Division of the Ministry of Agriculture, Fisheries and Food, 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. Senior Research Assistant of this fund. K.S. acknowledges the receipt of a Fellowship of the Belgian Instituut tot Aanmoediging van het Wetenschappelijk Onderzoek in Nijverheid en Landbouw. Furthermore, we want to acknowledge grant 7.0047.90 from the Nationaal Fonds voor Wetenschappelijk Onderzoek-Levenslijn fund.

* Corresponding author; e-mail fiaaa32@blekul11.bitnet; fax 32-16-322976.

Abbreviations: HCA, hydrophobic cluster analysis; LECRPA, cDNA clone encoding *Robinia pseudoacacia* lectin; LoLI, *Lathyrus ochrus* isolectin I; PHA, *Phaseolus vulgaris* agglutinin; RPbAI, *Robinia pseudoacacia* bark agglutinin I; RPbAII, *Robinia pseudoacacia* bark agglutinin II.

units. A cDNA clone encoding the 29-kD lectin polypeptide was isolated by the same research group using immunological screening of a cDNA library from *Robinia* bark (Yoshida et al., 1994).

To explain the apparent discrepancies in the literature concerning the molecular structure of the black locust bark lectin, this issue was reinvestigated by a detailed analysis of both the lectins and the corresponding cDNA clones. This paper describes the characterization and molecular cloning of two different lectins from *R. pseudoacacia* bark, called RPbAI and RPbAII. Evidence is presented that RPbAI is composed of two homologous 29- and 31.5-kD lectin polypeptides that associate in all possible combinations into five different tetrameric isolectins (similar to the isoforms of the *Phaseolus vulgaris* agglutinin). RPbAII is a homotetramer of 26-kD subunits, which show considerable sequence similarity to the 29- and 31.5-kD lectin polypeptides, but is devoid of agglutination activity. Isolation and sequencing of the cDNA clones encoding each of the three polypeptides composing the *Robinia* lectins confirmed our biochemical analyses. In addition, it could be demonstrated that the 31.5- and 29-kD polypeptides of RPbAI exhibit 80% homology with each other, whereas they show only 36% sequence identity with the 26-kD subunits of RPbAII. It appears, therefore, that the lectin composition of the black locust bark is far more complicated than was previously believed.

MATERIALS AND METHODS

Plant Material

Branches of *Robinia pseudoacacia* trees were collected locally during autumn. The inner bark was collected using a knife, taking care to remove the outer, corky bark tissue. This tissue was then stored at -70 or -20°C prior to being utilized.

Extraction and Isolation of the *Robinia* Bark Lectins

The *Robinia* lectins were isolated from partially purified bark extracts by affinity chromatography on immobilized fetuin. Bark tissue was homogenized in 20 mM acetic acid containing 0.2 g/L ascorbic acid (10 mL/g fresh weight) using a Waring blender. The homogenate was squeezed through cheesecloth, adjusted to pH 5.0 (with 1 N NaOH), and centrifuged at 10,000g for 10 min. The resulting supernatant was taken off and filtered through glass wool in order to remove the floating particles. After adding 1.5 g/L CaCl_2 , the extract was brought to pH 9.0 (with 1 N NaOH) and kept overnight in the cold room. The resulting precipitate was removed by centrifugation (3,000g for 10 min) and the cleared extract was adjusted to pH 3.8 (with 1 N HCl). After standing overnight in the cold room, the extract was cleared by centrifugation (10,000g for 10 min), adjusted to pH 7.0 (with 1 N NaOH), and brought to 1 M ammonium sulfate by adding the solid salt. Then the solution was degassed under vacuum and recentrifuged (10,000g for 10 min). Afterwards, the cleared supernatant was loaded on a column of fetuin-agarose equilibrated with 1.0 M ammonium sulfate (pH 7.0). After passing the extract the column

was washed with 1.0 M ammonium sulfate until the A_{280} fell below 0.01 and the lectin was desorbed with 20 mM diaminopropane (pH 11). Using the procedure described here, all agglutinating activity present in the crude extracts was retained on the column and could be eluted by raising the pH.

Since the lectin preparation obtained by affinity chromatography on the fetuin-agarose column still contains some impurities as well as phenolic compounds, it was further purified by hydrophobic interaction chromatography on a phenyl-Sepharose column. Therefore, the lectin preparation was desorbed from the affinity column was adjusted to pH 7.0 using 1 N HCl and brought to 1.0 M ammonium sulfate by adding the solid salt. After centrifugation at 15,000g for 20 min, the supernatant was loaded on a column of phenyl-Sepharose equilibrated with 1.0 M ammonium sulfate. Impurities were removed from the column by washing with 1.0 M ammonium sulfate and the lectin was desorbed using a linear gradient of 1.0 M ammonium sulfate to water. Fractions were collected and the A_{280} was determined. Fractions containing the lectin were used for further analysis of the lectin by ion-exchange chromatography.

Ion-Exchange Chromatography

Lectin samples were analyzed by ion-exchange chromatography on a Mono-S column (Type HR 5/5 from Pharmacia, Uppsala, Sweden) using a Pharmacia fast protein liquid chromatography system and 25 mM Mes (pH 5.0) as running buffer. After loading the samples dialyzed in 25 mM Mes buffer, the column was washed with 4 mL of the same buffer and the proteins were eluted using a linear gradient (40 mL) of increasing NaCl concentration (0–0.4 M) in this buffer. Peak fractions were collected manually, their agglutination activity was determined, and the lectin polypeptides were analyzed by gel electrophoresis.

Gel Filtration

Gel filtration of purified *Robinia* lectins was done on a Pharmacia Superose 12 column equilibrated with PBS containing 0.1 M Gal. Lectin samples (200 μL containing about 250 μg of pure protein) were loaded on the column and chromatographed at a flow rate of 20 mL/h. Molecular mass reference markers were catalase (240 kD), aldolase (160 kD), BSA (68 kD), ovalbumin (45 kD), chymotrypsinogen (25 kD), and Cyt *c* (12.5 kD).

Electrophoresis

Lectin preparations were analyzed by SDS-PAGE using 12.5 to 25% (w/v) acrylamide gradient gels as described by Laemmli (1970).

Agglutination Assays

Agglutination assays were carried out in small glass tubes in a final volume of 0.1 mL containing 80 μL of a 1% suspension of red blood cells and 20 μL of crude extracts or lectin solutions (each serially diluted with 2-fold incre-

ments). Agglutination was controlled visually after 1 h at room temperature.

The carbohydrate binding specificity of the lectins was determined using some glycoproteins (thyroglobulin, fetuin, asialofetuin, and ovomucoid) and a series of simple sugars. The sugars tested were: Glc, Gal, galactosamine, GlcNAc, N-acetylgalactosamine, Man, lactose, melibiose, Fuc, Ara, amylose, Rib, Fru, trehalose, sorbose, Xyl, Suc, maltose, and sorbitol (all sugars of the D configuration with the exception of L-Fuc and L-sorbose).

