Glycoconjugate Journal 18, 321–329, 2001

Chemical characterization of the lectin from *Amaranthus leucocarpus syn. hypocondriacus* by 2-D proteome analysis

Pedro Hernández[1], Henri Debray[2], Heidi Jaekel[2], Yonathan Garfias[1], Maria del Carmen Jiménez[1], Salvador Martínez-Cairo[3] and Edgar Zenteno[4]*

[1]*Departamento de Bioquímica, Instituto Nacional de Enfermedades Respiratorias, Tlalpan D.F., 14080, México,* [2]*Laboratoire de Chimie Biologique, UMR du CNRS No. 8576, Université des Sciences et Technologies de Lille, 59655 Villeneuve d'Ascq cedex, France,* [3]*Unidad de Investigación, Centro Médico Nacional S XXI, IMSS, 06725 México,* [4]*Laboratorio de Inmunología, Departamento de Bioquímica, Facultad de Medicina, UNAM, PO Box 70159, México 04510*

In this work, we characterized chemically the *N*-acetyl-D-galactosamine specific lectin from *Amaranthus leucocarpus* syn *hypocondriacus* lectin (ALL). It is a dimeric glycoprotein composed by three isoforms with pI at 4.8, 4.9, and 5.2. Circular dichroism analysis indicated that the secondary structure of ALL contains 45% of β-sheet and 5% of α-helix. Amino acid sequence of the purified lectin and its isoforms was determined from peptides obtained after trypsin digestion by MALDI-TOF (Matrix assisted laser desorption ionization-time of flight). The tryptic peptides prepared from the purified lectin and the three isoforms showed different degrees (80 to 83%) of identity with the amino acid sequence belonging to a previously described high nutritional value protein from *A. hypocondriacus* not shown at the time to be a lectin. Furthermore, analysis of tryptic peptides obtained from ALL previously treated with peptide *N*-glycosidase, revealed a 93% identity with the aforementioned protein. Presence of *N*-glycosidically linked glycans of the oligomannosidic type and, in minor proportion, of the *N*-acetyllactosaminic type glycans was determined by affinity chromatography on immobilized Con A.

Keywords: *Amaranthus leucocarpus*, plant lectins, T and Tn-specific lectin, glycoproteins, proteome analysis, circular dichroism, *Amaranthus* lectins

Introduction

Lectins with specificity for O-glycosidically linked glycans (containing the Galβ1-3 GalNAcα1-0Ser/Thr and GalNAcα1-0Ser/Thr determinants) have been widely used in fractionation of glycoproteins and cellular subpopulations [1–3]. Lectins from *Amaranthaceae* family have been identified in *A. caudatus* [4], *A. cruentus* [5] and *A. leucocarpus* [6]. All of these lectins have a high proportion of essential amino acids such as lysine, leucine, threonine, phenylalanine, valine and sulfur amino acids and interact specifically with GalNAc (*N*-acetyl-D-galactosamine) [4–6]. These lectins have been proved to be useful tools to recognize O-glycosidically linked glycans in different tissues, tumors, and cells [7–9]. *A. leucocarpus* syn *hypocondriacus* is a Mexican representative of the *Amaranthaceae* family. It possesses a lectin that induces immunosuppression in animals [8] and recognizes non-stimulated murine peritoneal macrophages, and naive $CD4^+$ mouse and human lymphocytes [10,11]. The lectin from *A. leucocarpus* (ALL) agglutinates human erythrocytes with the M phenotype [7] and interacts with O-glycosidically linked proteins in some neurodegenerative processes [12].

It has recently been suggested that ALL may recognize a particular conformation of some O-glycan determinants adjacent to the peptidic backbone of the specific cellular receptor, conferring the lectin its particular biological properties [13]. In this work, we determined the main characteristics of the subunits that compose this lectin and the amino acid sequence of each isoform was analyzed from peptides obtained by trypsin digestion by MALDI-TOF (Matrix assisted laser desorption ionization-time of flight).

*To whom correspondence should be addressed: E. Zenteno. Dep. Bioquímica, Facultad de Medicina UNAM, P.O. Box 70159, 04510, México D.F. Fax 52 (5) 616-24-19; E-mail: ezenteno@servidor.unam.mx

Materials and methods

Lectin extraction

Amaranthus leucocarpus syn *hypocondriacus* seeds were obtained in Tulyehualco (Mexico) and the lectin was purified by affinity chromatography on a column containing human erythrocyte type O^M membranes physically entrapped in Sephadex G-25 (Sigma Fine Chemical Co., St. Louis, MO, USA) [6].

Enzymatic treatment of ALL with PNGase

ALL (500 µg) in 150 µl of 30 mM phosphate buffer, pH 8.4, was boiled 5 min at 100°C with 0.2% SDS and 0.5% β-mercaptoethanol. After cooling at room temperature, 4 µl Triton X-100/mg protein and 300 µl of distilled water/mg protein were added. The treated sample was incubated overnight at 37°C with 80 mU/mg protein of *N*-peptide glycosidase F (PNGase) from *Flavobacterium meningosepticum* (provided by the late Dr. A. L. Tarentino, Wadsworth Center for Laboratories and Research, Albany, USA) [14]. The mixture containing deglycosylated lectin was electrophoresed in SDS-PAGE for trypsin treatment (see further).

