JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY
ARTICLE

J. Agric. Food Chem. 2010, 58, 12957–12963 12957
DOI:10.1021/jf103296n

Amaranth (*Amaranthus hypochondriacus*) Vicilin Subunit Structure

Alejandra Quiroga,[†] E. Nora Martínez,*[,†] Hélène Rogniaux,[‡] Audrey Geairon,[‡] and M. Cristina Añón[†]

[†]Facultad de Ciencias Exactas, Centro de Investigación y Desarrollo en Criotecnología de Alimentos (CIDCA), Universidad Nacional de La Plata y Consejo Nacional de Investigaciones Científicas y Técnicas, calle 47 y 116, 1900 La Plata, Argentina, and [‡]UR1268 Biopolymères Interactions Assemblages, Plate-Forme BIBS, INRA, Rue de la Geraudiere, 44300 Nantes, France

The 7S-globulin fraction is a minor component of the amaranth storage proteins. The present work provides new information about this protein. The amaranth 7S-globulin or vicilin presented a sedimentation coefficient of 8.6 ± 0.6 S and was composed of main subunits of 66, 52, 38, and 16 kDa. On the basis of mass spectrometry (MS) analysis of tryptic fragments, the 52, 38, and 16 kDa subunits presented sequence homology with sesame vicilin, whereas the 66 kDa subunit showed sequence similarity with a putative vicilin. Several characteristics of the 66 kDa subunit were similar to members of the convicilin family. Results support the hypothesis that the 7S-globulin molecules are composed of subunits coming from at least two gene families with primary products of 66 and 52 kDa, respectively. According to the present information, amaranth vicilin may be classified into the vicilin group that includes pea, broad bean, and sesame vicilins, among others.

KEYWORDS: Amaranth; vicilin; storage protein; peptide sequence; 7S-globulin; *Amaranthus hypochondriacus*

INTRODUCTION

Amaranth is a dicotyledonous plant that is considered a pseudocereal because of its high production of seeds. The seeds have proteins with a better balanced content of the essential amino acids than that of cereals and legumes (*1*).

Globulins constitute one of the major protein fractions of amaranth seeds and are mainly represented by legumin-like species (*2*−*6*). A minor globulin component, the 7S-like fraction, has been less studied. Some previous studies have revealed the presence of this fraction (*7*,*8*,*3*) which was purified and studied in more detail by Marcone (*9*).

The 7S-9S globulins, grouped as vicilins, are generally characterized by a trimeric organization, with subunits between 40−60 kDa and a molecular mass of 150−200 kDa. These globulins are frequently N-glycosylated and lack disulfide bridges joining their subunits. Some of them have been deeply studied and their structures are X-ray resolved (*10*−*12*). Notably, vicilins detected in other plants possess a complex subunit composition over an extended range of molecular weights (*13*). It is not clear whether these 7S-globulin subunits, with molecular weights below 45 kDa, came from limited proteolysis of a primary gene product during seed maturation or resulted from proteolytic degradation during the extraction and purification procedure. The hypothesis of polypeptide processing during maturation was postulated based on experiences with *Pisum sativum* L. and *Vicia faba L* (*13*, *14*). It was proposed (*15*) that 7S-globulins could be classified into two groups on the basis of whether the subunits undergo post-translational proteolysis (*Pisum sativum* L, *Vicia faba* L.) or not (*Glycin max*, *Phaseolus vulgaris* L). Globulins from the latter group were more extensively glycosylated.

The amaranth 7S globulin would belong to the first group since it was found to be composed of a great number of subunits from 15 to 90 kDa (*3*, *7*−*9*). Its molecular mass, as determined by gel filtration, was near 200 kDa and its pI was between 5.2 and 5.8 (*9*). In spite of this information, knowledge of the subunit composition and structure of amaranth 7S-globulin is by far less advanced than that of amaranth 11S globulin. A more extensive characterization of the 7S-globulin would improve our knowledge of the biology of amaranth proteins and its relationship with other storage proteins.

The study reported here provides new information about the subunit composition of amaranth 7S-globulin, which allows the postulation that the amaranth vicilin molecules are composed of subunits coming from at least two gene families. The product of one of them may be processed by post-translational proteolysis generating polypeptides that remain in the molecular structure. These results may help to disclose the influence of this protein component on the functional and physiological properties of amaranth proteins, especially from the standpoint of their use as nutritional ingredients.

MATERIALS AND METHODS

Plant Material. Seeds of *Amaranthus hypochondriacus* (cultivar 9122) were obtained from Estación Experimental del Instituto Nacional de Tecnología Agropecuaria (INTA), Anguil, La Pampa, Argentina.

*Corresponding author. Tel/fax: +54 221 4254853/+54 221 424 7289. E-mail: noramar@quimica.unlp.edu.ar.

Published on Web 11/30/2010

Seeds were ground in an Udy mill (Facultad de Ciencias Agrarias y Forestales, Universidad Nacional de La Plata, Argentina) and screened by a 0.092 mm mesh. Flour was suspended in hexane 10% (w/v), defatted for 24 h under continuous stirring, air-dried at room temperature, and stored at 4 °C until used.

Protein Extraction. Partially purified 7S- and 11S-globulins were obtained as described in ref (*16*) with minor modifications. Briefly, the flour was treated twice with water to extract albumins and then twice with 10 mM Na_2HPO_4, 0.1 M NaCl, pH 7.5 to extract 7S-globulin. The 11S-globulin fraction was extracted afterward with 10 mM Na_2HPO_4, 0.8 M NaCl, pH 7.5. Treatments were performed at room temperature using a 1:10 (w/v) flour/solution or residue/solution ratio, and after each treatment the residue was separated by centrifugation at 9000*g* for 20 min at room temperature. Supernatants containing 7S-globulin or 11S-globulin fractions were each adjusted to pH 6 with 2 N HCl. The resulting precipitates were suspended in water, neutralized with 0.1 N NaOH, and freeze-dried.