Amino Acid Analyses and Carbohydrate Composition

The amino acid composition and the carbohydrate content of the lectins were determined as described previously (Van Damme et al., 1987).

Amino Acid Sequence Analysis

Protein sequencing was conducted on an Applied Biosystems (Foster City, CA) model 477A protein sequencer interfaced with an Applied Biosystems model 120A on-line analyzer.

RNA Isolation

Total cellular RNA was prepared from plant material stored at -80°C essentially as described by Finkelstein and Crouch (1986). Alternatively, small-scale isolation of total RNA was performed according to the protocol described by Wadsworth et al. (1988).

Construction and Screening of cDNA Library

A cDNA library was constructed from total RNA isolated from *Robinia* bark using the cDNA synthesis kit from Pharmacia. cDNA fragments were inserted into the *EcoRI* site of the multifunctional phagemid pT₇T₃ 18U (Pharmacia). The library was propagated in *Escherichia coli* XL1 Blue (Stratagene, La Jolla, CA).

Recombinant lectin clones were screened using the random-primer-labeled cDNA clone encoding PHA-L (Hoffman and Donaldson, 1985) as a probe. In a later stage cDNA clones encoding the *Robinia* lectins were used as probes. Hybridization using the PHA-L cDNA clone was done overnight at 40°C as reported previously (Van Damme et al., 1992a). 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 Pfeiffer (1987) and sequenced by the dideoxy method (Sanger et al., 1977).

Northern Blot

RNA electrophoresis was performed according to Maniatis et al. (1982). Approximately $40\ \mu\text{g}$ of total RNA was denatured in glyoxal and DMSO and separated in a 1.2% (w/v) agarose gel. Following electrophoresis the RNA was transferred to Immobilon N membranes (Millipore, Bedford, MA) and the blot was hybridized using a random-

primer-labeled lectin cDNA insert. Hybridization was performed as reported by Van Damme et al. (1992b). An RNA ladder (0.16–1.77 kb) was used as a marker.

Genomic DNA Analysis

Total DNA from *Robinia* bark was isolated according to the procedures described by Dellaporta et al. (1983) and de Kochko and Hamon (1990). The DNA preparation was treated with RNase in order to remove any contaminating RNA. Approximately $50\ \mu\text{g}$ 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) and hybridized using the ^{32}P -labeled cDNA insert encoding the lectin.

Computer Analyses

DNA sequences were analyzed using programs from PC Gene (Intelligenetics, Mountain View, CA) and Genepro (Riverside Scientific, Seattle, WA).

Modeling of *Robinia* Lectins

The HCA plots (Gaboriaud et al., 1987; Lemesle-Varloot et al., 1990) were generated on a Macintosh LC using the program HCA-Plot2 (Doriane, Paris, France). The HCA plots were used to delineate the structurally conserved regions along the amino acid sequences of RPA1 and the model proteins LoLI and pea lectin.

Molecular modeling of RPA1, the protein encoded by LECRPA1, was carried out on a Silicon Graphics Iris 4D25G station using the programs InsightII, Homology, and Discover (Biosym Technologies, San Diego, CA). The coordinates of pea lectin (Einspahr et al., 1986) and LoLI (Bourne et al., 1990a) were used to build the RPA1 model. As a result, residues 237 to 254 were omitted to build the RPA1 three-dimensional model. The program TurboFrodo (Rousset and Cambillau, Marseille, France) was run on a Silicon Graphics Indigo 3000 to perform the superimposition of the models.

RESULTS

Nomenclature of the *R. pseudoacacia* Bark Lectins

As is demonstrated below, the bark of black locust contains two different lectins. To distinguish these bark lectins from the previously described seed lectin, which is usually referred to as *R. pseudoacacia* agglutinin or RPA, they are indicated as *R. pseudoacacia* bark agglutinins I and II (RPbAI and RPbAII). The composing subunits of 31.5 and 29 kD of RPbAI are referred to as lectin polypeptides a and b, respectively. Lectin polypeptide c refers to the 26-kD subunit of RPbAII. For modeling of RPbAI, the deduced amino acid sequence of a cDNA clone encoding the a polypeptide of RPbAI is used. To avoid any confusion, this deduced sequence is designated as RPA1.

Purification and Biochemical Characterization of the Major Lectin RPbAI

Preliminary experiments with crude extracts from *R. pseudoacacia* bark indicated that the agglutinating factor could be inhibited by high concentrations of *N*-acetylgalactosamine and several glycoproteins such as fetuin. Therefore, the bark lectin was purified by affinity chromatography on a column of fetuin-agarose. Under the conditions used (i.e. in the presence of 1.0 M ammonium sulfate), virtually all initial agglutinating activity was recovered in the protein fractions desorbed from the column with 20 mM diaminopropane. Since SDS-PAGE of the lectin preparations obtained after affinity chromatography indicated the presence of contaminating proteins, hydrophobic interactions chromatography was performed to remove the impurities from the lectin. As shown in Figure 1, SDS-PAGE of the lectin fractions obtained after hydrophobic interactions chromatography on phenyl-Sepharose and elution with a linear gradient from 1.0 M ammonium sulfate to water yielded virtually pure lectin. The overall yield of lectin was about 10 mg/g dry weight, which corresponds to 20% of the total soluble bark protein. Such high yields of lectin could be expected because the bark lectin polypeptides are the most prominent polypeptide bands visible on the gel after SDS-PAGE of a crude extract from bark tissue (Peumans et al., 1986).

It should be mentioned here that we tried to purify the *Robinia* bark lectin from extracts made in PBS on a column of immobilized *N*-acetylgalactosamine as well as on fixed red blood cells as described by Horejsi et al. (1978). How-

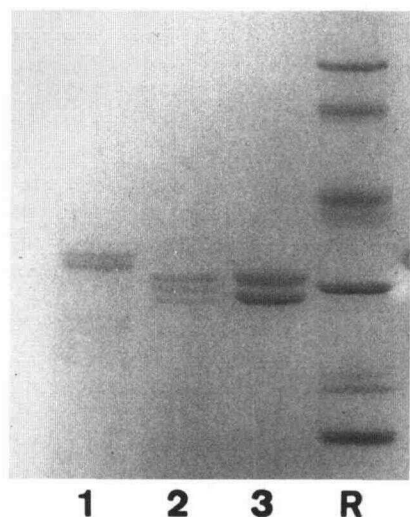


Figure 1. SDS-PAGE of purified lectins from the bark of different *R. pseudoacacia* trees. Lane 3 was loaded with the lectin purified from the crude extract obtained from a tree with a low (normal) level of RPbAII, whereas the lectin preparation in lane 2 was from an individual with a high level of both RPbAI and RPbAII. In lane 1, a sample of partially purified (total) PHA was loaded. About 25 μ g of each lectin was loaded on the gel. Molecular mass reference proteins are shown in lane R: lysozyme (14 kD), soybean trypsin inhibitor (20 kD), carbonic anhydrase (30 kD), ovalbumin (43 kD), BSA (67 kD), and phosphorylase b (94 kD).

