Purification, Crystallization, and Preliminary X-ray Characterization of a 36 kDa Amaranth Globulin

Nora L. Vasco-Méndez,[†] Manuel Soriano-García,*[,‡] Abel Moreno,[‡] Rutilo Castellanos-Molina,[†,§] and Octavio Paredes-López*[,∥]

Departamento de Biotecnología y Bioingeniería, Centro de Investigación y de Estudios Avanzados del IPN, Unidad Zacatenco, Avenida Instituto Politécnico Nacional 2508, Col. San Pedro Zacatenco, 07360 Mexico, D.F., Departamento de Bioestructura, Instituto de Química, Circuito exterior s/n, Ciudad Universitaria, UNAM, 04510 Mexico D.F., and Departamento de Biotecnología y Bioquímica, Centro de Investigación y de Estudios Avanzados del IPN, Unidad Irapuato, Apartado Postal 629, Irapuato, Gto. 36500, Mexico

The purpose of this study was to purify, crystallize, and characterize by X-ray diffraction an amaranth globulin for its subsequent structure elucidation. A 36-kDa amaranth globulin was extracted by sequential precipitation and purified by gel filtration and cationic exchange columns. It was crystallized at 18 °C from 4 M sodium formate. Suitable crystals for X-ray analysis were found to belong to the tetragonal crystal system with cell dimensions of $a = b = 195.5$ Å and $c = 164.14$ Å. Two possible tetragonal space groups $P4_12_12$ or $P4_32_12$ were determined. The crystals diffracted up to 2.5 Å.

Keywords: *Amaranth; globulin; storage protein; crystallization*

INTRODUCTION

Storage proteins are of particular importance for human beings because of the high protein content of the seeds. These can be used as an important source of dietary protein for developing societies. Because of their high level of expression and accumulation, storage proteins determine the nutritional value of the seed and dictate the requirements for supplements when the seed is used as foodstuff for man or domestic animals (Lawrence et al., 1990).

It is well-known that amaranth seed has protein with a better balance content of essential amino acids than that of cereals and legumes. The lysine content of the amaranth species is relatively high (3.2−6.4%) compared with the most common cereals (2.2−4.5%), and the sulfur amino acids concentration (2.6−5.5%) is higher than in the most important legumes (1.4%) (Paredes-López et al., 1990; Segura-Nieto et al., 1994; Gorinstein et al., 1996). In fact, its balance is closer to the optimum required amount for humans (FAO/WHO, 1973).

Genetic engineering is a good way to improve grain crops; nutritional value and functional properties could be altered according to human requirements, but these improvements have been hampered by the lack of detailed three-dimensional structural information (Lawrence et al., 1990, 1994). There are two cloning examples of amaranth proteins. The first was published by Raina and Datta (1992); they cloned a gene encoding a well-balanced amino acid albumin. The second work was done by Barba de la Rosa et al. (1996) with a globulin 11S-like protein, also with good nutritional value. Amaranth proteins were proposed by both groups to isolate the corresponding genes to be used for improvements of amino acid deficiencies in many seed proteins. However, it is also important to know the three-dimensional structure of these proteins, as basic information for the achievement of engineered mutants with novel properties in their seeds without damage to any other biological function.

Structure elucidation of storage proteins is seen as a fundamental requirement for the rational design of structural functional mutants with enhanced nutritional and functional properties (McPherson, 1992; Utsumi, 1992; Shewry, 1995). A very important step before the three-dimensional structure elucidation of a protein is the crystallization; some storage proteins have been crystallized, but only a few of them have been resolved because of the exceptional quality and size that is needed to carry out high-resolution X-ray diffraction analysis (McPherson, 1992). Although the 11S globulin of the Brazil nut was crystallized in 1859 by Maschke [cited by Shewry et al. (1995)], this crystal and other 11S globulin crystals are generally small and disordered and have failed to provide any detail of the protein's structure. Related to amaranth globulin structure, Marcone et al. (1994), using electron microscopy and computer image analysis techniques, found that the most abundant storage globulin was arranged into two annular hexagonal rings; this study constitutes the unique structure evidence for amaranth globulins. Here we present the purification, partial biochemical characterization, crystallization, and preliminary X-ray characterization of an amaranth globulin crystal.