Purification of lectin glycopeptides

ALL glycopeptides (ALLgp) were prepared by incubating 10 mg lectin in 0.5 ml of 150 mM Tris-acetate pH 8 and 15 mM $CaCl_2$ with 1 mg pronase (from *Streptomyces griseus*, Sigma) for 48 h at 36°C. The mixture was boiled for 5 min to stop the enzymatic reaction and the glycopeptides were filtered on a column containing Bio-Gel P-2 (100 × 1.2 cm ID) equilibrated with distilled water. The glycopeptides were then lyophilized until use [7]. Purification of ALL glycopeptides was performed by affinity chromatography on a Con A-Sepharose column (10 × 1 cm) (Sigma), containing 10 mg of Con A/ml Sepharose. ALLgp (200 µg) in acetate buffer (5 mM sodium acetate, 0.14 M NaCl, and 5 mM $CaCl_2$, pH 5.2) were deposed onto the column. The unretained fraction was eluted with the acetate buffer; addition of 200 mM α-methyl-mannoside eluted the retained fraction. Each collected fraction was desalted on a Biogel P-2 column equilibrated with distilled water and lyophilized. ALL glycopeptides, not recognised by ConA, were applied to a column (10 × 1 cm) containing the erythroagglutinin from alubia (*Phaseolus coccineus*, 10 mg/ml Sepharose), under similar conditions as for Con A.

Analytical methods

Protein concentration was determined by the method of Bradford [15], using bovine serum albumin as standard. Carbohydrate composition was determined with the hepta-fluorobutyrate derivatives of O-methyl-glycosides from the lectin and its glycopeptides, obtained after methanolysis in 0.5 M methanol-HCl for 24 h at 80°C; lysine (Sigma) was used as internal standard. The samples were analysed by gas-chromatography using a capillary column (25 × 0.32 mm) of 5% Silicone OV 210, (Applied Science Lab., Buffalo, NY), in a Varian 2100 gas chromatograph (Orsay, France), as described by Zanetta *et al.* [16]. Two-dimensional electrophoresis (2-D) of the purified ALL was performed as follows: For the first dimension, isoelectric focusing was performed with 10 µg of purified ALL in cylindrical 5% polyacrylamide gels (0.2 × 10 cm) containing 2% (v/v) of pH 3.5 to 10 ampholines (Pharmacia) and 1% Triton X-100 [17]. Gels were run at 600 V for 14 h at 4°C. Then, each gel was deposed (SDS-PAGE), for a second dimension, on 12% separating and 4% stacking gel in a slab gel apparatus according to the method of Laemmli [18]. Coomassie brilliant blue R-250 was used for staining. The pH gradient was established after isoelectric focusing in blank gels by slicing the gels into 0.4 cm sections, and eluting the ampholines with distilled water for 12 h and measuring the pH of the solution. The gradient was linear over the pH 4 to 8.5 range.

Amino acid sequencing

Amino acid sequencing of ALL was determined by MALDI-TOF on peptide fragments obtained by trypsin digestion of the purified ALL. After electrophoresis of the purified lectin or of the *N*-peptide glycosidase F-treated lectin on SDS-PAGE, or after 2-D electrophoresis, the bands or spots were excised with a scalpel, and each band or spot containing 200 pM of protein was digested with 0.5 µg trypsin (Promega sequencing grade) in 500 µl ammonium bicarbonate, pH 8.0, at 37°C, for 24 h. The enzyme digest was then evaporated to dryness using a Speed Vac concentrator (Savant Instruments Inc., N.Y) and re-evaporated with water (2 × 100 µl). Samples were prepared by mixing directly onto the target 1 µl of the reaction products (containing 50 pM of peptides) and 1 µl of a 2,5-dihydroxy-benzoic acid matrix (12 mg/ml in acetonitrile/H_2O, 80:20, v/v) and allowing the mixture to crystallize at room temperature. Positive ions of the peptides were measured by MALDI-TOF on a Vision 2000 time-of-flight mass spectrometer (Finnigan MAT, Bremen, Germany) equipped with a 337 nm UV laser. Mass spectra were acquired in reflectron mode under 8 keV acceleration voltage and positive detection. Control assays were performed using trypsin alone to identify self-digested peptide mass and with angiotensin I as standard (Mr 1296.7) [19]. The mass of $[M + H]^+$ ions from peptides produced by tryptic digestion was compared with those obtained from NCBInr (Swiss-Prot 10/04/99) database. Putative glycosylation sites were determined with the Prosite Pattern Search.

Circular dichroism

CD spectrum of ALL was recorded on an AVIV (New York, USA) spectropolarimeter. The protein concentration was 300 µg of ALL/ml of PBS (30 mM sodium phosphate, 0.15 M sodium chloride, pH 7.0) and the scanning range was 180 to 240 nm in 5 mm path length cells. Prediction of secondary structure, using *A. leucocarpus* amino acid

sequence, was made with SOPMA [20], GOR4 [21], and PSIPRED [22] programs.

Hemagglutinating activity and sugar specificity

Human erythrocytes group O^M from healthy donors were obtained from the Central Blood Bank, IMSS, Mexico. Hemagglutinating activity, using 25 µl of lectin in phosphate buffered saline (PBS: 0.01 M sodium phosphate, 0.15 M sodium chloride, pH 7.2), was assayed in microtiter U plates (NUNC, Denmark) by the two-fold serial dilution procedure, and with 25 µl of a 2% (v/v) erythrocyte suspension in PBS. The hemagglutinating titer is reported as the inverse of the last dilution with agglutinating activity. Sugar specificity of the lectin was determined by comparing the inhibitory activity of GalNAc on the hemagglutination induced by the lectin and against human erythrocytes.