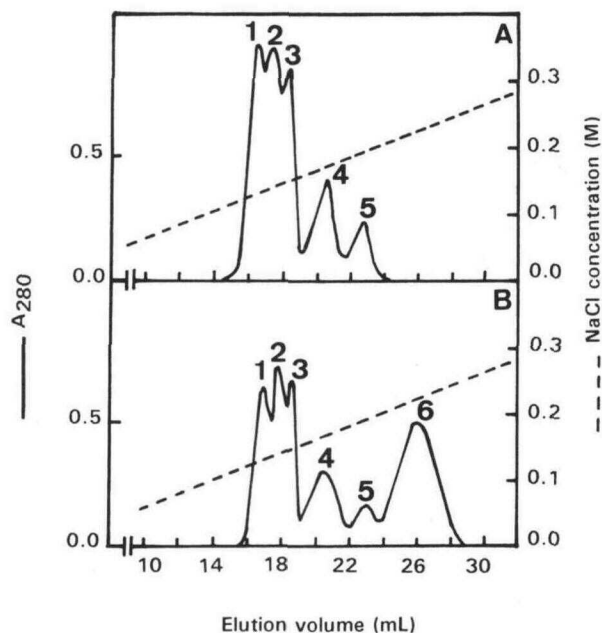


Figure 2. Ion-exchange chromatography of the bark lectins from *R. pseudoacacia* on a Pharmacia Mono-S column. About 1 mg of the purified lectin (dissolved in 25 mM Mes buffer, pH 5.0) was chromatographed on a Mono-S column (Pharmacia, type HR5/5) equilibrated with 25 mM Mes, pH 5.0, and eluted using a linear gradient (40 mL) of increasing NaCl concentration (0–0.4 M) in the same buffer. The flow rate was 2 mL/min. Patterns shown in A (RPbAI) and B (RPbAI + RPbAII) were obtained with the lectin preparations shown in lanes 2 and 3, respectively, of Figure 1.

ever, these attempts were not very successful in that the lectin was not retained on the *N*-acetylgalactosamine column and did not firmly bind to the red blood cells. The use of 1.0 M ammonium sulfate during the affinity chromatography of the extract on fetuin-agarose offers a double advantage: first, the enzyme activity in the bark extracts is drastically reduced, which prolongs the lifetime of the column, and second, the presence of ammonium sulfate increases the affinity of the lectin for the column.

The molecular structure of the *Robinia* lectin was determined by SDS-PAGE and gel filtration. Gel electrophoresis revealed two polypeptide bands of 29 and 31.5 kD (Fig. 1), whereas the native lectin eluted from the Superose 12 column in a single symmetrical peak with an apparent molecular mass of 120 kD (results not shown). It can be concluded, therefore, that RPbAI is a tetramer composed of 31.5- (a) and 29-kD (b) subunits.

Ion-exchange chromatography of RPbAI on a Pharmacia Mono-S column revealed the occurrence of five different isolectins (Fig. 2). Through SDS-PAGE, we traced the origin of the isolectins. Whereas isolectins 1 and 5 contain exclusively a and b subunits, respectively, isolectins 2, 3, and 4 contain both a and b polypeptides in different ratios (Fig. 3). Since all (native) isolectins elute with an apparent molecular mass of 120 kD from the Superose 12 gel filtration column, we assume that they are tetramers. It can be concluded, therefore, that RPbAI is composed of 31.5-kD (a)

and 29-kD (b) subunits that associate into tetramers in all five possible combinations (i.e. a_4 , a_3b_1 , a_2b_2 , a_1b_3 , and b_4). In this respect, the isolectin composition of RPbAI is reminiscent of that of PHA, which is also a mixture of five isolectins that find their origin in the association E and L subunits into tetramers (thereby giving rise to isolectins L_4 , L_3E_1 , L_2E_2 , L_1E_3 , and E_4) (Feldsted et al., 1977; Leavitt et al., 1977).

The different isoforms of the RPbAI differ also in their agglutination properties. Agglutination assays indicated that isolectin a_4 is completely devoid of agglutination activity, whereas all other isoforms exhibit lectin activity. Estimations of the specific agglutination activity (which is expressed as the lowest concentration at which agglutination of trypsin-treated rabbit erythrocytes still occurs) further demonstrated that isolectin b_4 has the highest (0.2 $\mu\text{g}/\text{mL}$) and isolectin a_3b_1 has the lowest (4.0 $\mu\text{g}/\text{mL}$) activity. As could be expected, the activity of isolectins a_2b_2 and a_1b_3 (1.6 and 0.8 $\mu\text{g}/\text{mL}$, respectively) is intermediate between that of isoforms b_4 and a_3b_1 . The differences in specific agglutination activity of the RPbAI isolectins are reminiscent of the well-known differential hemagglutinating activity of the PHA isolectins (Leavitt et al., 1977). It is striking, indeed, that the 31.5-kD subunit confers no agglutination activity to the lectin (like the L subunit of PHA), whereas an increasing ratio of the 29-kD subunit results in strongly enhanced lectin activity (like the E subunit of PHA).

The carbohydrate binding specificity of (total) RPbAI was assessed by hapten inhibition assays using a whole series of simple sugars and some glycoproteins. Of all monosaccharides tested only *N*-acetylgalactosamine gave an inhibition of the agglutination (Table I). Assays with

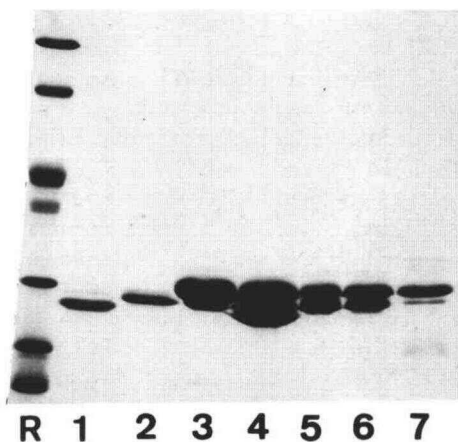


Figure 3. SDS-PAGE of different isolectins separated by ion-exchange chromatography of the *Robinia* bark lectins on a Mono-S column (Fig. 2). Samples are loaded as follows: lane 1, peak 6 (c_4); lane 2, peak 5 (b_4); lane 3, peak 2 (a_3b_1); lane 4, total lectin preparation containing 31.5- (a), 29- (b) and 26-kD (c) lectin polypeptides; lane 5, peak 4 (a_1b_3); lane 6, peak 3 (a_2b_2); lane 7, peak 1 (a_4). Molecular mass markers are shown in lane R; they are the same as in Figure 1.