MATERIALS AND METHODS

Plant Material. *Amaranthus hypochondriacus* (Mercado cultivar) was obtained from the Instituto Nacional de Investigaciones Forestales y Agropecuarias (INIFAP), Chapingo,

* Authors to whom correspondence should be addressed.
[†] Departamento de Biotecnología y Bioingeniería.
[‡] Departamento de Bioestructura (fax 52-5-616-2203; e-mail soriano@servidor.unam. mx).
[§] Deceased November 1997.
[∥] Departamento de Biotecnología y Bioquímica (fax 52-462-45846; e-mail oparedes@irapuato.ira.cinvestav.mx).

10.1021/jf9809131 CCC: $18.00
Published on Web 02/09/1999

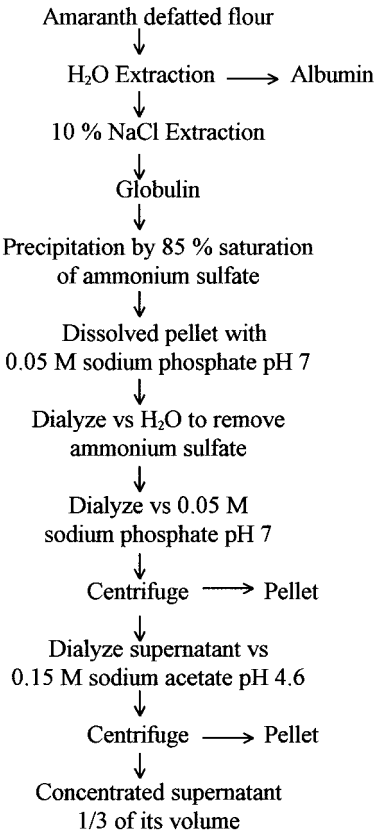

Figure 1. Protein fractionation scheme.

Mexico. Seed meal was obtained by grinding whole seeds at 6 °C in a mill (Tekmar, Germany) attaining an average particle size of ~60 mesh. The sample was defatted by extraction with an *n*-hexane suspension (1:10, w/v) under continuous stirring for 24 h. The defatted meal was separated by centrifugation at 8000*g* for 20 min at 4 °C. The flour was then dried at room temperature and stored in a hermetic bowl at 4 °C until it was used.

Protein Fractionation. Globulin was extracted according to Osborne's (1924) sequential method and fractionated according to a modification of the Blagrove and Gillespie (1975) method. Albumin was first extracted with distilled water (1:10, w/v) by two stirring steps of 1.5 h at 4 °C and then centrifuged. The pellet was then resuspended and extracted overnight as the first stirring step with 10% NaCl with the same indicated ratio; the second step was for 1 h. After centrifugation, the supernatants of both steps were pooled. Globulin fraction was precipitated with saturated (85%) ammonium sulfate ($(NH_4)_2SO_4$) and then centrifuged, and the pellet was resuspended with 0.05 M sodium phosphate buffer at pH 7. The concentrated globulin fraction was dialyzed at 4 °C against distilled water to remove ammonium sulfate. A second 4 °C dialysis was done against 0.05 M sodium phosphate buffer at pH 7. Insoluble particles were removed by centrifugation, and the supernatant was again dialyzed but this time against 0.15 M sodium acetate buffer at pH 4.8 for 2 days. After this time, the precipitate was removed by centrifugation. The supernatant was dialyzed against distillate water at pH 7 overnight, changing the water at pH 7 several times. Finally, the precipitate was again removed by centrifugation and the supernatant concentrated by ultrafiltration up to one-third of its volume using a stirred ultrafiltration cell (Amicon Division, Beverly, MA) with a cutoff membrane of 10 kDa (Figure 1). This concentrated globulin subfraction, which is soluble in water after several pH precipitations, were then chromatographed. All dialysis processes were made at 4 °C using cellulose membranes with 8000 PM of cutoff. Each centrifuged step was performed at 8000g for 20 min at 6 °C, and any buffer and water used were added with 0.02% sodium azide (NaN_3) to prevent protein contamination.