Results

Identification of lectin isoforms

The lectin from *Amaranthus leucocarpus* was purified by affinity chromatography using a column containing glutaraldehyde treated stroma from human erythrocytes type O. SDS-PAGE of the purified lectin under non denaturing conditions indicated that it is a 70 kDa protein, whereas, after treatment with ß-mercaptoethanol and boiling for 10 min, the lectin renders a band of 35 kDa, indicating its two subunits composition, as determined by SDS-PAGE (Figure 1). Two-dimensional (2-D) electrophoresis of *A. leucocarpus* lectin confirms that it is a protein of 35 kDa, presenting three main isoforms with a pI of 4.8, 4.9, and 5.2 (further identified as isoform I, II and III, respectively) (Figure 2). Hemagglutinating activity of the purified lectin was confirmed in the presence of human erythrocytes with the M phenotype and 2 µg/ml of ALL were necessary to agglutinate a 2% solution of erythrocytes. The hemagglutinating activity of the lectin was inhibited specifically by 200 mM of GalNAc. Other monosaccharides such as GlcNAc, Gal, Man, L-Fuc, or Neu5Ac at 200 mM concentration did not affect the hemagglutinating activity of the lectin.

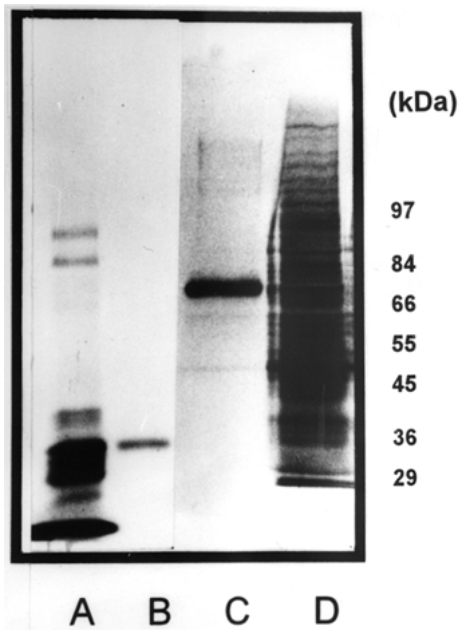

Figure 1. SDS-polyacrylamide electrophoresis of the purified *Amaranthus leucocarpus* lectin: A) *A. leucocarpus* crude extract (60 µg); B) Purified lectin (5 µg) after treatment with ß-mercaptoethanol; C) Purified lectin (5 µg) without reductive treatment. D) *A. leucocarpus* crude extract without reductive treatment. Molecular weight markers: ß-galactosidase *Escherichia coli* (116 kDa), phosphorylase B (97 kDa), fructose-6-phosphate kinase (84 kDa), bovine serum albumin (66 kDa), glutamic dehydrogenase (55 kDa), ovalbumin (45 kDa), glyceraldehyde-3-phosphate dehydrogenase (36 kDa), carbonic anhydrase (29 kDa).

Amino acid sequence

Amino acid sequence of the purified lectin and each isoform was analysed from peptides obtained after trypsin digestion, by MALDI-TOF, and compared with the relative values obtained from the NCBInr (Swiss-Prot 10/01/99) database. Digestion of the purified lectin with trypsin yields 15 peptides with an m/z range of 508.5 to 2810.7 (Table 1), which were

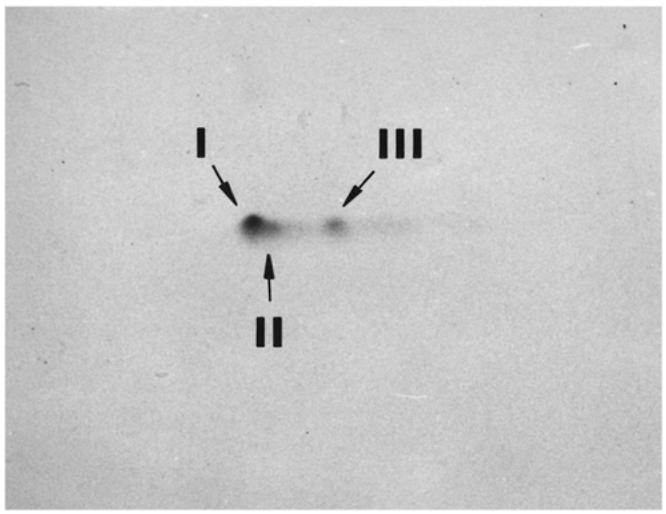

Figure 2. Two-dimensional electrophoresis of the purified *Amaranthus leucocarpus* lectin. Isoelectric focusing (in the first dimension) of 10 µg purified ALL was performed in cylindrical polyacrylamide (at 5% concentration) gels containing 2% (v/v) of pH 3.5 to 10 ampholines and 1% Triton X-100. Gels were run at 600 V for 14 h at 4°C. The gel was deposed for a second dimension on an SDS-PAGE slab gel. The pH gradient was established after isoelectric focusing in blank gels by slicing the gels into 0.4 cm sections, eluting ampholines with distilled water for 12 h and measuring the pH of the solution. The gradient was linear over the pH 4 to 8.5 range, and indicated spots correspond to pH 4.8, 4.9 and 5.2 for ALL-I to III, respectively.

Table 1. Amino acid sequence of the lectin from *Amaranthus leucocarpus* (syn. *hypocondriacus*) determined from tryptic peptides by MALDI-TOF

m/z	*Isoform*	*Residue No**	*Sequence*
1861.2	II, III	2–18	(–)AGLPVIMCLKSNNNQKY
2119.1	I, III	21–38	YQSDNIQQYGLLQFSADK
1586.3	I, II,III	39–52	ILDPLAQFEVEPSK
550.9	III	68–71	YLVR
2259.4	I, II, III	72–90	WSPNHYWITASANEPDENKS
3848.8**		112–146	LLHVQLGYTENYTVGGS FVSYLFAESSQIDTGSK
1159.7	II, III	148–156	DVFHVIDWK
866.8	I, II, III	157–163	SIFQFPK
758.7	I	164–169	TYVTFK
2810.7	III	175–198	YLGVITINQLPCLQ FGYDNLNDPK
1822.5	III	199–214	VAHQMFVTSNGTICIK
508.5	III	221–223	FWR
1715.3	I,II,III	224–238	LSTDNWILVDGNDPR
1121.9	I,II,III	239–248	ETNEAAALFR
1861.2	III	249–264	SDVHDFNVISLLNMQK
694.8	I	265–269	TWFIK