Table I. Carbohydrate binding specificity of the *R. pseudoacacia* bark lectin RPbAI

The carbohydrate binding specificity of the *Robinia* lectin RPbAI was determined by hapten inhibition assays using a series of simple sugars and some glycoproteins. The following sugars were not inhibitory at a final concentration of 100 mM: Man, Glc, Gal, galactosamine, GlcNAc, lactose, melibiose, Fuc, Ara, Rib, Fru, trehalose, sorbose, Xyl, Suc, maltose, and sorbitol.

Sugar/Glycoprotein	Concentration Required for 50% Inhibition of the Agglutination of Trypsin-Treated Rabbit Erythrocytes in the Presence of RPbAI (5 $\mu\text{g}/\text{mL}$)
<i>N</i> -Acetylgalactosamine	25 mM
Fetuin	500 $\mu\text{g}/\text{mL}$
Asialofetuin	500 $\mu\text{g}/\text{mL}$
Ovomucoid	30 $\mu\text{g}/\text{mL}$
Thyroglobulin	15 $\mu\text{g}/\text{mL}$

some animal glycoproteins showed that the agglutination activity of RPbAI was inhibited by low concentrations of ovomucoid and thyroglobulin (Table I).

Occurrence of a Second Lectin in *R. pseudoacacia* Bark

Analysis of purified lectin preparations from bark of different *Robinia* trees revealed that in some cases a third 26-kD polypeptide co-purified with the RPbAI lectin polypeptides of 29 and 31.5 kD. SDS-PAGE of a crude extract from the bark of a tree that yielded such a lectin preparation revealed, in addition to the 29- and 31.5-kD lectin polypeptides, a third major polypeptide band of 26 kD (results not shown). Since these results were indicative of the occurrence of two lectins in some *Robinia* trees, a lectin preparation from such a tree (which contains roughly equal amounts of the a and b polypeptides of RPbAI and the unidentified 26-kD polypeptide [Fig. 1]) was further analyzed. Ion-exchange chromatography on a Mono-S column demonstrated that this lectin preparation yielded one additional peak in comparison to pure RPbAI (Fig. 2). Since the additional peak eluted well behind the isoforms of RPbAI, it could readily be purified in reasonable amounts for further analysis. SDS-PAGE of the resulting protein demonstrated that it contained exclusively 26-kD polypeptides, whereas gel filtration on a Superose column indicated that it eluted at virtually the same position as RPbAI. It can be concluded, therefore, that this protein (which is referred to as the second *R. pseudoacacia* bark lectin, or RPbAII) is a homotetramer composed of four identical 26-kD subunits (referred to as the type c lectin polypeptides).

When tested with trypsin-treated rabbit or human erythrocytes, the RPbAII showed no agglutination activity (just like isolectin a_4 of RPbAI). However, since it co-purifies with RPbAI upon affinity chromatography on immobilized fetuin, the carbohydrate-binding activity of RPbAII is evident.

Carbohydrate Content of the *Robinia* Lectin Polypeptides

Determinations of the total carbohydrate content of RPbAI yielded values of 0.4 mol% GlcN and 2.3 mol% Man,

indicating that on average each native (tetrameric) lectin molecule contains a single oligosaccharide side chain. In an attempt to determine which lectin polypeptides are glycosylated, lectin preparations were treated with endoglycosidase H or blotted and assayed for carbohydrate using the Boehringer DIG glycan detection kit. Overnight treatment of RPbAI with endoglycosidase H revealed no visible shift in molecular mass of any of the lectin polypeptides compared to the control. With the DIG glycan detection kit, a weak staining of the a polypeptide was obtained (results not shown). However, the intensity of this staining was much lower than that of the E and L subunits of PHA (each of which contain two carbohydrate chains [Sturm and Chrispeels, 1986]). Taken together, these results indicate that possibly one carbohydrate side chain is present on the a polypeptide. Most likely, however, only a fraction of the a polypeptides is glycosylated since (a) the DIG staining is very weak and (b) no shift in molecular mass was seen after treatment with endoglycosidase H.

Since estimations of the total carbohydrate content of RPbAII (using the phenol sulfuric acid method) and staining of the c lectin polypeptide with the DIG glycan detection kit yielded negative results, we assume that this protein is not glycosylated.

N-Terminal Sequences of the *Robinia* Lectin Polypeptides

To determine the N-terminal sequences of the *Robinia* lectin polypeptides, the lectin was run on an SDS polyacrylamide gel and blotted onto a nitrocellulose filter. After staining the blot, the lectin polypeptides were excised and used for automated amino acid sequencing. As shown in Figure 4, highly homologous N-terminal sequences were obtained for the a and b polypeptides of RPbAI, whereas the c polypeptide of RPbAII yielded an N-terminal sequence that shows almost no sequence homology to the a and b polypeptides, at least not within the first 20 amino acids.

Molecular Cloning of the Lectins from *R. pseudoacacia*

Previous studies on the seasonal changes of the lectin content of *Robinia* bark have shown that the level of this protein and its mRNA drastically increases during autumn (Nsimba-Lubaki and Peumans, 1986; Yoshida et al., 1994). Therefore, a cDNA library was constructed from total RNA isolated from *Robinia* bark tissue collected during this season. Since legume lectins exhibit a considerable sequence homology (Sharon and Lis, 1990), initial screenings of the cDNA library were performed using the random-primer-labeled cDNA clone encoding PHA-L (Hoffman and

Donaldson, 1985). In a later stage cDNA clones encoding the *Robinia* lectins were used to screen the cDNA library.

Multiple lectin cDNA clones were isolated from the cDNA library and their sequences were analyzed. Three classes of lectin cDNA clones corresponding to the a and b polypeptides of RPbAI and the c polypeptide of RPbAII were identified. The complete nucleotide sequences and deduced amino acid sequences of a representative of each class were determined and compared. As shown in Figures 5 and 6, the lectin cDNA clones LECRPA1, LECRPA2, and LECRPA3 clearly differ from each other at some positions in their nucleotide sequences (Fig. 5) as well as in their deduced amino acid sequences (Fig. 6).

A first cDNA clone, LECRPA1, contains an 856-bp open reading frame encoding a 285-amino acid precursor with one possible initiation codon at position 1 of the deduced amino acid sequence. Translation starting with this Met residue results in a lectin precursor with a calculated molecular mass of 30,928 D, which after co-translational cleavage of the signal peptide of 31 amino acids yields a lectin precursor polypeptide of 27,330 D with an N-terminal amino acid sequence identical to the one determined for the a polypeptide of RPbAI. The estimated pI for the lectin polypeptide encoded by LECRPA1 is 5.04. The deduced amino acid sequence of the lectin cDNA clone LECRPA1 contains two putative glycosylation sites at positions 147 and 188.