Fast Protein Liquid Chromatography (FPLC) Gel Filtration. The partially purified 7S-like and 11S globulins were analyzed at room temperature in a Superose 6B HR 10/30 column using a FPLC System (Pharmacia LKB, Uppsala, Sweden). Samples (4 mg of protein) were dissolved in 0.2 mL of buffer A (32.5 mM K_2HPO_4−2.6 mM KH_2PO_4, pH 7.5, 0.4 M NaCl) and were eluted with the same buffer at a flow rate of 0.2 mL/min. Fractions of 0.5 mL were collected, and the elution profile (absorbance at 280 nm) was obtained. The column was calibrated with gel filtration calibration kits (GE-Healthcare, Buckinghamshire, UK).

Curves were processed and data were evaluated using the Pharmacia AB, FPLC director and FPLC assistant software.

Ultracentrifugation. Ultracentrifugation was performed in a Beckman Coulter Optima LE-80K ultracentrifuge using a SW 41 Ti rotor. Samples of 3−10 mg of 7S-like or 11S-globulin in 0.3 mL of buffer A were layered on top of 5−20% (w/v) sucrose linear gradients in buffer A and centrifuged at 15 °C for 16 h at 38 000 rpm. Sedimentation constants of the peaks were calculated by comparison with the standard proteins catalase, aldolase, and ribonuclease A with known sedimentation constants (11.3S, 7.3S, and 2.0S, respectively) (*17*, *18*). Mean values of four samples (two replicates from two preparations) analyzed in two runs are reported. Gradients were fractionated, and the UV-absorbance (280 nm) of each fraction (0.4 mL) was determined in a Beckman DU650 spectrophotometer.

Sodium Dodecyl Sulfate−Polyacrylamide Gel Electrophoresis (SDS−PAGE). Fractions coming from gel filtration FPLC were analyzed in gels (12% w/v polyacrylamide, running gel; 4% w/v polyacrylamide, stacking gel) arranged in minislabs (BioRad Mini Protean II Model, CA).

Fractions from ultracentrifugation gradients were analyzed in gels (13.5% w/v polyacrylamide, running gel; 4% w/v polyacrylamide, stacking gel) arranged in Mini VE GE-Healthcare (Uppsala Sweden).

Runs were carried out according to the Laemmli (*19*) method as modified by ref (*20*). Molecular masses of polypeptides were calculated using the LMW protein standards (GE-Healthcare). Values were obtained from at least three replicates.

Samples: proteins from gel filtration or ultracentrifugation fractions were precipitated with 15% trichloroacetic acid. Precipitates were washed with acetone and dissolved in sample buffer (0.125 M Tris-hydroximethilaminomethane (Tris)-HCl, pH 6.8, 20% (v/v) glycerol, 1% (w/v) SDS, 5% (v/v) 2-mercaptho ethanol (2-ME) and 0.05% (w/v) bromophenol blue) and were heated (100 °C, 3 min). Gels were fixed and stained with Coomassie Brilliant Blue Stain.

Two-Dimensional Electrophoresis (IEF−SDS−PAGE). The first-dimension isoelectric focusing (IEF) was run using 18-cm linear immobiline pH gradient (IPG) strips (pH 3−10) in the IPGphor system (GE Healthcare UK Limited, England). Precipitates from the sucrose gradient fractions (approximately 100 μg of protein) were dissolved in 340 μL of the rehydration buffer (7 M urea, 2 M thiourea, 2% 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate (CHAPS), 20 mM dithiotreithol (DTT), 0.5% v/v IPG and 0.002% w/v bromophenol blue) and the IPG strips were rehydrated with these samples. The voltage settings for IEF were 500 V for 1 h, 1000 V for 1 h, and 8000 V for 4 h until reaching a final condition of 30 000 V/h.

Following IEF, the gel strips were incubated with equilibration buffer (2% SDS, 6 M urea, 75 mM Tris-HCl pH 8.8, 30% v/v glycerol, 1% w/v DTT, and 0.01% w/v bromophenol blue) for 1 h, followed by equilibration for a further hour with the same solution containing 2.5% w/v iodoacetamide instead of DTT. The strips were placed onto 15% polyacrylamide gels and were run in a vertical electrophoresis device (Ettan DALTsix, GE Healthcare, Uppsala Sweden) with the buffer system described for SDS−PAGE. Strips were overlaid with agarose sealing solution (0.25 M Tris base, 1.92 M glycine, 1% w/v SDS, 0.5% w/v agarose, 0.002% w/v bromophenol blue). Runs were performed in two steps, step 1: 10 mA/gel, 80 V, W/gel = 1. Time 1 h; step 2: 12 mA/gel, 150 V, W/gel = 2. Time 15−17 h.

All gels were fixed and stained with Coomassie Brilliant Blue.

In-Gel Enzymatic Digestion of Proteins. The protein spots indicated in **Figure 3** were excised and washed with 25 mM NH_4HCO_3, followed by dehydration with 50% (v/v) acetonitrile in 25 mM NH_4HCO_3. The proteins therein were then reduced with 10 mM DTT and alkylated with 55 mM iodoacetamide. The gel bands were further washed and dehydrated as described above and finally dried in a vacuum centrifuge. For protein digestion 15 μL of trypsin solution (Promega, Madison, WI, USA, 12.5 ng/L in 25 mM NH_4HCO_3) were added to each sample and the mixture was incubated overnight at 37 °C. The resulting peptide mixture was acidified with 1% (v/v) formic acid and stored at −20 °C until analysis.

Mass Spectrometry. Mass-spectrometry analyses were conducted by the platform "Biopolymers- Structural Biology" located at the INRA Center of Angers-Nantes, France (http://www.angers-nantes.inra.fr/plateformes_et_plateaux_techniques/plate forme_bibs).