*Amino acid residues are numbered according to the determined sequence of the high nutritional value protein from *A. hypocondriacus* [24]. The fifteen tryptic peptides identified in the isoforms were also identified in the purified lectin.
**This peptide was obtained after deglycosylation and trypsin digestion of the purified lectin.

common for the tryptic peptides obtained from the lectin isoforms. When we compared the m/z values with the relative values obtained from the NCBInr database, an identity of 83% was found with a protein of high nutritional value identified in *Amaranthus hypocondriacus* seeds (AH) and the matched peptides cover 60.5% (184/304 amino acids) of the AH protein (Table 2); however, treatment of the purified lectin with PNGase before trypsin digestion enabled us to identify a new fragment at m/z 3848.8, which corresponds to a 35 amino acid long peptide (Table 1), and the identified peptides revealed now a 93% identity with the aforementioned AH globular protein. The matched peptides, obtained after deglycosylation, covered 72% (219/304 amino acids) of the *A. hypocondriacus* protein amino acid sequence (Table 2). Tables 2 and 3 depict also the three putative *N*-glycosylation sites, determined with the Prosite Pattern Search, present at 89, 123, and 207 amino acid positions. Analysis of the peptides obtained by trypsin digestion of *A. leucocarpus* isoforms, separated after 2-D electrophoresis, indicated relevant differences in the digestion profile and in the degree of identity with the *A. hypocondriacus* protein. ALL-I and II isoforms yielded 8 and 7 main peptides, respectively, and ALL III yielded 13 peptides. As for *A. leucocarpus* lectin, the m/z of the tryptic peptides prepared from the isoforms were identified at 508.5 to 2810.7; however, as indicated in Tables 1 and 3, some fragments were common for the three isoforms. Tryptic peptides obtained from the lectin isoforms showed identity with *A. hypocondriacus* protein in different proportions: ALL-I and III isoforms showed 83% identity and ALL-II showed 80% with *A. hypocondriacus* protein. As showed in the clustal alignment indicated in Table 3, the three isoforms presented common peptides after tryptic digestion; however, some peptides were observed in a specific isoform, *e.g.,* peptide 164–169 was observed in isoform I, peptide 21–38 was specific for isoforms I and III, 2–18 for isoforms II and III, and peptides 68–71 and 175–198 for isoform III.

Circular dichroism

The CD spectrum of *A. leucocarpus* lectin was characterized by a maximum negative absorption at 198–202 nm and positive absorption at 185 nm in the far UV region (Figure 3). Analysis of these spectra with PROCEC 3.0 program revealed that the secondary structure of ALL contains 45% of β-sheet and 5% of α-helix. Similar results were obtained using SOPMA (35% of β-sheet and 7% of α-helix), GOR4 (40% of β-sheet and 7% of α-helix), and PSIPRED (39% of β-sheet and 6% of α-helix) programs.

Carbohydrate composition of ALL

The monosaccharide analysis indicated that the purified lectin as well as the glycopeptides obtained by pronase digestion contained galactose, mannose, and *N*-acetyl-D-glucosamine; in minor concentration, xylose residues were also identified. *N*-acetyl-D-galactosamine, fucose, rhamnose, or glucuronic acid were not identified in any of the fractions tested (Table 4).

Table 2. Identity of the amino acid sequence of the lectin *Amaranthus leucocarpus* (ALL) and the high nutritional value protein from *Amaranthus hypocondriacus* (AH)

```
                          10         20         30         40         50
                           |          |          |          |          |
1 AH          MAGLPVIMCLKSNNNQKYLRYQSDNIQQYGLLQFSADKILDPLAQFEVEP
2 ALL         *AGLPVIMCLKSNNNQKY---QSDNIQQYGLLQFSADKILDPLAQFEVEP
                          60         70         80         90        100
                           |          |          |          |          |
1 AH          SKTYDGLVHIKSRYTNKYLVRWSPNHYWITASANEPDENKSNWACTLFKP
2 ALL         SK---------------YLVRWSPNHYWITASANEPDENKS---------
                         110        120        130        140        150
                           |          |          |          |          |
1 AH          LYVEEGNMKKVRLLHVQLGYTENYTVGGSFVSYLFAESSQIDTGSKDVFH
2 ALL         ------------LLHVQLGYTENYTVGGSFVSYLFAESSQIDTGSKDVFH
                         160        170        180        190        200
                           |          |          |          |          |
1 AH          VIDWKSIFQFPKTYVTFKGNNGKYLGVITINQLPCLQFGYDNLNDPKVAH
2 ALL         VIDWKSIFQFPKTYVTFK-----YLGVITINQLPCLQFGYDNLNDPKVAH
                         210        220        230        240        250
                           |          |          |          |          |
1 AH          QMFVTSNGTICIKSNYMNKFWRLSTDNWILVDGNDPRETNEAAALFRSDV
2 ALL         QMFVTSNGTICIK------FWRLSTDNWILVDGNDPRETNEAAALFRSDV
                         260        270        280        290        300
                           |          |          |          |          |
1 AH          HDFNVISLLNMQKTWFIKRFTSGKPEFINCMNAATQIVDETAILEIIELG
2 ALL         HDFNVISLLNMQKTWFIK--------------------------------

1 AH          SNN
2 ALL         ---
```

AH corresponds to the determined sequence of the high nutritional value protein from *A. hypocondriacus* [24]. In bold letters, the amino sequence of ALL matching with A. hypocondriacus protein. In cursive letters, the amino acid sequence identified after deglycosylation and trypsin treatment. *The NH_2-terminal end of ALL is blocked. Putative *N*-glycosylation sites (underlined) were determined with the Prosite Pattern Search.