The second cDNA clone LECRPA2 encodes a 286-amino acid precursor with a calculated molecular mass of 31,211 D, which after cleavage of the signal peptide (31 amino acids) is converted into a 27,600-D lectin polypeptide with an N-terminal amino acid sequence similar to the one determined for the b polypeptide of RPbAI and an estimated pI of 4.95. The sequence of this lectin polypeptide contains only one putative N-glycosylation site, the position of which coincides with the first possible glycosylation site in LECRPA1.

The third cDNA clone, LECRPA3, encodes a 272-amino acid precursor with one possible initiation site at position 13. Translation starting at this site yields a lectin precursor with a calculated molecular mass of 27,878 D, which after cleavage of the signal peptide (17 amino acids) is converted into a 25,970-D lectin polypeptide, the N-terminal sequence of which resembles the sequence determined for the c polypeptide of RPbAII. Although the N-terminal sequences determined for the 26-kD lectin polypeptide and the sequence deduced from the cDNA clone LECRPA3 are not identical, LECRPA3 is believed to encode a lectin polypeptide, since it shows a striking sequence homology to LECRPA1 and LECRPA2. Furthermore, the differences in the N-terminal sequences may arise from the expression of different lectin genes (the presence of which can be inferred from Southern blot analysis). The estimated pI of the polypeptide encoded by LECRPA3 (6.5) is higher than the pI of the lectin polypeptides encoded by LECRPA1 and LECRPA2, which is in good agreement with the elution position of this polypeptide from a Mono-S column in Mes buffer, pH 5.0 (Fig. 2). Within the coding sequence of

type a (31.5 kD)	TGSLG FSFPK FAPNQ PYLI
type b (29 kD)	TGSLG FSFPK FMPNE PDL
type c (26 kD)	AEGIS FDFTK FTQSD XITL

Figure 4. N-terminal amino acid sequences of the different *Robinia* bark lectin polypeptides. No amino acid could be determined at position 16 of the 26-kD polypeptide.

Table II. Sequence comparison of the deduced amino acid sequences of LECRPA1 and LECRPA2 toward other legume lectins

The percentage identity (I%) and homology (H%) of the lectin polypeptides RPA1 and RPA2 encoded by LECRPA1 and LECRPA2, respectively, toward other legume lectins was calculated.

Polypeptide		PHA-E	PHA-L	Con A	LoLI	EcoL	GsIV	PNA
RPA1	I%	56.6	56.8	42.9	53.0	49.5	45.2	44.3
	H%	73.3	70.0	67.4	70.7	66.9	63.4	65.5
RPA2	I%	55.9	54.4	44.4	53.9	50.2	48.1	42.9
	H%	69.4	66.8	67.1	69.8	68.0	66.7	66.5

Modeling of the *Robinia* Lectin Polypeptide Encoded by LECRPA1

At this time the three-dimensional structure of only a limited number of legume lectins has been studied in detail. Some of these are the lectins from jack bean (*Canavalia ensiformis*, Con A), garden pea (*Pisum sativum*, PsA), fava bean (*Vicia faba*, favin), *Griffonia simplicifolia* (GsIV), *Erythrina corallodendron* (EcoL), *Lens culinaris* (LcA), and LoLI.

The lectin polypeptide encoded by LECRPA1 (referred to as RPA1) has been modeled from the coordinates of two other closely related lectins: the pea lectin (Einspahr et al., 1986) and LoLI (Bourne et al., 1990a). Thus, it was assumed that RPA1 is a single-chain lectin, which is consistent with the molecular masses of the 29- and 31.5-kD subunits of RPbAI.

The HCA plots of LoLI and RPA1 were very similar, thus indicating that both proteins exhibit very close three-dimensional structures (Fig. 7). In this respect, the 13 anti-parallel strands of β sheet, which form the front (7 strands) and back (6 strands) faces of all the legume lectin monomers resolved so far by x-ray crystallography (Rougé et al.,

1991), were easily delineated along the amino acid sequence of RPA1 by comparison with the LoLI plot. They were subsequently used to build the three-dimensional model of RPA1. For this purpose a limited number of gaps were created and four insertions were introduced.

The three-dimensional model of RPA1 resembles those of other legume seed lectins. Basically, the lectin is built up of 13 strands of β sheet connected by loops and turns, forming the dome-shaped monomer (Fig. 8). As shown in Figure 9, this model is superimposable with only minor differences to those of other legume lectins, such as LoLI, Con A, or EcoL of *E. corallodendron* (Shaanan et al., 1991). Two putative N-glycosylation sites, Asn¹¹⁶-Lys¹¹⁷-Ser¹¹⁸ and Asn¹⁵⁷-Trp¹⁵⁸-Thr¹⁵⁹, occur along the amino acid sequence of RPA1, where they are located in exposed loops. Site Asn¹¹⁶-Lys¹¹⁷-Ser¹¹⁸ is most exposed on the surface of the lectin and thus the most likely to be glycosylated (which is in good agreement with the presence of about 2.3 mol% Man and 0.4 mol% GlcN in RPbAI). Assuming that the amino acid residues forming the carbohydrate site of RPA1 occupy the same position as in LoLI (Bourne et al.,