Peptides were analyzed by liquid chromatography−mass spectrometry/mass spectrometry (LC-MS/MS) through the use of a Switchos-Ultimate II capillary LC system (LC Packings/Dionex, Amsterdam, The Netherlands) coupled to a hybrid quadrupole orthogonal acceleration time-of-flight mass spectrometer (Q-TOF Global, Micromass/Waters, Manchester, UK). Chromatographic separation was conducted on a reverse-phase capillary column (Pepmap C18, 75-μm × 15-cm, LC Packings) with a linear gradient from 2% to 40% (v/v) acetonitrile in 50 min, followed by an increase to 50% acetonitrile within 10 min, at a flow rate of 200 nL/min.

Mass data acquisitions were piloted by the Mass Lynx software (Micromass/Waters): the MS data were recorded for 1 s on the mass-to-charge (m/z) range 400−1500, and the three most intense ions (doubly, triply, or quadruply charged) were selected and fragmented in the collision cell (MS/MS measurements).

Protein Identification−Databank Searching from LC-MS/MS Data and De Novo Sequencing. Raw data were processed by means of the Protein Lynx Global Server v. 2.1 software (Micromass/Waters). Protein identification was performed by comparing the collected LC-MS/MS data against Uniprot databank restricted to *Viridiplantae* species (Uniprot release 2010-04 containing 11 134 468 sequences; taxonomy filtered databank contained 751 663 sequences). The mass tolerance was set at 150 ppm for parent ions (MS mode) and 0.3 Da for fragment ions (MS/MS mode), and one missed cut per peptide was allowed. Databank searches were performed through the use of the Mascot server v. 2.2 program (Matrix Science). Validation of proteins was done first by setting the requirements to a minimum of two MS/MS spectra matching the databank sequence with individual MASCOT ion scores above the significance threshold (threshold score of 40, $p < 0.05$). This did not yield any protein identification. The requirement for validation was then lowered to one MS/MS spectrum matching the databank sequence with a significant score. This again did not produce any hit. Consequently, all MS/MS spectra were subjected to a de novo sequencing through the use of the Protein Lynx Global Server v. 2.1 software. The results were then fed into the OVNIp program (*21*). This program enables the filtration of the de novo sequences to keep a nonredundant list of sequence tags having high scores as determined by the Protein Lynx Global Server program. These tags were then submitted to a sequence-homology search by means of the MS-BLAST program, with the PAM30MS matrix and the NrDB95 databank provided on the MS-BLAST server located at EMBL (released on March 17/2005: 2 078 555 entries). The protein sequences producing high-scoring segment pairs were considered.

Multiple Sequence Alignments. The sequenced peptides of all the spots were analyzed by the multiple sequence alignment program ClustalW2 (http://www.ebi.ac.uk/Tools/clustalw2/). Sequences with the best match were selected and the polypeptides with sequence homology

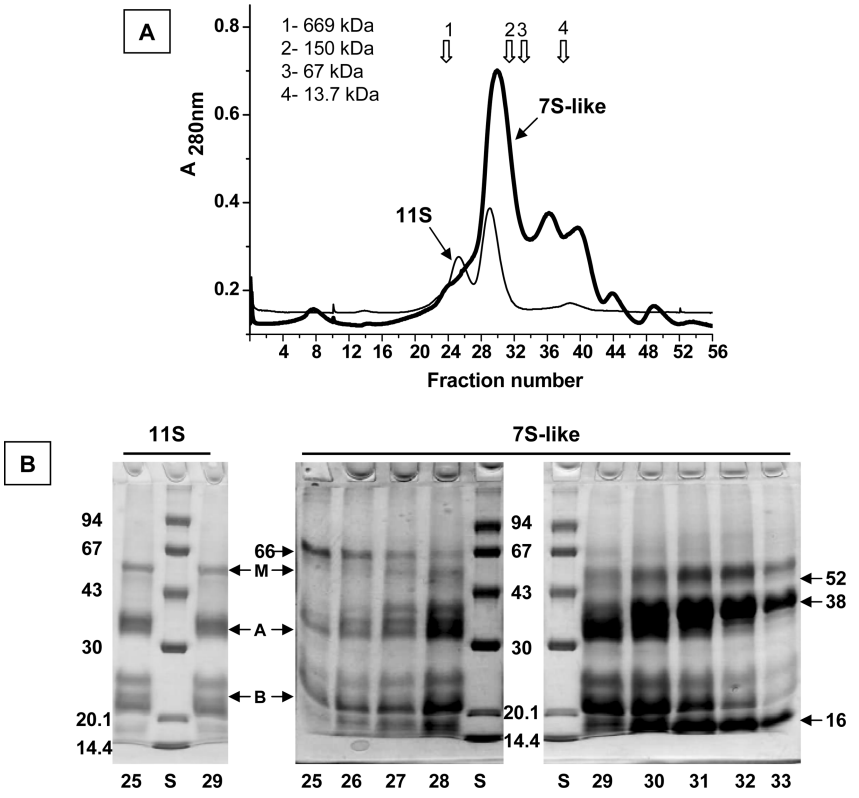

Figure 1. (**A**) Chromatographic profiles on FPLC of the partially purified 7S-like globulin and 11S-globulin. (**B**) SDS−PAGE patterns of 0.5 mL fractions corresponding to 11S-globulin chromatogram (11S) and 7S-like chromatogram (7S-like). The 7S-like polipeptides of 66.0 ± 1.0, 52 ± 1.1, 37.9 ± 0.3, and 15.8 ± 0.4 kDa are labeled in the figure as 66, 52, 38, and 16 respectively. M, A, and B 11S globulin polypeptides are also indicated. Lane S: corresponds to standard proteins pattern and on the left side are the standard molecular masses.

were grouped. These polypeptides were arranged in **Figure 5** as fragments of a common precursor by locating their sequences in a hypothetic primary polypeptide which was constructed using the sesame vicilin sequence as a guide.