The sugar composition of the glycopeptides purified by affinity chromatography on ConA-Sepharose was Man and GlcNAc. The glycopeptides that were not retained on the ConA-Sepharose column were chromatographed on alubia erythroagglutinin-Sepharose. With this column, we identified a slight and retarded interaction with ALL glycopeptides; most of the glycopeptides were obtained practically in the void volume, and this fraction contained Man, GlcNAc, Gal, and Xylose residues (Table 4).

Discussion

Lectins from the *Amaranthaceae* family, such as *A. caudatus* (ACA) and *A. cruentus,* have been shown to be composed of tightly associated homodimers of 33 kDa subunits [3,5,23]. The lectin from *A. leucocarpus* is a dimeric glycoprotein of 35 kDa subunits, composed of three main isoforms with a pI ranging from 4.8 to 5.2, as demonstrated by 2-D electrophoresis. The hemagglutinating activity and sugar specificity for GalNAc of the three isoforms is identical, indicating that the active site of the lectin seems to be preserved among the three isoforms.

As demonstrated previously, the amino terminal in *A. leucocarpus* lectin is blocked [13]. In this work, we determined the amino acid sequence of the purified lectin, as well as of each isoform, by analyzing the peptides obtained through trypsin digestion by MALDI-TOF and comparing them with the relative values obtained from the NCBInr database (Swiss-Prot 10/01/99). Our results indicate that the tryptic peptides prepared from purified lectin and the three isoforms show different degree (80 to 83%) of identity with the amino acid sequence of the high nutritional value globular protein identified in *A. hypocondriacus* [24]. Although the nutritional value of the 35 kDa protein identified in *A. hypocondriacus* has been cited only for its amount of essential amino acids [24], previous works showed that removal of the lectin does not affect the nutritional value of *Amaranthus* seed meal [25,26]. The identified amino acid sequence for the *Amaranthus hypocondriacus* globular protein has not been associated with any functional activity [24], but the present work shows that this identified protein is in fact a lectin.

Lectins from *Amaranthaceae* share the capacity to interact with GalNAc [4,5]; however, ALL shows slight differences in

Table 3. Clustal alignment of the tryptic peptides obtained from the three *A. leucocarpus* lectin isoforms

```
                                  10        20        30        40        50
                                   |         |         |         |         |
1 AH                    MAGLPVIMCLKSNNNQKYLRYQSDNIQQYGLLQFSADKILDPLAQFEVEP
2 ALLI                  --------------------YQSDNIQQYGLLQFSADKILDPLAQFEVEP
3 ALLII                 -AGLPVIMCLKSNNNQKY--------------------ILDPLAQFEVEP
4 ALLIII                -AGLPVIMCLKSNNNQKY--YQSDNIQQYGLLQFSADKILDPLAQFEVEP
                                  60        70        80        90       100
                                   |         |         |         |         |
1 AH                    SKTYDGLVHIKSRYTNKYLVRWSPNHYWITASANEPDENKSNWACTLFKP
2 ALLI                  SK--------------------WSPNHYWITASANEPDENKS---------
3 ALLII                 SK--------------------WSPNHYWITASANEPDENKS---------
4 ALLIII                SK----------------YLVRWSPNHYWITASANEPDENKS---------
                               110       120       130       140       150
                                   |         |         |         |         |
1 AH                    LYVEEGNMKKVRLLHVQLGYTENYTVGGSFVSYLFAESSQIDTGSKDVFH
2 ALLI                  ------------------------------------------------
3 ALLII                 ------------------------------------------------DVFH
4 ALLIII                ------------------------------------------------DVFH
                               160       170       180       190       200
                                   |         |         |         |         |
1 AH                    VIDWKSIFQFPKTYVTFKGNNGKYLGVITINQLPCLQFGYDNLNDPKVAH
2 ALLI                  -----SIFQFPKTYVTFK---------------------------------
3 ALLII                 VIDWKSIFQFPK-------------------------------------
4 ALLIII                VIDWKSIFQFPK-----------YLGVITINQLPCLQFGYDNLNDPKVAH
                               210       220       230       240       250
                                   |         |         |         |         |
1 AH                    QMFVTSNGTICIKSNYMNKFWRLSTDNWILVDGNDPRETNEAAALFRSDV
2 ALLI                  ----------------------LSTDNWILVDGNDPRETNEAAALFR---
3 ALLII                 ----------------------LSTDNWILVDGNDPRETNEAAALFR---
4 ALLIII                QMFVTSNGTICIK------FWRLSTDNWILVDGNDPRETNEAAALFRSDV
                               260       270       280       290       300
                                   |         |         |         |         |
1 AH                    HDFNVISLLNMQKTWFIKRFTSGKPEFINCMNAATQIVDETAILEIIELG
2 ALLI                  -------------TWFIK--------------------------------
3 ALLII                 ------------------------------------------------
4 ALLIII                HDFNVISLLNMQK------------------------------------

1 AH                    SNN
2 ALLI                  ---
3 ALLII                 ---
4 ALLIII                ---
```

AH corresponds to the determined sequence of the high nutritional value protein from *A. hypocondriacus*. Putative glycosylation sites are underlined.