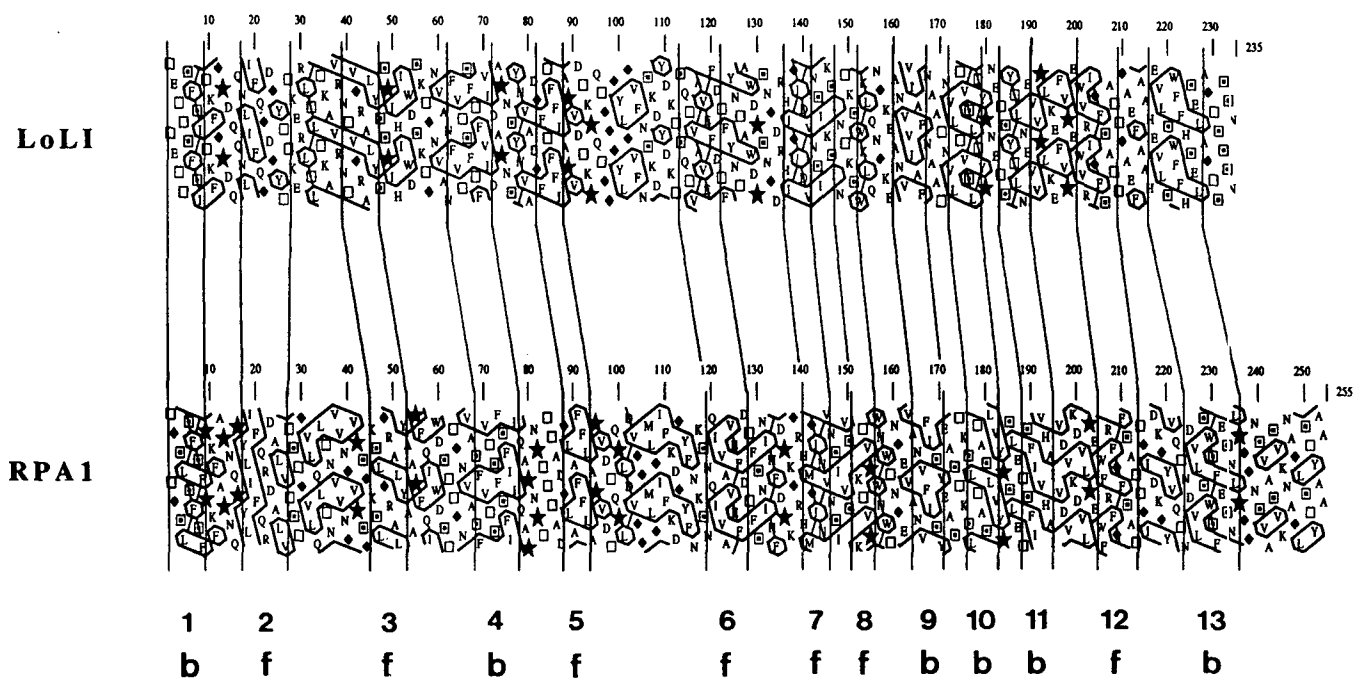


Figure 7. Comparison of the HCA plots of LoLI and RPA1. The 13 strands of β sheet delineated on the LoLI sequence from the three-dimensional model of the protein are easily recognized on the RPA1 sequence. The strands of β sheet forming the front (f) and the back (b) faces of the lectin monomers are numbered from 1 to 13.

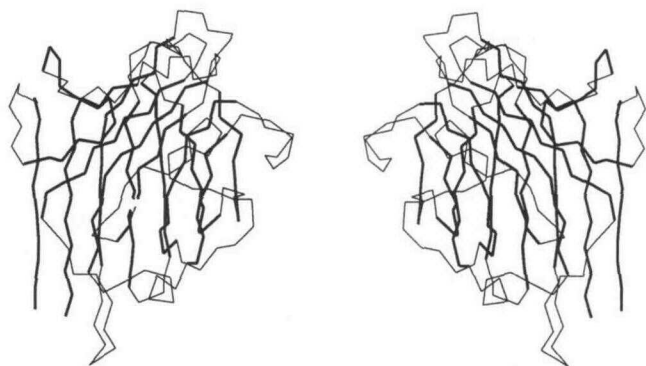


Figure 8. Three-dimensional models of RPA1 showing the front (left) and the back (right) faces of the monomer. Strands of β sheet are in thick lines and loops are in thin lines.

1990b), the site should correspond to residues Asp⁸⁷ (Asp⁸¹ in LoLI), Gly¹⁰⁵ (Gly⁹⁹ in LoLI), Phe¹²⁹ (Phe¹²³ in LoLI), Asn¹³¹ (Asn¹²⁵ in LoLI), Gly²¹⁵ (Gly²⁰⁹ in LoLI), Ile²¹⁶ (Ala²¹⁰ in LoLI), and Asp²¹⁷ (Glu²¹¹ in LoLI). These few amino acid changes could explain the differences in sugar-binding specificities between both lectins.

Northern Blot Analysis

To determine the total length of the lectin mRNAs, a blot containing about 50 μ g of total RNA from *Robinia* bark was hybridized with the labeled lectin cDNA LECRPA1. As shown in Figure 10, hybridization yielded a single band of approximately 1200 nucleotides, which is in good agreement with the length of the cDNA clones isolated from the cDNA library constructed from this RNA. Hybridization of the blot with specific oligonucleotide probes designed for the N-terminal coding sequences of each lectin cDNA clone (Fig. 5) revealed that LECRPA1 and LECRPA2 are expressed at high levels, whereas LECRPA3 yielded only a



Figure 9. Superimposition of the three-dimensional models of RPA1 (thick lines), Con A, LoLI, and EcoL. Strands of β sheet are in thick lines and loops are in thin lines. The overall α -carbon tracing of strands of β sheet of the lectins is very similar, despite some marked differences occurring with a few loops.

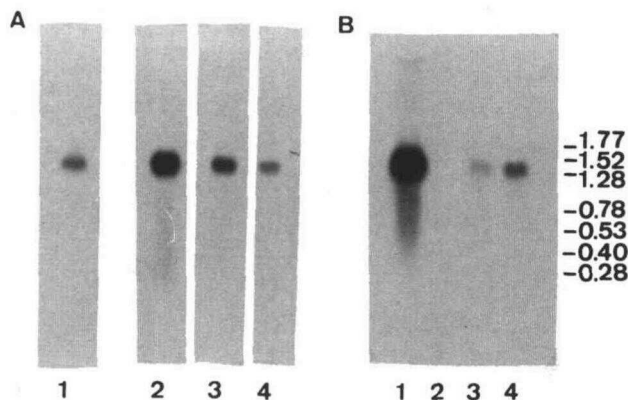


Figure 10. Northern blot of total RNA isolated from *R. pseudoacacia*. A, Total RNA isolated from bark was hybridized with a random-primer labeled cDNA clone encoding LECRPA2 (lane 1) or a specific oligonucleotide probe (see Fig. 5) derived from the N-terminal sequence of LECRPA1 (lane 2), LECRPA2 (lane 3), or LECRPA3 (lane 4). The blots showing lanes 1, 2, and 3 were exposed to x-ray film for 6 h, and the blot showing lane 4 was exposed for 18 h. B, Total RNA was isolated from seeds (lane 1), leaves (lane 2), flowers (lane 3), and bark from roots (lane 4). The blot was hybridized using the ³²P-labeled cDNA insert LECRPA1. About 40 μ g of total RNA was loaded in each lane. Numbers on the right show RNA size (kb).

weak hybridization signal. It is evident, therefore, that the RPBaII is expressed at much lower levels than RPBaI, at least at the time of sampling of the bark.

To study the expression of RPBaI in different tissues a blot containing total RNA from different tissues was hybridized with the labeled lectin clone LECRPA1. Hybridization of the blot revealed a very strong hybridization signal in the seed. Lower levels of lectin mRNAs were also seen in the flower and the bark of the roots, whereas no hybridization signal could be detected in leaf RNA.