RESULTS

Chromatographic Analysis. The partially purified 7S-like globulin fraction was analyzed by size exclusion chromatography and ultracentrifugation, and results were compared with those for the 11S globulin. FPLC profiles (**Figure 1A**) show that the 7S-like globulin main peak eluted at a rather higher volume than that of the 11S-globulin. The chromatographic profile of the 7S-like globulin fraction also shows two other peaks, eluting at higher volumes than the main peak, that correspond to small polypeptides (SDS−PAGE profile not shown).

In the SDS−PAGE analysis of the chromatographic fractions of the 7S-like globulin (**Figure 1B**, 7S-like), fraction 25, which eluted at a lower volume than that of the main peak, exhibited the M, A, and B polypeptides characteristic of 11S globulin (*5*, *22*) (which were also present in 11S fractions 25 and 29, as shown in **Figure 1B**), together with an additional band at 66.0 ± 1.0 kDa. Fractions eluting at higher volumes also contained additional bands at 52 ± 1.1, 37.9 ± 0.3, and 15.8 ± 0.4 kDa. These bands increased their relative intensity in the patterns corresponding to increasing eluting volumes, while A, B, and 66.0 ± 1.0 kDa bands faded away. These results suggest that, unlike the 11S preparation, the 7S-like globulin preparation is composed of several globulins. Therefore, the main peak of the 7S-like fraction might be composed of 11S molecules, eluting at rather lower volumes, same as in the 11S preparation, and 7S molecules, eluting at higher volumes. These molecules would be integrated by the 52 ± 1.1, 37.9 ± 0.3, and 15.8 ± 0.4 kDa polypeptides.

Ultracentrifugation Analysis. Ultracentrifugation profiles are shown in **Figure 2A** and the SDS−PAGE patterns of the 7S gradient fractions are depicted in **Figure 2B**. According to the results, the 7S-like preparation was composed of three species, a 2.0 ± 0.3 S fraction (peak I, **Figure 2A**) with polypeptides of 23.7 ± 1.8, 20.5 ± 0.9, 18.2 ± 1.3, 16.8 ± 1.2, 14.2 ± 1.5, and 10.5 kDa (**Figure 2B**, fractions 6 and 7), an 8.6 ± 0.6 S component (peak II, **Figure 2A**) with main bands of 66.0 ± 1.0, 52 ± 1.1, 37.9 ± 0.3, and 15.8 ± 0.4 kDa and minor bands at approximately 35, 31, 23, and 21 kDa (**Figure 2B**, fractions 17, 18, and 19), and an 12.2 ± 1.2 S species (peak III, **Figure 2A**) with the major bands of 52−55, 33−35, and 21−22−24 kDa corresponding to the amaranth 11S-globulin (**Figure 2B**, fractions 23 and 24) (*5*, *22*).

According to these results, ultracentrifugation provided a better separation of the 7S-like and 11S-globulins. Therefore, the ultracentrifugation 8.6 ± 0.6 S fraction was further analyzed.

MS-Analysis of Tryptic Fragments. Species in this fraction were resolved by two-dimensional electrophoresis and the spots were analyzed by LC-MS/MS using a conventional proteomic approach in which protein identity was assessed by comparing the masses of collected peptides with reported sequences in the Uniprot databank pertaining to green plants. However, this error-intolerant databank search did not yield any valid protein hits (see Material and Methods for description of validation criteria). Therefore, the remaining, nonassigned, MS/MS mass spectra were submitted to a de novo sequencing, producing several stretches of amino acids of various lengths which were then subjected to a tentative alignment against the sequences reported in the NrDB95 databank by means of the MS-BLAST program. As shown in **Figure 3** and **Table 1**, with this "error-tolerant" approach, we could assign several spots of 66.0 ± 1.0, 52 ± 1.1, 37.9 ± 0.3, 34.9 ± 0.5, and 15.8 ± 0.4 kDa — labeled P66, P52, P38

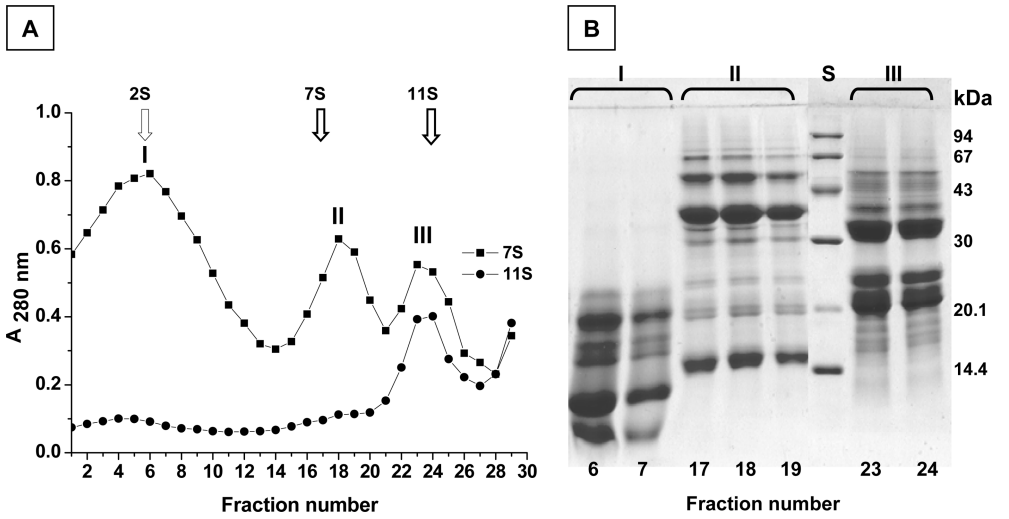

Figure 2. (**A**) Sedimentation profiles of the partially purified 7S-like globulin and 11S-globulin. Peaks of the 7S profile are labeled I, II, and III. Arrows indicate the absorbance peaks of protein markers, 2S, ribonuclease A; 7S, aldolase; 11S, catalase. (**B**) SDS−PAGE patterns of 7S gradient fractions under peak I (fractions 6 and 7), peak II (fractions 17, 18, and 19), and peak III (fractions 23 and 24). Lane S: corresponds to standard proteins pattern and on the right side under kDa are the standard molecular masses.