Protein Purification. The concentrated globulin subfraction was purified by gel filtration and cation exchange chromatography. Samples of <8 mL with variable concentration between 8 and 15 mg/mL were applied on a filtration column. Gel filtration was done using a Sephacryl S-200 superfine gel (Pharmacia Biotech, Uppsala, Sweden) column (2.0 × 80.0 cm) equilibrated with 0.05 M sodium phosphate buffer at pH 7 containing 0.02% sodium azide (NaN_3), with a molecular weight fractionation between 10 and 80 kDa. A standard curve was obtained using the low molecular weight gel filtration kit (Pharmacia Biotech), which contains blue dextran 2000 (2000 kDa) for void volume, bovine serum albumin (BSA, 67 kDa), ovalbumin (43 kDa), chymotrypsinogen A (25 kDa), and ribonuclease A (13.7 kDa). Buffer was eluted from the column at room temperature with a flow rate of 12 mL/h; 5 mL fractions were collected, and the absorbance at 280 nm was measured. Protein peaks were separately pooled, concentrated, and dialyzed against 0.5 M sodium acetate buffer at pH 5 and at 5 °C for further purification by a strong cation exchange column using an Econo-Pac S Ion Exchange 5 mL cartridge (Bio-Rad, Richmond, CA). Samples containing ≤15 mg were applied on the cationic column, and buffer (0.5 M sodium acetate at pH 5) was eluted at a flow rate of 30 mL/h at room temperature; a gradient of sodium salt (NaCl) ranging from 0 to 1 M was used. Two milliliter fractions were collected, and 280 nm absorbance was measured. Peaks were then individually pooled, concentrated, and dialyzed exhaustively against 0.5 M sodium acetate buffer at pH 5 at 4 °C using a stirring ultrafiltration cell (Amicon Division) with a cutoff membrane of 10 kDa. Each peak was again rechromatographed by a cation exchange column under the same conditions, pooled, concentrated, and dialyzed exhaustively against water.

Gel Electrophoresis. Nondenaturing gels were used to examine the native protein. This was carried out in a Mini-Protean II electrophoresis cell (Bio-Rad), using 7% polyacrylamide gel. Native gel was loaded with 50−60 μg of protein by lane. Sodium dodecyl sulfate−polyacrylamide gel electrophoresis (SDS−PAGE) was carried out using the Phast System apparatus (Pharmacia Biotech) according to the Laemmli (1970) method. SDS−PAGE with or without β-mercaptoethanol (β-ME) was run on 12.5% gels according to separation method 2 using the separation technique file 111 of the Phast System user's manual. All SDS−PAGE with or without β-ME was overloaded with 80−90 μg to observe the protein impurities except SDS−PAGE−βME, which was used to compare the purified 36 kDa globulin before (3 μg) and after crystallization. Molecular weight standards (low molecular weight calibration kit, Pharmacia Biotech) were used to estimate the protein molecular weight. The gels were stained with Coomassie brilliant blue R-250.

Isoelectric Focusing. Isoelectric focusing was performed by the technique described in Phast System technical file 100 using IEF 3-9 gels (Pharmacia Biotech). The isoelectric focusing method contains three steps: a prefocusing step, in which the pH gradient is generated; a sample application step, in which ~4 μg was loaded per lane; and a focusing step, in which the proteins are separated according to their charge. Isoelectric focusing calibration broad p*I* kit (pH 3−9, Pharmacia Biotech) was used to calculate the protein isoelectric point. The gel was stained with Coomassie brilliant blue R-250.

Crystallization. Suitable crystals for X-ray diffraction were obtained at 18 °C by vapor diffusion technique of 6 μL hanging drops against 750 μL of the precipitating agent in a linbro plate well sealed with vacuum grease. Each drop was made by mixing 3 μL of amaranth globulin solution of 10 mg/mL and 3 μL of the precipitating agent. The crystal screen sparse matrix kit (Hampton Research) was used to search for the initial crystallization conditions of purified proteins.

X-ray Diffraction. All diffraction data were collected on a RAXIS−IIC image plate area detector mounted on a Rigaku RU-200 rotating-anode X-ray generator operated at 50 kV and 100 mA. Copper Kα radiation was used with a 0.3 mm focus cup and a 0.6 mm collimator. The crystal detector distance