Southern Blot Analysis

The occurrence of three different types of lectin polypeptides in *Robinia* and the cloning of their corresponding cDNAs already suggest that at least three different genes are involved in the expression of the lectins in the bark tissue. Since the analysis of different clones of each type of lectin cDNA revealed some minor differences in their sequences, it seemed worthwhile to look for the possible presence of lectin gene families. Therefore, genomic DNA was subjected to Southern blot analysis. As shown in Figure 11, hybridization of a blot containing restriction fragments of genomic DNA from *Robinia* with labeled LECRPA1 revealed a complex pattern of bands indicating that the lectin is encoded by a small family of genes. A similar complex pattern of bands was also obtained when the same blot was hybridized with the labeled lectin cDNA clone LECRPA3, indicating that both lectins are under the control of a family of closely related genes.

DISCUSSION

Although the occurrence of a lectin in the bark of *R. pseudoacacia* has been known for a long time, its molecular

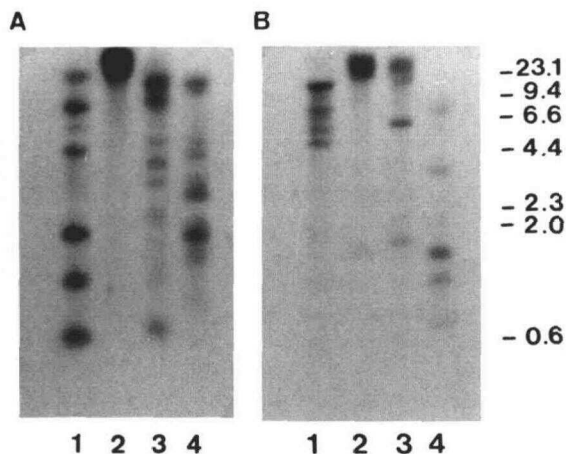


Figure 11. Southern blot of genomic DNA isolated from bark of *R. pseudoacacia*. DNA was digested with *Eco*RI (lane 1), *Pst*I (lane 2), *Bam*HI (lane 3), and *Hind*III (lane 4) and hybridized with the ³²P-labeled cDNA insert LECRPA1 (A) or LECRPA3 (B). Numbers on the right show DNA size (kb).

structure has not been fully elucidated. Indeed, although Horejsi et al. (1978) described the black locust bark lectin as a tetrameric protein composed of 31.5- and 29-kD subunits, more recent studies by Tazaki et al. (1992) and Yoshida et al. (1994) claimed that the same lectin is built up of four identical 29-kD polypeptides.

The results presented here, which are based on a detailed analysis of the carbohydrate-binding proteins and their corresponding cDNA clones, demonstrate that the lectin composition of the black locust bark is far more complicated than one could expect from the data described in the literature. Although part of it can be explained by the presence of two different lectins, the occurrence of multiple isoforms of the major lectin undoubtedly contributes to the apparent complexity of the lectin mixture in *Robinia* bark.

The elucidation of the isolectin composition of RPbAI, which is the result of the association of a and b subunits in all five possible combinations into tetramers, is important for two reasons. First, it gives an explanation for the discrepancies in the literature concerning the molecular structure of the major bark lectin of *Robinia*. It is evident that the lectin described by Horejsi et al. (1978) was a mixture of isolectins, whereas Tazaki et al. (1992) characterized isolectin b₄. Second, the unraveling of the isolectin pattern of RPbAI demonstrates a striking similarity with the isolectins of PHA, not only for what concerns the origin of the isoforms but also for their differential agglutination activity (Feldsted et al., 1977; Leavitt et al., 1977). Analysis of several lectin cDNA clones proved that the a and b polypeptides of RPbAI are encoded by different genes, which show approximately 80% sequence identity. In this respect our results clearly differ from those reported by Yoshida et al. (1994), who found only one type of lectin cDNA (namely that encoding the b subunit).

A comparison of the sequences of the cDNA clones encoding RPbAI shows a high degree of sequence identity to the previously reported legume lectins. In addition, mod-

eling of the 31.5-kD polypeptide predicts that it has a three-dimensional structure that is strongly related to the three-dimensional models determined for a few legume lectins. Since the molecular masses of the deduced amino acid sequences of the mature lectin encoded by the cDNA clones LECRPA1 and LECRPA2 are almost identical (27.3 and 27.6 kD, respectively), the differences in molecular mass between the lectin polypeptides on SDS-PAGE probably reside in a different degree of glycosylation. Taking into consideration the weak staining for carbohydrate of the 31.5-kD a polypeptide, it is very likely that (at least a fraction of) the 31.5-kD a lectin polypeptides carry one carbohydrate side chain, which confirms the results of (a) the molecular modeling of RPA1 and (b) determinations of the carbohydrate content of RPbAI. Alternatively, the subunit heterogeneity of RPbAI may arise from differential posttranslational proteolytic processing. Recent studies of the lectin in leaves and stems of the legume *Dolichos biflorus* have shown that the α and β subunits composing the lectin are derived from a single gene product by differential processing at the COOH terminus (Etzler, 1994).

Analysis of purified lectin preparations and crude extracts of different *Robinia* trees led to the discovery of a second bark lectin, RPbAII. Characterization of RPbAII demonstrated that it consists of a single molecular species of homotetramers of 26-kD subunits. Molecular cloning of the cDNA encoding the 26-kD polypeptide revealed that it shares only 36% sequence identity with the cDNA clones encoding RPbAI. There is no doubt, therefore, that RPbAI and RPbAII are two different proteins that are encoded by different genes.

RPbAII has not been reported yet in the literature. Although its discovery followed the analysis of the total lectin from a tree that contained large quantities of both RPbAI and RPbAII, RPbAII has been found in the bark of many other trees, albeit in much lower concentrations. Indeed, although the 26-kD lectin polypeptide is not visible upon SDS-PAGE of total lectin preparations of most trees, it can easily be distinguished when the same preparations are enriched in RPbAII by ion-exchange chromatography on a Mono-S column (by taking the fractions that elute at the end of the NaCl gradient) (results not shown). Moreover, it should be emphasized here that the cDNA clone encoding LECRPA3 has been isolated from a cDNA library constructed with mRNA from a tree with the total lectin that does not yield a visible 26-kD polypeptide upon SDS-PAGE. It should be mentioned, however, that the cDNA clones encoding LECRPA3 are about 100-fold less abundant than the cDNA clones encoding LECRPA1 and LECRPA2 (as is illustrated by the results of a northern blot, whereby each of the three groups of cDNAs is detected by a specific oligonucleotide probe) (Fig. 10).

Analysis of crude bark extracts from numerous *Robinia* trees indicated that most individuals contain barely detectable levels of RPbAII and that the presence of high concentrations of this lectin is rather exceptional. At present there is no explanation why only a few trees accumulate large quantities of RPbAII except that there must be genetic differences. Irrespective of the underlying mechanisms,

however, the observed differences in lectin patterns emphasize the need for a thorough analysis of numerous individuals within a population of a wild species before general conclusions can be drawn.