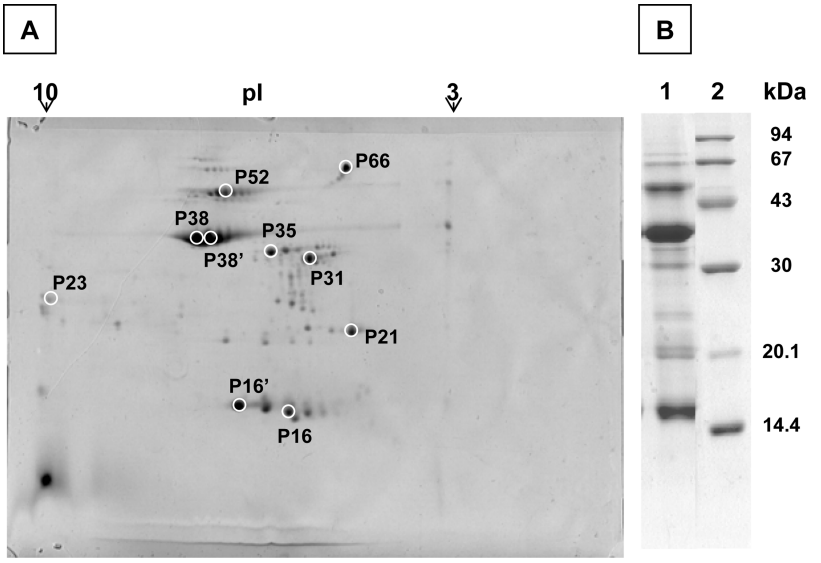

Figure 3. (**A**) 2D-electrophoresis pattern of the 7S globulin (peak II, fraction 17 of the 7S sedimentation gradient). On top, pI range. Spots analyzed are indicated in the pattern. (**B**) Lane 1, SDS−PAGE pattern of the 7S globulin same sample as 2D electrophoresis), lane 2, molecular mass standard proteins, kDa standard molecular masses.

and P38′, P35 and P16, respectively — to vicilin-like proteins. Although the complete sequence of the protein present within each spot is not derived from these data, a functional description can be extrapolated by similarity to other protein sequences. Thus, according to the found homologies, these spots correspond to subunits of the amaranth vicilin (7S-globulin). On the other hand, the spots P31, P23, P21, and P16′ (**Figure 3** and **Table 1**) of approximately 31, 23, 21, and 16 kDa were not assigned to a vicilin protein by this methodology. According to the results of the MS-BLAST alignment, the P31, P23, and P16′ spots correspond to contaminant 11S-globulin acid and basic subunits, respectively, whereas the P21 spot corresponds to a contaminant body membrane protein. In a second attempt to disclose the subunit composition of the 7S-globulin, sequences corresponding to the 2D spots were analyzed using the ClustalW2 multiple sequence alignment program (*23*). Sequence homologies were found among subunits P52, P38 and P38′, P35, P16 and P16′. Many of these homologous sequences also aligned with the sesame 7S-globulin (**Table 1**). Amaranth subunits also present homologous sequences which do not align with sesame's; these sequences are shown in **Figure 4**. Although P16′ was found by MS-Blast alignment to present homology with an 11S polypeptide, on the basis of results found with ClustalW2 alignment, it is considered to belong to the 7S-globulin. In contrast, no sequence homology was found between P66 and the other polypeptides.

DISCUSSION

These results suggest that (i) amaranth vicilins come from a multigenic family, and (ii) similar to sesame vicilin (*13*, *15*), some of the observed amaranth vicilin subunits (P52, P38−38′, P35, P16′ and P16) derive from the posttranslational hydrolysis of the product of a single gene. In support of this theory, **Figure 5** shows a possible distribution of peptides (location of the peptides) in the primary sequence taking as a guide the sesame vicilin sequence.

Table 1. Protein Alignment of Trypsin-Digested 7S-Globulin Polypeptides Analyzed by MS-BLAST