the recognition of oligosaccharidic sequences. The lectin from *A. caudatus* and ALL can bind T and Tn antigens, and substitutions in these antigens of their Gal C3 and GalNAc C6 by neuraminic acid are well tolerated by both lectins; however, in contrast to ACA, substitutions by *N*-acetyl lactosaminyl groups at the same positions in the typically class II core of O-glycosidically linked glycans are not tolerated by ALL [7]. These results indicate that the specificity of ALL is directed toward GalNAc present in the inner core of class I O-glycosidically linked glycans and not to the middle of a longer oligosaccharide. The ability of lectins to bind carbohydrate structures found in glycoconjugates depends on their 3-D structure [27,28]. Secondary structure of *A. leucocarpus* lectin, as determined by Circular Dichroism, is rich in β-sheet (45%) along with 5% α-helix; however, *A. caudatus* lectin possesses 40 and 2% of β-sheet and α-helix motifs (Protein Data Bank, version 1999). Considering that both lectins are very similar in their secondary structures, although not identical, our results suggest that the 3-D structure of ALL and, most probably the sugar binding site, are influenced by the presence of *N*-linked glycans; this could explain the subtle differences in

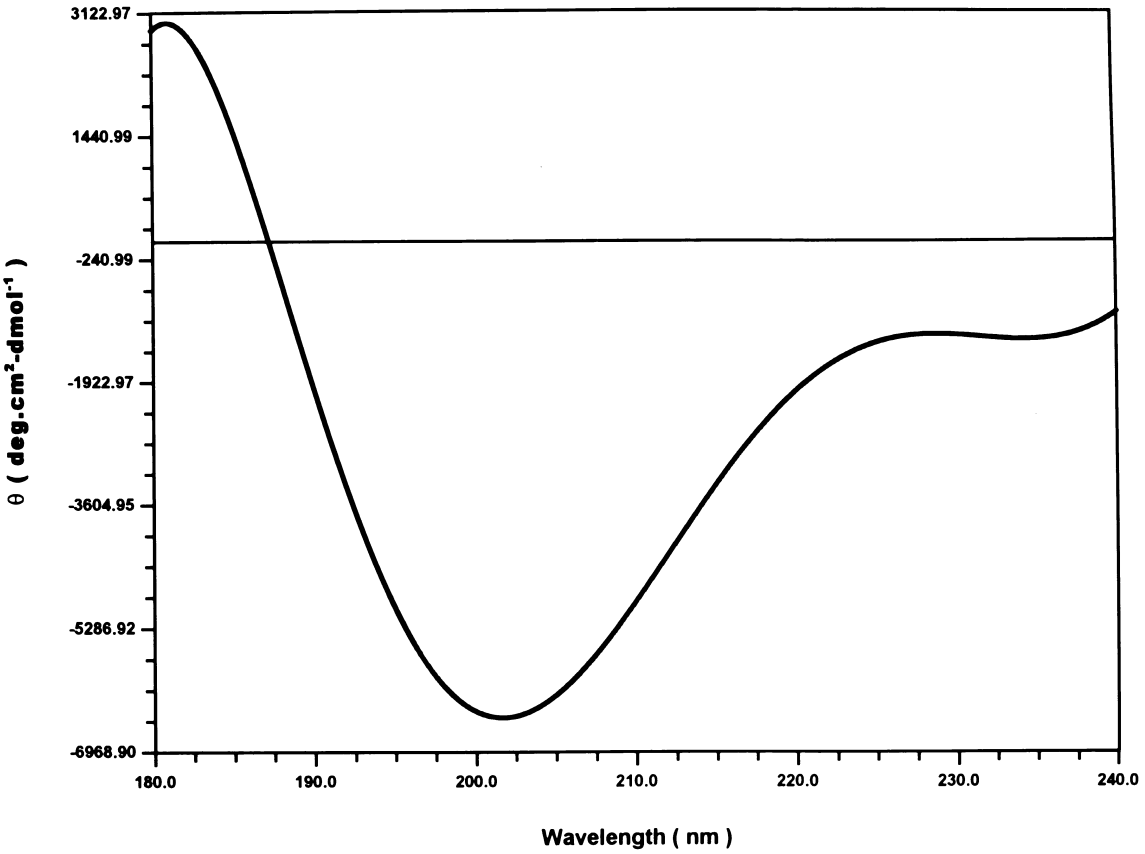

Figure 3. Circular dichroism spectrum of *A. leucocarpus* lectin. CD spectrum of the lectin was recorded on an AVIV spectropolarimeter. The protein concentration was 300 μg of ALL/ml of 30 mM PBS, pH 7.0; the scanning range was 180 to 240 nm in 5 mm path length cells. Analysis of this spectrum with PROCEC 3.0 program revealed that the secondary structure of ALL contains 45% of β-sheet and 5% of α-helix.

Table 4. Carbohydrate composition of *Amaranthus leucocarpus* lectin glycopeptides purified by affinity chromatography on ConA- and Alubia erythroagglutinin-Sepharose 4B column

	*Molar ratio**		
Sugar residue	*ALLgp*	*ALLgp ConA*	*ALLgpAE***
Xylose	0.7	0.0	0.3
Mannose	3.0	3.0	3.0
Galactose	4.0	0.0	2.7
GlcNAc	3.5	2.1	3.7

*Molar ratio is related to 3 Man residues.
**ALLgp corresponds to the glycopeptides obtained from total lectin after pronase treatment. ALLgp ConA corresponds to the fraction retained on Con A- or Alubia erythroagglutinin (AE)-Sepharose column.

the recognition of oligosaccharide sequences with the non-glycosylated lectin from *A. caudatus* [4,29].