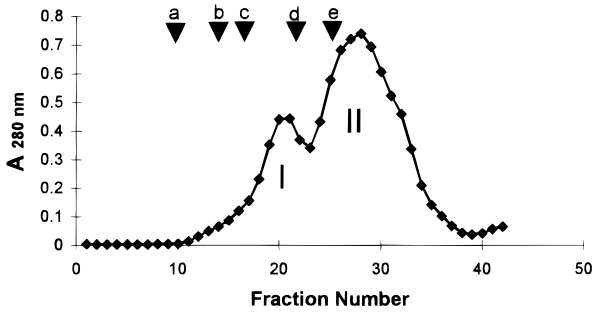

Figure 2. Gel filtration elution pattern of crude globulin subfraction after sequential precipitation of total globulin. Peak I eluted approximately at 36 kDa and peak II around 12 kDa. Elution positions of blue dextran and standard proteins are indicated by arrows: a, blue dextran (M_r 2000000); b, bovine serum albumin (M_r 67000); c, ovoalbumin (M_r 43000); d, chymotrypsinogen A (M_r 25000); e, ribonuclease A (M_r 13700).

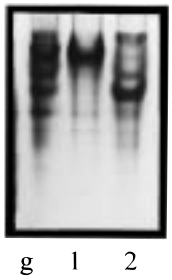

Figure 3. Native PAGE: lane g, crude globulin subfraction before separation on gel filtration column; lane 1, peak I; lane 2, peak II.

was 160 mm, and a data collection time of 30 min per frame was used. Data were reduced and merged using the RAXIS software.

RESULTS AND DISCUSSION

Protein Fractionation and Purification. The final globulin subfraction that is soluble in water, after total globulin fractionation, by sequential precipitation was chromatographed on a gel filtration column. The globulin subfration eluted into two wide peaks (Figure 2). Peak I eluted around 36 kDa and peak II at ~12 kDa. Each peak was pooled and concentrated separately. A native elecrophoresis was carried out to compare the initial globulin subfraction with each peak separately (Figure 3). The globulin subfraction pattern on lane g shows four well-separated major bands; peak I on lane 1 shows only one wide band, which corresponds to the major molecular weight (36 kDa) of the original globulin subfraction. The electrophoretic pattern on lane 2 shows three bands of low molecular weight that eluted in peak II; one of them (~12 kDa) appears to be in larger quantities. When the fractions were run on SDS−PAGE with (Figure 4) and without (Figure 5) reduction by β-ME, the same native pattern was recognized. As can be seen in both gels, a globulin of ~36 kDa in peak I and a globulin of ~12 kDa in peak II are monomeric proteins because the same pattern bands are present under native and denatured conditions. Those proteins appear similar to that described by Martínez et al. (1997), as free subunits of 16 and 39 kDa found in lower sedimentation constant species during separation of 7S and 11S. Each peak was concentrated separately and dialyzed against 0.1 M sodium acetate buffer at pH 5 in a stirred ultrafiltration cell to be chromatographed

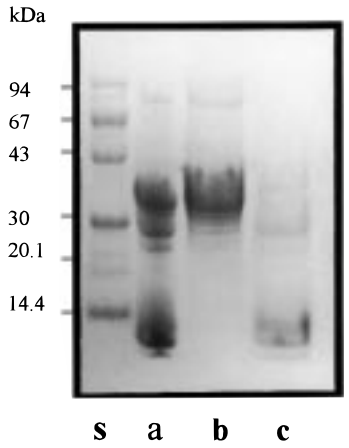

Figure 4. SDS−PAGE with reduction by β-ME: lane a, crude globulin subfraction; lane b, peak I separated by gel filtration column; lane c, proteins eluted in peak II; lane s, standard molecular masses [phosphorylase *b* (94 kDa), BSA (67 kDa), ovalbumin (43 kDa), carbonic anhydrase (30 kDa), soybean trypsin inhibitor (20.1 kDa), α-lactoalbumin (14.4 kDa)].

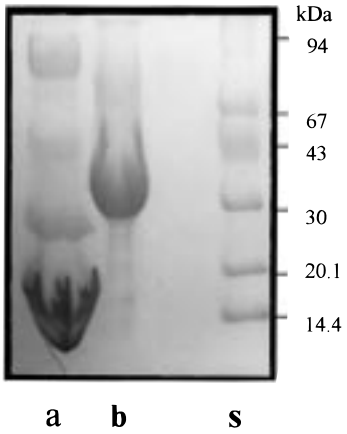

Figure 5. SDS−PAGE without reduction: lane a, proteins eluted in peak II; lane b, proteins eluted in peak I; lane s, standard proteins.