ACKNOWLEDGMENT

We are indebted to Dr. M.J. Chrispeels for supplying us with the cDNA clone encoding PHA-L.

Received September 12, 1994; accepted November 28, 1994.

Copyright Clearance Center: 0032-0889/95/107/0833/11.

The GenBank/EMBL accession numbers for the sequences reported in this article are U12782, U12783, and U12784.

LITERATURE CITED

- Bourne Y, Abergel C, Cambillau C, Frey M, Rougé P, Fontecilla-Camps JC (1990a) X-ray crystal structure determination and refinement at 1.9 Å resolution of isolectin I from the seeds of *Lathyrus ochrus*. *J Mol Biol* **214**: 571–584
- Bourne Y, Roussel A, Frey M, Rougé P, Fontecilla-Camps JC, Cambillau C (1990b) Three-dimensional structures of complexes of *Lathyrus ochrus* isolectin I with glucose and mannose: fine specificity of the monosaccharide binding site. *Proteins* **8**: 365–376
- de Kochko A, Hamon S (1990) A rapid and efficient method for the isolation of restrictable total DNA from plants of the genus *Abelmoschus*. *Plant Mol Biol Rep* **8**: 3–7
- Dellaporta SL, Wood J, Hicks JB (1983) A plant DNA miniprep: version II. *Plant Mol Biol Rep* **1**: 19–21
- Einspahr H, Parks EH, Suguna K, Subramanian E, Suddath FL (1986) The crystal structure of pea lectin at 3.0-Å resolution. *J Biol Chem* **261**: 16518–16527
- Etzler ME (1994) Isolation and characterization of subunits of DB58, a lectin from the stems and leaves of *Dolichos biflorus*. *Biochemistry* **33**: 9778–9783
- Feldsted RL, Egorin MJ, Leavitt RD, Bachur NR (1977) Recombinations of subunits of *Phaseolus vulgaris* isolectins. *J Biol Chem* **252**: 2967–2971
- Finkelstein RR, Crouch ML (1986) Rapeseed embryo development in culture on high osmoticum is similar to that in seeds. *Plant Physiol* **81**: 907–912
- Gaboriaud C, Bissery V, Benchetrit T, Mornon JP (1987) Hydrophobic cluster analysis: an efficient new way to compare and analyse amino acid sequences. *FEBS Lett* **224**: 149–155
- Goldstein IJ, Poretz RD (1986) Isolation, physicochemical characterization and carbohydrate-binding specificity of lectins. In IE Liener, N Sharon, IJ Goldstein, eds, *The Lectins: Properties, Functions and Applications in Biology and Medicine*. Academic Press, San Diego, CA, pp 33–248
- Hoffman LM, Donaldson DD (1985) Characterization of two *Phaseolus vulgaris* phytohemagglutinin genes closely linked on the chromosome. *EMBO J* **4**: 883–889
- Horejsi V, Haskovec C, Kocourek K (1978) Studies on lectins. XXXVIII. Isolation and characterization of the lectin from black locust bark (*Robinia pseudoacacia* L.). *Biochim Biophys Acta* **532**: 98–104
- Laemli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T₄. *Nature* **227**: 680–685
- Leavitt RD, Feldsted RL, Bachur NR (1977) Biological and biochemical properties of *Phaseolus vulgaris* isolectins. *J Biol Chem* **252**: 2961–2966
- Lemesle-Varloot L, Henrissat B, Gaboriaud C, Bissery V, Morgat A, Mornon JP (1990) Hydrophobic cluster analysis: procedure to derive structural and functional information from 2-D-representation of protein sequences. *Biochimie* **72**: 555–574
- Maniatis T, Fritsch EF, Sambrook J (1982) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
- Mierendorf RC, Pfeffer D (1987) Direct sequencing of denatured plasmid DNA. *Methods Enzymol* **152**: 556–562
- Nsimba-Lubaki M, Peumans WJ (1986) Seasonal fluctuations of lectins in barks of elderberry (*Sambucus nigra*) and black locust (*Robinia pseudoacacia*). *Plant Physiol* **80**: 747–751
- Peumans WJ, Nsimba-Lubaki M, Broekaert WF, Van Damme EJM (1986) Are bark lectins of elderberry (*Sambucus nigra*) and black locust (*Robinia pseudoacacia*) storage proteins? In LM Shannon, MJ Chrispeels, eds, *Molecular Biology of Seed Storage Proteins and Lectins*. American Society of Plant Physiologists, Rockville, MD, pp 53–63
- Rougé P, Cambillau C, Bourne Y (1991) The three-dimensional structure of legume lectins. In DC Kilpatrick, E Van Driessche, TC Bog-Hansen, eds, *Lectin Reviews*, Vol 1. Sigma Chemical Co., St. Louis, MO, pp 143–159
- Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain terminating inhibitors. *Proc Natl Acad Sci USA* **74**: 5463–5467
- Shaanan B, Lis H, Sharon N (1991) Structure of a legume lectin with an ordered N-linked carbohydrate in complex with lactose. *Science* **254**: 862–866
- Sharon N, Lis H (1990) Legume lectins—a large family of homologous proteins. *FASEB J* **4**: 3198–3208
- Sturm A, Chrispeels MJ (1986) The high mannose oligosaccharide of phytohemagglutinin is attached to Asn¹² and the modified oligosaccharide to Asn⁶⁰. *Plant Physiol* **81**: 320–322
- Tazaki K, Yoshida K (1992) The bark lectin from *Robinia pseudoacacia*: Purification and partial characterization. *Plant Cell Physiol* **33**: 125–129
- Van Damme EJM, Allen AK, Peumans WJ (1987) Leaves of the orchid twayblade (*Listera ovata*) contain a mannose-specific lectin. *Plant Physiol* **85**: 566–569
- Van Damme EJM, Goldstein IJ, Vercammen G, Vuylsteke J, Peumans WJ (1992a) Lectins of members of the Amaryllidaceae are encoded by multigene families which show extensive homology. *Physiol Plant* **86**: 245–252
- Van Damme EJM, Smeets K, Torrekens S, Van Leuven F, Goldstein IJ, Peumans WJ (1992b) The closely related homomeric and heterodimeric mannose-binding lectins from garlic are encoded by one-domain and two-domain lectin genes, respectively. *Eur J Biochem* **206**: 413–420
- Wadsworth GJ, Redinbough MG, Scandalios JG (1988) A procedure for the small-scale isolation of plant RNA suitable for RNA blot analysis. *Anal Biochem* **172**: 279–283
- Yoshida K, Baba K, Yamamoto N, Tazaki K (1994) Cloning of a lectin cDNA and seasonal changes in levels of the lectin and its mRNA in the inner bark of *Robinia pseudoacacia*. *Plant Mol Biol* **25**: 845–853