Spot	Peptides [a]	Accession number [b]	Protein [b]	Score [c]
P66	LLDAQGK, VLQTLSK, ASVLQTLSK, NHPQFLVGK, SLQVVSALR, CPSAPLNQECK, VFMGFSTTSK, FRLEEGDVFTLPR, QAGEEER- EQEEWQQQEEE, RGEEEEER(-)EEEAQQR, EEEQQQEEEE(-)K	Q9ZU69	Putative vicilin (*Arabidopsis thaliana*)	375
P52	LQLFR, FKNFKR, SVPFGGK, REQE(-)R, SPGPLQLFR, NNEQLVLVK, LLEGLENYR, FREQEGNVR, LAVVLNGR, LFSQQSEQLLR, EQLSALTHEK, ESTTPVHYER, GHFEMACP, EEQLSALTH, LAFASSA(-)R, QEEEFFFPGPR, QSNAFGTLFETD(-)R	Q9AUD0	7S globulin (*Sesamum indicum*)	246
P38	REQEGNVR, SVPFGGK, DSGPLQLFR, NNEQLVLVK, LLEGLENYR, LLNPVSNPSGK, FREQEGNVR, PAGVTYL(-)R, LFSQQSEQLLR, LFSQGA(-)R, GHFEMA(-)PHVSK, GHFEMA(-)K, GALSALT, EQLSALTHEKR, HQHHQ(-)R, LAFASSA, VLPAGVT(-)R, VLPARTA(-)R, FGTLFETDFDDRR, AGSNAFGTLFE(-)R, QSNAFGTLFE(-)R, ESGAPY(-)K	Q9SEW4	Vicilin-like protein (*Juglans regia*)	179
		Q9AUD0	7S globulin (*Sesamum indicum*)	115
P38'	HADPFGGK, DSGPLQLFR, LLEGLENYR, LVGVLLNPVSNPSGK, LVGYTSVFR, AFSTELLEAAYK, LFSQQSEGALLR, PAGVTAYL, -FEMACP(-)TLPNFF, ASEEQLSALTH, LPLLNYLQL, QGEGTVSLV, VMVLPAGVTA, QSNAFGTLM(-)SQQSE(-)RR, FQENGELVDL, EEEEVGGESGAPYVFDEQH-	Q9AUD0	7S globulin (*Sesamum indicum*)	291
P35	SLAMSGR, ESVLQR, LGLDRFR, LLEGLENYR, --NEKLVLVK, LLNPVSNP(-)K, LFSQQSEGALLR, APHYNSR, LGLTLLMQE(-)R, EQLSALT, AVLTLLMKENR, LAFASS, QSNAFGTLFE, SESFESLYSTQ, VVLVEEGSGYL, FDSESFESLYS	Q9AUD0	7S globulin (*Sesamum indicum*)	200
P31	LLEGLENYR, VVCPSVLTNK, FREQE(-)R, LFSENNLL(-)R, CPSVLTNKECK(-)R, LLECTSCAEEQER, LDQLSAN, SVLTPKECK(-)R, ECTSCAEE(-)R, EEDLEQG, MMEEDLEQEE(-)R, MMEEELEQEE	Q8W3X8	PSV membrane protein (*Cucurbita maxima*)	189
		Q9FYW3	11S globulin (*Lycopersicon esculentum*)	117
P23	QELKYNR, HLDMSLEK, DSDVGVYTR, DAQELKYNR, VNEHKLPLLK, LVGYTSVFR, QHQTFFLPSR, EESLCTFKLR, GFEYVSFK, AGNDGF, SLFYLHSSTKAE(-)K, QTFFL(-)R, QSPLVGYTSVFR, NPGFEYVSFK, LSTSFNVPDTTL, LLQAAL, EEEEVGGESGA	Q40346	Legumin precursor (*Magnolia salicifolia*)	180
P21	GPLQLFR, PSVLTNKECK, VFTVPR, SLFYLHSS(-)K, LHSST(-)K, LECTSCAEWAER, LECTSCAE, LSTSFNV, SFSPELLQAALNLP(-)K, VEE, DTEVSVVLEG(-)R, LSKAFG, MEEDLEQEE	Q8W3X8	PSV membrane protein (*Cucurbita maxima*)	183
P16'	FKNFKR, NLFKNFKR. LVGYTSVFR, LAFASSAEE, LFAGFDTR, VLLNSYQLSPR, QEEEFFF, VLPQYF, ESTTPVHYERLTA(-)R, QELKVPN(-)TFFLPSR	Q9XHP0|	11S globulin (*Sesamum indicum*)	103
P16	LFKNFKR, ESTTPVHYER, GFGLNA(-)R, LAFASSAE(-)R, QEEEFFFPGPR, TVFEEFFFP(-)R	Q9SPL4	Vicilin-like protein (*Macadamia integrifolia*)	128

[a] Tryptic peptide sequences. Underlined sequences aligned to the protein of the same row; shaded sequences aligned with sesame vicilin. [b] UniProtKB entries. [c] MS-BLAST total score.

```
P52  :  SPGPLQLFR    FREQEGNVR              LAFASSA    ESTTP VHYER      F KNFKR
P38:    DS- GPLQLFR   REQEGNVR               LAFASSA
P38':   DS- GPLQLFR                LVGYTSVFR LAFASSA
P35:                               LVGYTSVFR LAFASSA
P16':                              LVGYTSVFR LAFASSAE E ESTTP VHYERLTA  NLF KNFKR
P16:                               LVGYTSVFR  LAFASSAE  ESTTP VHYER     LF KNFKR
```

Figure 4. Alignment of P52, P38, P38′, P35, P16′, and P16 sequences that do not align with sesame vicilin.

In **Figure 5** the subunits are arranged according to their equivalent sequences and their lengths. P16′ and P16 subunits would correspond to the C-terminus region of the precursor as suggested by their sequence homology with the sesame and P52 sequences. On the other hand, the lack of those sequences in P38−38′ and P35 suggests that the C-terminal regions of these subunits are located relatively far from the sesame C-terminus. The existence of a sequence (**LAFASSA**, see **Figure 4**) shared by the five subunits indicates that they overlap in that region, suggesting that the large (P38−38′ and P35) and small (P16′ and P16) fragments are products of proteolysis occurring at different cleavage points.

Figure 5. Location of amaranth vicilin polypeptides sequences in a hypothetical primary sequence using the sesame vicilin as a guide. On the left side are the names of the proteins, on top, sesame vicilin (Se 7S) followed by the amaranth 7S subunits P52, P38 and P38′, P35, P16 and P16′. According to the determined molecular masses, the length of the polypeptides was proposed and their N- and C-termini were marked with an X. The number of residues of each chain is on the right side. Sequences equal to sesame's are underlined and highlighted.

Overall, these findings further support the classification of the purified protein as a 7S globulin or vicilin, and are in accordance with Marcone (*9*) who reported the existence of a 7S globulin-like storage protein composed by polypeptides with molecular masses ranging from 15.6 to 90.1 kDa. In this work, we have characterized the amaranth vicilin as having a sedimentation constant of 8.6 ± 0.6 S and containing at least four major subunits that would be the product of at least two gene families encoding different primary sequences. According to these results, the amaranth vicilin could be classified into the vicilins group that includes pea, broad bean, and sesame vicilins, among others, with subunits undergoing post-translational proteolysis. In the amaranth vicilin, the 52 kDa polypeptide would be the primary gene product, and the lower molecular weight subunits would be the fragments resulting from cleavage at different processing sites. It cannot be ruled out that some of the vicilin polypeptides, especially those found in minor proportion (P35 and P16′), result from the action of some intrinsic protease during the extraction steps. In this regard, the fact that aspartic proteases have been detected in amaranth protein preparations must be considered (*24*).