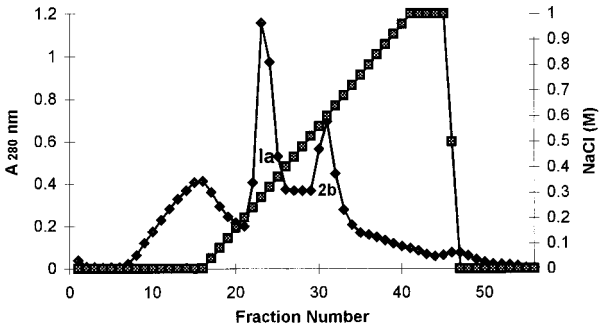

Figure 6. Elution profile of peak I in a cationic exchange column. Peak Ia eluted at ~0.25 M NaCl and peak 2b at ~0.5 M NaCl.

on a cationic exchange column. Only one band, in peak I, was separated into three peaks (Figure 6): a small peak that passes freely throughout the column; a second peak (1a), the most abundant, which is released at ~0.25 M of the gradient; and a third peak (2b) released near 0.55 M of the NaCl gradient. Peak 1a was pooled, concentrated, and dialyzed against sodium acetate buffer to be rechromatographed under the same conditions. The uniquely sharp peak was rechromatographed and then concentrated and dialyzed against distillate water until it reached 10 mg/mL. The same procedure

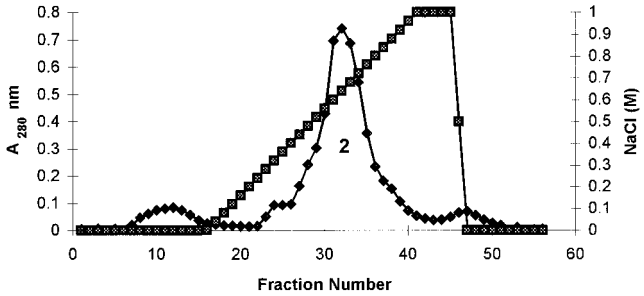

Figure 7. Elution profile of peak II in a cationic exchange column. Peak 2 eluted near 0.6 M NaCl gradient.

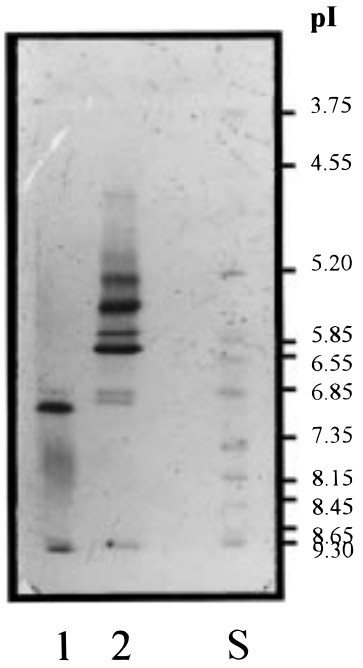

Figure 8. Isoelectric focusing on polyacrylamide gel of peaks 2 and 1a after a cationic exchange column: lane 1, peak II, which contains the 12 kDa globulin; lane 2, peak 1a, which contains the 36 kDa globulin; lane S, isoelectric focusing standards [methyl red (p*I* 3.75), soybean trypsin inhibitor (p*I* 4.55), β-lactoglobulin A (p*I* 5.20), bovine carbonic anhydrase B (p*I* 5.85), human carbonic anhydrase B (p*I* 6.55), horse myoglobin acidic (p*I* 6.85), horse myoglobin basic (p*I* 7.35), lentil lectin acidic (p*I* 8.15), lentil lectin middle (p*I* 8.45), lentil lectin basic (p*I* 8.65), trypsinogen (p*I* 9.30)].

was done for peak II; the cationic exchange pattern (Figure 7) shows a wide peak (peak 2), which eluted near 0.6 M of the NaCl gradient; it was then dialyzed against water and concentrated up to 10 mg/mL. Both globulins were isoelectric focused (Figure 8) to determine their isoelectric points. As can be seen on lane I, globulin of 12 kDa is practically pure and it is a neutral protein because it possesses an isoelectric point of pH 7. On the other hand, peak 1a (36 kDa globulin) does not have a pure protein, due to the presence of four major bands, with isoelectric points of 6.3, 6.0, 5.7, and 5.3 pH, and other two small bands at about p*I* 6.9 and 6.8 pH, all of which are acidic proteins; their relative quantities are 6.3 > 6.0 > 5.7 > 5.3 > 6.9 > 6.8. Because the six bands shown in Figure 8 have the same molecular weights but different p*I* values, we think that they are isoforms with the same structure; this type of polymorphism is characteristic of storage proteins (Shewry, 1995). Globulins of 36 and 12 kDa were assayed to find initial crystallization conditions. Although the 36 kDa globulin was not completely pure, it