ALL is a glycoprotein that contains 8% of sugar by weight [13], and our work shows that three putative *N*-glycosylation sites are actually present in the amino acid sequence of ALL. Treatment of the lectin with PNGase F and trypsin reveals a 35 amino acid-long cryptic peptide (residues 112 to 146), which has not been identified in the native lectin or its isoforms. The PNGase F is an amidase/amidohydrolase that cleaves the β-aspartylglycosylamine bond of asparagine-linked oligosaccharides. The enzyme has a broad substrate specificity, but both the amino and carboxyl groups of the asparagine residue have to be in peptide linkage, while the oligosaccharide must consist at least of the N,N′-diacetylchitobiose core. The enzyme is highly sensitive to modifications of this core; an α1-3-fucose substituent on the asparagine-proximal GlcNAc residue completely blocks PNGase F activity [31]. ALL lacks fucose residues and, at the present, it has not been determined whether xylose, identified in the lectin, could interfere with the enzyme activity, suggesting that exposure of new peptide sequences to trypsin effect could be the result of glycan elimination. It is worthy to note that the cryptic peptide 112 to 146 has not been identified in the native lectin or in the three isoforms, suggesting that the three lectin isofoms could be glycosylated at the predicted 89 and 123 amino acid positions; moreover, only isoform III seems to be not glycosylated at position 207, since it was the only isoform that showed an exposed peptide sequence to the trypsin treatment.

The presence of *N*-glycosidically linked glycan(s) was also suggested by the carbohydrate composition of the glycopeptides obtained from the lectin and those purified by affinity chromatography on Con A and *Phaseolus coccineus* erythroagglutinin. Con A possesses high affinity for glycans with the trimannosidic core containing man α(1,6) (man α(1,3)) man substituted by two GlcNAc residues [30]. ALL-glycopeptides interact (with low affinity) with the erythroagglutinin from alubia (*Phaseolus coccineus*) with a specificity directed toward the Gal (ß1,4 or ß1,3) GlcNAc (ß1-2) Man (α1-) saccharidic sequence present in triantennary *N*-acetyllactosamine-type *N*-glycopeptides [32]. Our data obtained by lectin-affinity chromatography indicate that ALL contains *N*-glycosidically linked glycans of the oligomannosidic-type and, in minor proportions, biantennary *N*-acetyllactosaminic-type glycans. Moreover, our data strongly suggest that *N*-glycosylation seems to be a common feature for the three ALL-isoforms, since, for total ALL, amino acid sequence 112 to 146 was not revealed until the lectin had been previously treated with PNGase F. The oligomannosidic structure seems to be common to glycan moieties found in other plant lectins, such as jacalin [33], *Sophora japonica* [34], *Erythrina americana* [35], and *Viscum album* lectin I [36]. The specific role of glycan moieties of the *A. leucocarpus* lectin, such as its effect on the topology, the internal flexibility of the molecule, or its fine sugar specificity, remains to be determined.

Acknowledgments

We thank Prof. Nathan Sharon for critical reading of this manuscript, Dr. Anthony L. Tarentino of Wadsworth Center for Laboratories and Research, Albany USA, for providing the PNGase from *Flavobacterium meningosepticum* and Marie Christine Slomianny from USTL, for MALDI analysis. This work was supported in part by CONACyT (27609 M), DGAPA-UNAM (PAPIIT-IN224598) and Program ECOS Mexico-France (M97B05).

References

1 Sharon N, Lis H, Legume lectins, a large family of homologous proteins, *FASEB J* **4**, 3198–208 (1990).
2 Weiss I, Drickramer K, Structural basis of the lectin-carbohydrate recognition, *Annu Rev Biochem* **65**, 441–73 (1996).
3 Van Damme EJ, Peumans WJ, Barre A, Rouge P, Plant lectins: A composite of several distinct families of structurally and evolutionary related proteins with diverse biological roles, *Crit Rev Biochem Mol Biol* **33**, 209–58 (1998).
4 Rinderle SJ, Goldstein IJ, Matta KL, Ratcliffe RM, Isolation and characterization of amaranthin, a lectin present in the seeds of *Amaranthus caudatus*, that recognizes the T- (or cryptic T)-antigen, *J Biol Chem* **264**, 16123–31 (1989).
5 Calderon de la Barca A, Vazquez-Moreno L, *Amaranthus cruentus* lectin: Purification, stability and some biochemical properties, *J Food Biochem* **12**, 117–26 (1988).
6 Zenteno E, Ochoa JL, Isolation and characterization of *Amaranthus leucocarpus* lectin, *Phytochemistry* **27**, 313–7 (1988).
7 Zenteno E, Lascurain R, Montaño LF, Vazquez L, Debray H, Montreuil J, Specificity of *Amaranthus leucocarpus* lectin, *Glycoconjugate J* **9,** 204–8 (1992).
8 Zenteno E, Ochoa JL, Parra C, Montaño LF, Rayon I, Maldonado G, Ruiz B, Carvajal R, Mitogenic immunosuppressive and phagocytic activity of *Machaerocerus eruca* and *Amaranthus leucocarpus* lectin. In *Lectins-biology, Biochemistry, Clinical Biochemistry*, Vol. 4, edited by Bög-Hansen TC, Breborowicz J, (Walter de Gruyter, 1985), pp. 537–46.
9 Boland CR, Chen YF, Rinderle SJ, Resau JH, Luk GD, Lynch HT, Goldstein IJ, Use of the lectin from *Amaranthus caudatus* as a histochemical probe of proliferating colonic epithelial cells, *Cancer Res* **51**, 657–65 (1991).
10 Lascurain R, Chávez R, Gorocica P, Pérez A, Montaño LF, Zenteno E, Recognition of $CD4^+$ mouse medullary thymocyte subpopulation by *Amaranthus leucocarpus* lectin, *Immunology* **83**, 410–3 (1994).
11 Lascurain R, Porras F, Báez R, Chávez R, Martínez S, Álvarez G, Zenteno E, *Amaranthus leucocarpus* lectin recognizes human naive T cell subpopulations, *Immunol Invest* **26**, 579–87 (1997).
12 Guevara J, Espinosa B, Zenteno E, Vázquez L, Luna J, Perry G, Mena R, Altered glycosylation pattern of proteins in Alzheimer disease, *J Neuropath Exp Neurol* **57**, 905–14 (1998).
13 Hernández P, Bacilio M, Porras F, Juárez S, Debray H, Zenteno E, Ortiz B, A comparative study on the purification of the *Amaranthus leucocarpus* syn. *hypocondriacus* lectin, *Prep Biochem Biotechnol* **29**, 219–34 (1999).
14 Kuster B, Wheler SF, Hunter AP, Dweek RA, Harvey DJ, Sequencing of *N*-linked oligosaccharides directly from protein gels: in-gel deglycosylation followed by matrix-assisted laser desorption/ionization mass spectrometry and normal-phase high-performance liquid chromatography, *Anal Biochem* **250**, 82–101 (1977).
15 Bradford MM, A rapid and sensitive method for quantitation of microgram quantities of proteins utilizing the principle of protein dye-binding, *Anal Biochem* **72**, 248–54 (1976).
16 Zanetta JP, Timmerman P, Leroy Y, Gas–liquid chromatography of the heptafluorobutirate derivatives of O-methyl-glycosides on capillary columns: a method for the quantitative determination of monosaccharide composition of glycoproteins and glycolipids, *Glycobiology* **9**, 255–66 (1999).
17 Catsimpoolas N, Micro isoelectric focusing in polyacrylamide gel columns, *Anal Biochem* **26**, 480–4 (1968).
18 Laemmli UK, Favra M, Maturation of the head of bacteriophage T4. I. DNA packaging events, *J Mol Biol* **80**, 575–99 (1973).
19 Hellman U, Wernstedt C, Gonez J, Heldin CH, Improvement of an "In-Gel" digestion procedure for the micropreparation of internal protein fragments for amino acid sequencing, *Anal Biochem* **224**, 451–5 (1995).
20 Geourjon C, Deleage G, SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments, *Comput Appl Biosci* **11**, 681–4 (1995).
21 Garnier J, Gibrat JF, Robson B, GOR secondary structure prediction method version IV. In *Methods in Enzymology*, Vol. 266, edited by Doolittle RF, (1996), pp. 540–53.
22 Jones DT, Protein secondary structure prediction based on position-specific scoring matrices, *J Mol Biol* **292**, 195–202 (1999).