Unlike the 52 kDa protein, the 66 kDa polypeptide presented several characteristics similar to members of the convicilin family, or the α and α′ polypeptides from β-conglycinin, which are larger and more acidic than the other group (vicilin and β conglycinin β subunit) (*25*−*28*). Thus, our results suggest that in amaranth, as in other plants (*27*), there are two size classes of vicilins that do not arise from the differential processing of a common precursor transcript, but result from the transcription of distinct genes. Convicilins, in the same way as β-conglycinin α and α′ subunits, differ from vicilins or β subunit in the presence of an N-terminal additional sequence of amino acids corresponding to the first exon. It has been proposed (*25*, *29*) that the N-terminal extension of the convicilin gene evolved from the vicilin gene mainly by a series of duplications of short internal sequences and triplet expansions, and, as in the case of soybean, the β-subunit genes are the progenitor vicilin genes. The present results do not allow us to conclude that the amaranth 66 kDa subunit gene is related to the 52 kDa subunit one as described for the vicilins because we have not found sequence homology between these two polypeptides. Nevertheless, considering the previously mentioned methodological limitations, such possibility cannot be excluded.

The chromatographic results (**Figure 1**A,B) provided evidence of the presence of vicilin molecules of different sizes and subunit composition. Large vicilin molecules eluting in the same fractions as 11S-globulin molecules are mainly composed of the P66 subunits (**Figure 1**B, 7S-like fraction 25). On the other hand, the main subunits of the smallest vicilin molecules are P52, P38−38′ and P16 (**Figure 1**B, 7S-like fractions 32 and 33). There are also vicilin molecules of intermediate molecular weight that contain P66 and P16 (**Figure 1**B, 7S-like fraction 26), and others with P66, P52, P38−38′ and P16 in different proportions (**Figure 1**B 7S-like fractions 27−31). These results, which suggest the existence of vicilin molecular isoforms composed of different subunit arrangements, are in agreement with data showing that pea vicilin molecules carry convicilin and vicilin polypeptides in different proportions (*27*). The molecular heterogeneity is typical of storage proteins and also has been described in other vicilins (*25*, *30*, *31*).

Regarding the molecular masses of amaranth vicilins, the chromatographic results obtained in the present study seem to contradict those obtained by ultracentrifugation, since the sedimentation profile revealed only one species of vicilin with a unique sedimentation constant. A possible explanation for this discrepancy is that the shape of vicilin molecules in our preparation is less spherical than the shape of the 11S-globulin molecules. Therefore, both globulins (11S and vicilin) present a differential hydrodynamic behavior which is reflected in the different results obtained by chromatography and ultracentrifugation. Moreover, differences in MW and shape of the different vicilin isoforms would cause them to be separated by chromatography and not by ultracentrifugation.

ABBREVIATIONS USED

2-ME, 2-mercaptho-ethanol; CHAPS, 3-[(3-cholamidopropyl)-dimethylammonio]-1-propanesulfonate; DTT, dithiotreithol; IPG, immobiline pH gradient; SDS−PAGE, sodium dodecyl sulfate−polyacrylamide gel electrophoresis; FPLC, fast protein liquid chromatography; IEF, isoelectric focusing; LC-MS/MS, liquid chromatography−mass spectrometry/mass spectrometry; Tris, Tris-hydroximethil-aminomethane; P66, P52, P38−P38′, P35, and P16, polypeptides of 66.0 ± 1.0, 52 ± 1.1, 37.9 ± 0.3, 34.9 ± 0.5, and 15.8 ± 0.4 kDa, respectively; P31, P23, P21, and P16′, polypeptides of approximately 31, 23, 21, and 16 kDa.

LITERATURE CITED

(1) Bressani, R. Composition and nutritional properties of amaranth. In *Amaranth: Biology, Chemistry and Technology*; Paredes-López, O., Ed.; CRC Press: Boca Raton, FL, 1994; pp 185−205.

(2) Segura-Nieto, M.; Barba de la Rosa, A. P.; Paredes-López, O. Biochemistry of amaranth proteins. In *Amaranth: Biology, Chemistry and Technology*; Paredes-López, O., Ed.; CRC Press: Boca Raton, FL, 1994; pp 75−106.

(3) Martínez, E. N.; Castellani, O. F.; Añón, M. C. Common molecular features among amaranth storage proteins. *J. Agric. Food Chem.* **1997**, *45*, 3832−3839.

(4) Gorinstein, S.; Jaramillo, N.; Medina, O.; Rogriques, W.; Tosello, G.; Paredes-Lopez, O. Evaluation of some cereals, plants and tubers through protein composition. *J. Prot. Chem.* **1999**, *18*, 687−692.

(5) Quiroga, A. V.; Martínez, E. N.; Añón, M. C. Amaranth globulin polypeptide heterogeneity. *Protein J.* **2007**, *26*, 327−333.

(6) Quiroga, A. V.; Martínez, E. N.; Rogniaux, H.; Geairon, A.; Añón, M. C. Globulin-p and 11S-globulin from *Amaranthus hypochondriacus* are two isoforms of the 11S-globulin. *Protein J.* **2009**, *28*, 457−467.

(7) Segura-Nieto, M.; Vázquez-Sánchez, N.; Rubio-Velázquez, H.; Olguín-Martinez, L.; Rodriguez-Nester, C.; Herrera-Estrella, L. Characterization of amaranth (*Amaranthus hypochondriacus* L.) seed proteins. *J. Agric. Food Chem.* **1992**, *40*, 1553−1558.

(8) Barba de la Rosa, A. P.; Paredes-López, O.; Gueguen, J. Characterization of amaranth globulins by ultracentrifugation and chromatographic techniques. *J. Agric. Food Chem.* **1992**, *40*, 937−940.

(9) Marcone, M. Evidence confirming the existence of a 7S globulin-like storage protein in *Amaranthus hypochondriacus* seed. *Food Chem.* **1999**, *65*, 533−542.