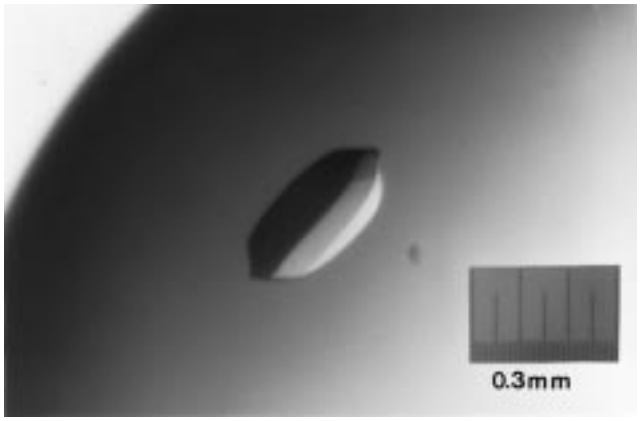

Figure 9. Crystal of 36 kDa globulin from *A. hypochondriacus*, 11× magnification.

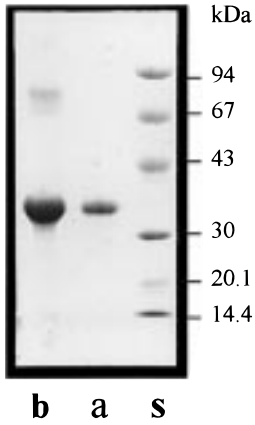

Figure 10. SDS−PAGE with β-ME of 36 kDa protein: lane a, protein after crystallization; lane b, protein before crystallization; lane s, standard proteins.

was possible to crystallize it and thereafter to obtain data for the preliminary crystal characterization. In addition, further experiments are still in progress with the 12 kDa globulin to try to achieve its crystallization.

Crystallization. Small colorless crystals of 36 kDa globulin were observed after 5 days. The precipitant agent was 4 M sodium formate (reagent 33, Hampton Research). An Izit test (Hampton Research) was done to verify the protein nature. The maximal crystal growth was ~3 weeks with dimensions of 0.3 × 0.2 × 0.15 mm (Figure 9). An SDS−PAGE with β-ME was carried out to compare the purified 36 kDa protein before crystallization and with the protein crystals formed. Lane b, in Figure 10, shows the protein after crystallization, and lane a shows the protein before it was crystallized; the same principal band is observed (36 kDa) in both cases. Many of these crystals were mounted in an X-ray capillary tube with an internal diameter of 0.5 mm to be diffracted at room temperature. Additionally, a 0.5 mm loop was used to perform the data acquisition at −160 °C. A total of 636653 data were measured with an R_{merge} of 0.11. Of these 636653 reflections, only 79582 were unique data corresponding to a completeness of 80%. The crystal belongs to the tetragonal system with two possible space groups, $P4_12_12$ or $P4_32_12$, with cell dimensions $a = b = 195.5$ Å and $c = 164.14$ Å. The unit cell is as large as some virus capside proteins (Koszelak et al., 1989). Similar cell dimensions have been found for phaseolin ($a = 126.8$ Å, $b = 134.5$ Å, $c = 161.6$ Å), a 7S globulin from *Phaseolus vulgaris* (Lawrence et al.,

1990), canavalin ($a = b = 136.8$ Å, $c = 75.7$ Å), a 7S globulin from *Canavalia ensiformis* (Ko et al., 1993), and edestin ($a = b = 215$ Å, $c = 80$ Å), an 11 S globulin from hemp seeds (Patel et al., 1994). An estimate of the number of molecules per unit cell can be made according to the Matthews (1968) method. Knowing the unit cell volume (6.27×10^6 Å^3), space group symmetry (eight asymmetric units for unit cell), and experimental relative weight of the monomer (~36 kDa), there are eight hexamers by unit cell. The solvent amount is 64% and 36% of protein; these percentages were calculated using the Matthews volume of 3.6 Å^3/Da.