23 Transue TR, Smith AK, Mo H, Goldstein IJ, Saper, MA, Structure of benzyl T-antigen disaccharide bound to *Amaranthus caudatus* agglutinin, *Nat Struct Biol* **4**, 779–83 (1997).
24 Raina A, Datta A, Molecular cloning of a gene encoding a seed-specific protein with nutritionally balanced amino acid composition from *Amaranthus, Proc Natl Acad Sci USA* **89**, 11774–8 (1992).
25 Calderon de la Barca A, Ochoa JL, Valencia ME, Effect of the extraction of a hemagglutinin on the nutritive value of *Amaranthus leucocarpus* seeds, *J Food Sci* **50**, 1700–8 (1985).
26 Calderon de la Barca A, Zenteno E, Ochoa JL, Valencia M, Carvajal R, The removal of *Amaranthus leucocarpus* lectin does not affect the nutritional value of the seed meal. In *Lectins-Biology, Biochemistry, Clinical Biochemistry*, Vol. 4, edited by Bög-Hansen TC, Breborowics J, (Walter de Gruyter, 1985), pp. 531–6.
27 Rini M, Lectin structure, *Annu Rev Biomol Struc* **24**, 551–77 (1995).
28 Sharma V, Surolia A, Analysis of carbohydrate recognition by legume lectin: Size of the combining site loops and their primary specificity, *J Mol Biol* **267**, 433–45 (1997).
29 Rinderle SJ, Goldstein IJ, Remsen EE, Physicochemical properties of Amaranthin, the lectin from *Amaranthus caudatus* seeds, *Biochemistry* **29**, 10555–61 (1990).
30 Debray H, Decout D, Strecker G, Spik G, Montreuil J, Specificity of twelve lectins towards oligosaccharides and glycopeptides related to *N*-glycosylproteins, *Eur J Biochem* **117**, 41–55 (1981).
31 Kuhn P, Guan C, Cui T, Tarentino AL, Plummer TH Jr, Van Roey P, Active site and oligosaccharide recognition residues of peptide-*N*-(*N*-acetyl-D-glucosaminyl)asparagine amidase F, *J Biol Chem* **270**, 29493–7 (1995).
32 Perez Campos E, Sierra C, Lascurain R, Espinoza B, Bouquelet S, Debray H, Zenteno E, The erythroagglutinin from *Phaseolus coccineus* Var. Alubia: chemical characterization, sugar specificity and effect on coagulation factors, *J Agric Food Chem* **45**, 3747–52 (1997).
33 Capon C, Piller F, Wieruszeski JM, Leroy I, Fournet B, The glycomannosidic structure from jacalin, *Artocarpus integrifolia* lectin, *Carbohydr Res* **199**, 121–7 (1990).
34 Fournet B, Leroy Y, Wieruszeski JM, Montreuil J, Poretz RD, Goldberg R, Primary structure of an *N*-glycosidic carbohydrate unit derived from *Sophora japonica* lectin, *Eur J Biochem* **166**, 321–4 (1987).
35 Ortega M, Sanchez C, Chacon E, Estrada R, Masso F, Montaño LF, Rendon J, Zenteno E, Purification and characterization of a lectin from *Erythrina americana* by affinity chromatography, *Plant Science* **72**, 133–40 (1990).
36 Debray H, Wieruszeski JM, Strecker G, Franz H, Structural analysis of the carbohydrate chains isolated from mistletoe (*Viscum album*) lectin I, *Carbohydr Res* **236**, 135–43 (1992).

Received 12 February 2001; revised 26 June 2001; accepted 9 August 2001