(10) Lawrence, M. C.; Izard, T.; Beuchat, M.; Blagrove, R. J.; Colman, P. M. Structure of phaseolin at 2.2 Å resolution. Implications for a common vicilin/legumin structure and the genetic engineering of seed storage proteins. *J. Mol. Biol.* **1994**, *238*, 748−776.

(11) Ko, T. P.; Day, J.; McPherson, A. The refined structure of canavalin from jack bean in two crystal forms at 2.1 and 2.0 Å resolution. *Acta Crystallogr.* **2000**, *D56*, 411−420.

(12) Maruyama, N.; Adachi, M.; Takahashi, K.; Yagasaki, K.; Kohno, M.; Takenaka, Y.; Okuda, E.; Nakagawa, S.; Mikami, B.; Utsumi, S. Crystal structures of recombinant and native soybean β-conglycinin β homotrimers. *Eur. J, Biochem.* **2001**, *268*, 3595−3604.

(13) Gatehouse, J. A.; Croy, R. R. D.; Morton, H.; Tyler, M.; Boulter, D. Characterisation and subunit structures of the vicilin storage proteins of pea (*Pisum sativum* L.). *Eur. J. Biochem.* **1981**, *118*, 627−633.

(14) Scholz, G.; Manteuffel, R.; Müntz, K.; Rudolph, A. Low molecular weight polypeptides of vicilin from *Vicia faba* L. Are products of proteolytic breakdown. *Eur. J. Biochem.* **1983**, *132*, 103−107.

(15) Orruño, E.; Morgan, M. R. A. Purification and characterisation of the 7S globulin storage protein from sesame (*Sesamum indicum* L.). *Food Chem.* **2007**, *100*, 926−934.

(16) Barba de la Rosa, A. P.; Fomsgaard, I.; Laursen, B.; Mortensen, A.; Olvera-Martínez, L.; Silva-Sánchez, C.; Mendoza-Herrera, A. González-Castañeda, J.; De León-Rodríguez, A. Amaranth (*Amaranthus hypochondriacus*) as an alternative crop for sustainable food production: Phenolic acids and flavonoids with potential impact on its nutraceutical quality. *J. Cerral Sci.* **2009**, *49*, 117−121.

(17) Martin, R. G.; Ames, B. M. A method for determining the sedimentation behavior of enzymes: application to protein mixtures. *J. Biol. Chem.* **1961**, *236*, 1372−1379.

(18) Erickson, H. P. Size and shape of protein molecules at the nanometer level determined by sedimentation, gel filtration, and electron microscopy. *Biol. Proc. Online*, **2009**, *11* (1). DOI: 10.1007/s12575-009-9008-x.

(19) Laemmli, U. K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **1970**, *227*, 680−685.

(20) Petruccelli, S.; Añón, M. C. Relationship between the method of obtention and the structural and functional properties of soy protein isolates. 1. Structural and hydration properties. *J. Agric. Food Chem.* **1994**, *42*, 2161−2169.

(21) Tessier, D.; Yclon, P.; Jacquemin, I.; Larré, C.; Rogniaux, H. OVNIp: an open source application facilitating the interpretation, the validation and the edition of proteomics data generated by MS analyses and de novo sequencing. *Proteomics* **2010**, *10*, 1794−1801.

(22) Chen, S.; Paredes-López, O. Isolation and characterization of the 11S globulin from amaranth seeds. *J. Food Biochem.* **1997**, *21*, 53−65.

(23) Larkin, M.; Blackshields, G.; Brown, N.; Chenna, R.; McGettigan, P.; McWilliam, H.; Valentin, F.; Wallace, I.; Wilm, A.; Lopez, R.; Thompson, J.; Gibson, T.; Higgins, D. Clustal W and Clustal X version 2.0. *Bioinformatics* **2007**, *23*, 2947−2948.

(24) Ventureira, J. 2010. Propiedades Estructurales y Funcionales de Preparados Proteicos de Amaranto Modificados y Soja-Amaranto, PhD Thesis, Facultad de Ciencias Exactas. Ed. UNLP. Argentina.

(25) Harada, J.; Barker, S.; Goldberg, R. Soybean β-conglycinin genes are clustered in several DNA regions and are regulated by transcriptional and posttranscriptional processes. *The Plant Cell* **1989**, *1*, 415−425.

(26) Domoney, C.; Casey, R. Another class of vicilin gene in *Pisum*. *Planta* **1990**, *182*, 39−42.

(27) O'Kane, F.; Happe, R.; Vereijken, J.; Gruppen, H.; Van Boekel, M. A. Characterization of pea vicilin. 1. Denoting convicilin as the α-subunit of the *Pisum* vicilin family. *J. Agric. Food Chem.* **2004**, *52*, 3141−3148.

(28) Saénz de Miera, L.; Ramos, J.; Pérez de la Vega, M. A comparative study of convicilin storage protein gene sequences in species of the tribe *Vicieae*. *Genome* **2008**, *51*, 511−523.

(29) Sáenz de Miera, L. E.; Pérez de la Vega, M. Evidence that the N-terminal extension of the Vicieae convicilin genes evolved by intragenic duplications and trinucleotide expansions. *Genome* **2001**, *44*, 1022−1030.

(30) Bernardo, A. E. N.; Garcia, R. N.; Adachi, M.; Angeles, J. G. C.; Kaga, A.; Ishimoto, M.; Utsumi, S.; Tecson-Mendoza, E. M. 8S Globulin of Mungbean [*Vigna radiata* (L.) Wilczek]: Cloning and characterization of its cDNA isoforms, expression in *Escherichia coli*, purification and crystallization of the major recombinant 8S isoform. *J. Agric. Food Chem.* **2004**, *52*, 2552−2560.

(31) Tang, C.; Sun, X. Physicochemical and structural properties of 8S and/or 11S globulins from mungbean [*Vigna radiata* (L.) Wilczek] with various polypeptide constituents. *J. Agric. Food Chem.* **2010**, *58*, 6395−6402.

Received for review August 24, 2010. Revised manuscript received November 9, 2010. Accepted November 15, 2010.