LITERATURE CITED

Barba de la Rosa, A. P.; Herrera-Estrella, A.; Utsumi, S.; Paredes-López, O. Molecular characterization, cloning and structural analysis of a cDNA encoding an amaranth globulin. *Plant Physiol.* **1996**, *149*, 527−532.

Blagrove, R. J.; Gillespie, J. M. Isolation, purification and characterization of the seed globulins of *Lupinus angustifolius*. *Aust. J. Plant Physiol.* **1975**, *2*, 13−27.

FAO. *Energy and Protein Requirements*; Nutrition Meeting Report Series *52*; Food and Agriculture Organization: Rome, 1973.

Gorinstein, S.; Zemser, M.; Friedman, M.; Vasco-Méndez, N. L.; Paredes-López, O. Denaturant-induced conformations of globulins. *J. Agric. Food Chem.* **1996**, *44*, 93−99.

Ko, T.-P.; Ng, J. D.; McPherson, A. The three-dimensional structure of canavalin from jack been (*Canavalia ensiformis*). *Plant Physiol.* **1993**, *101*, 729−744.

Koszelak, S.; Dodds, A.; McPherson, A. Preliminary analysis of crystals of satellite tabacco mosaic virus (STMV). *J. Mol. Biol.* **1989**, *208*, 323−326.

Laemmli, U. K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **1970**, *217*, 680−685.

Lawrence, M. C.; Suzuki, E.; Varghese, J. N.; Davis, P. C.; VanDonkelaar, A.; Tulloch, P. A.; Colman, P. M. The three-dimensional structure of the seed storage protein phaseolin at 3 Å resolution. *EMBO J.* **1990**, *9*, 9−15.

Lawrence, M. C.; Izard, T.; Beauchat, M.; Blagrove, R. J.; Colman, P. M. Structure of Phaseolin at 2.2 Å Resolution. Implication for a common vicilin/legumin structure and the genetic engineering of seed storage proteins. *J. Mol. Biol.* **1994**, *238*, 748−776.

Marcone, M. F.; Beniac, D. R.; Harauz, G.; Yada, R. Y. Quaternary structure and model from the oligomeric seed globulin from *Amaranthus hypochondriacus* K343. *J. Agric. Food Chem.* **1994**, *42*, 2675−2678.

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

Matthews, B. W. Solvent content of protein crystals. *J. Mol. Biol.* **1968**, *33*, 491−497.

McPherson, A. The role of protein crystals in biotechnology and industry. In *Frontiers in Bioprocessing II*; Conference Proceedings Series; Todd, P., Sikdar, S. K., Bier, M., Eds.; American Chemical Society: Washington, DC, 1992; pp 10−15.

Osborne, T. B. *The Vegetable Proteins*, 2nd ed.; Longmans, Green: New York, 1924.

Paredes-López, O.; Barba de la Rosa, A. P.; Hernández-López, D.; Carabez-Trejo, A. *Amaranto-Caracteristicas Alimentrarias y Aprovechamiento Agroindustrial*; OEA: Washington, DC, 1990.

Patel, S.; Cudney, R.; McPherson, A. Crystallographic characterization and molecular symmetry of edestin, a legumin from hemp. *J. Mol. Biol.* **1994**, *235*, 361−363.

Raina, A.; Datta, A. Molecular cloning of a gene encoding a seed-specific protein with nutritionally balanced amino acids composition from amaranthus. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 11774−11778.

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; Chapter 8, pp 76−95.

Shewry, P. R. Plant storage proteins. *Biol. Rev.* **1995**, *70*, 375−426.

Shewry, P. R.; Napier, J. A.; Tatham, A. S. Seed storage proteins: Structure and biosynthesis. *Plant Cell.* **1995**, *7*, 945−956.

Utsumi, S. Plant food protein engineering. *Adv. Food Nutr. Res.* **1992**, *36*, 89−208.

Received for review August 13, 1998. Accepted December 8, 1998. N.L.V.-M. thanks CONACYT-México for a fellowship and acknowledges M.S.-G. and O.P.-L. for all of the facilities to complete this research. M.S.-G. thanks DGAPA-UNAM, Project 000000.

JF9809131