Large-Scale Chromatography for the Isolation of 7S Globulin Enriched Fraction from Pigeon Pea Seeds

Journal of Agricultural and Food Chemistry

Locali-Pereira, Adilson Roberto; Solé-Jamault, Véronique; Davy, Joëlle; Cherkaoui, Mehdi; Rogniaux, Hélène et al

https://doi.org/10.1021/acs.jafc.5c01681

JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

pubs.acs.org/JAFC Article

Large-Scale Chromatography for the Isolation of 7S Globulin Enriched Fraction from Pigeon Pea Seeds

Adilson Roberto Locali-Pereira, Véronique Solé-Jamault,* Joëlle Davy, Mehdi Cherkaoui, Hélène Rogniaux, Hamza Mameri, Sophie Le Gall, Valérie Beaumal, Hanitra Rabesona, Sophie Laurent, Vânia Regina Nicoletti, Claire Berton-Carabin, and Adeline Boire

Cite This: *J. Agric. Food Chem.* 2025, 73, 11288−11302 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Pigeon pea (*Cajanus cajan*) contains around 60% globulin proteins, of which 7S globulin is the most abundant fraction. In this work, a purification protocol for pigeon pea 7S globulin was developed using large-scale chromatography. The process was designed on an analytical scale through desalting of the crude protein extract followed by anion exchange chromatography and size exclusion chromatography. Then, the process was scaled up to a large scale. The purified fraction presented a protein content of 89 g/100 g powder and yield close to 23%. The isolated 7S globulin showed two main subunits of 64 and 49 kDa, identified as the α- and β-chains of β-conglycinin by proteomic analysis. The developed protocol was shown to be suitable for purifying pigeon pea 7S globulin on a large scale, and it is relevant for isolating this fraction from other pulse seeds in sufficient quantities for their characterization and evaluation of functional properties.

KEYWORDS: *pulse crops, Cajanus cajan, plant-based proteins, 7S vicilin, anion exchange chromatography, size exclusion chromatography*

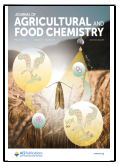
1. INTRODUCTION

The 7S globulin is the main protein fraction found in pigeon peas (*Cajanus cajan*), accounting for about 60% of its total protein composition.[1,2] Liquid chromatography techniques enable the isolation of the 7S globulin fraction from pigeon pea seeds, which has been investigated in analytical-scale procedures in a few studies. Krishna and co-workers were pioneers in investigating the composition of pigeon pea proteins using extraction methods followed by precipitation and/or purification by analytical chromatography.[2−4] In their first study, pigeon pea proteins were extracted using solubilization/centrifugation/precipitation cycles,[4] which led to the identification of vicilin- and legumin-type proteins. Then, they developed a purification protocol for pigeon pea 7S vicilin on an analytical scale.[4] First, a crude extract of globulins was obtained by alkaline extraction followed by isoelectric precipitation. 7S globulin was then purified using zonal isoelectric precipitation and injection on a weak anion exchange column. According to the authors, 7S globulin has a molecular weight of around 184 kDa in its native state, with a 70 kDa subunit and two 57 kDa subunits. More recently, other studies have isolated pigeon pea globulins[5,6] based on differential solubility and/or ammonium sulfate precipitation. While the purification of globulin proteins has been a crucial step in elucidating their composition and structure, the analytical-scale procedures reported in the literature yielded pigeon pea globulins on a milligram scale due to the working scale used. According to Gravel and Doyen,[7] the purity and yield of purified pulse globulins are often poorly described by authors in the literature, but usually, 11S and 7S globulins yields vary from 100 mg purified globulin per 10 g of flour to 800 mg per 4.2 g of protein isolate (1−19%), depending on the starting material and method used. Traditional extraction methods, such as salt precipitation, isoelectric precipitation, solvent extraction, and dialysis, have been commonly employed to isolate pigeon pea proteins. In a previous study by our group, pigeon pea protein concentrates were obtained using the alkaline solubilization method followed by isoelectric precipitation, resulting in an extraction yield of around 12% (mass of protein concentrate per mass of flour) and protein content ranging from 62 to 68%.[8] These methods often result in low yield and purity due to non-selective protein precipitation and coextraction of unwanted compounds such as organic acids and polyphenol compounds.[9−11] In addition, precipitation methods have resulted in modifications to the protein structures.[12]

In this study, fractionation by chromatography was chosen to obtain native protein fractions with a high purity. Besides, milder processes, which do not use protein precipitation steps, have not been investigated and are more suitable for evaluating the functionality of proteins. Moreover, to probe the functionality of purified proteins, it is necessary to purify them at the gram scale by upscaling from analytical- to large-scale chromatography.[13,14]

Large-scale chromatography can be defined as a unity operation used to isolate a compound in adequate purity and

Received: February 7, 2025
Revised: April 11, 2025
Accepted: April 15, 2025
Published: April 24, 2025



quantity for subsequent processes or experiments.[15] Chromatographic methods are commonly used to isolate 7S fractions from globulin protein concentrates or isolates,[7] and procedures on a large scale have been proposed for pea,[14,16,17] rapeseed,[18] and bean globulins.[19] The chromatographic steps may include (i) desalting of the crude protein extract to remove small molecules that can form complexes with proteins;[18] (ii) ion exchange chromatography, in particular anion exchange chromatography, which is most used for globulin fractionation as it is suitable for proteins with a acid isoelectric point;[7,14] and (iii) gel filtration separation, which allows separating proteins based on molecular size, removing aggregates and minor compounds. Dialysis/ultrafiltration and gel filtration of protein fractions are commonly used before freeze-drying to remove salt and increase protein concentration.[7]

Therefore, we developed a large-scale purification procedure for pigeon pea 7S globulin using a mild protein extraction process and the following chromatographic steps: desalting of the protein extract, separation on an anion exchange column, and subsequent separation on a gel filtration column. The purification stages were initially optimized on an analytical scale and then increased to a large scale. The biochemical properties of the resulting purified fractions were determined using sodium dodecyl sulfate−polyacrylamide gel electrophoresis (SDS-PAGE), fluorescence intrinsic measurements, and proteomic analysis.

2. MATERIALS AND METHODS

2.1. Materials. Pigeon pea seeds (*Cajanus cajan*) were purchased from a local supplier (Sao Paulo, Brazil). The Tris-glycine SDS running buffer and Tris base were supplied by ThermoFisher (Massachusetts, USA). Sodium phosphate, sodium carbonate, sodium chloride, citric acid, MES free acid, ammonium carbonate, potassium phosphate monobasic, disodium phosphate, sodium hydroxide, hydrochloric acid, and 2-β-mercaptoethanol were supplied by Sigma-Aldrich (Missouri, USA). Acetic acid was provided by Carlo Erba (Emmendingen, Germany). Coomassie Blue G250 staining was supplied by Serva (Heidelberg, Germany).

2.2. Protein Composition of Pigeon Pea Flour by SE-HPLC. The protein composition of pigeon pea flour was investigated by size exclusion high-performance liquid chromatography (SE-HPLC) under denaturing conditions.

To produce the flour (<300 μm), pigeon pea seeds were first ground using a disk mill (J.R. Araújo & Cia Ltda., Brazil) at 1730 rpm for 5 min. Proteins were separated according to their molecular weight and quantified by SE-HPLC following the protocol described by Morel et al. (2000) with some modifications.[20] First, 10 mg of flour or protein powder was used for protein extraction in 1 mL of 0.1 M sodium phosphate buffer at pH 6.8 containing 1% SDS. Extraction was performed on a rotary shaker (Heidolph Reax) (60 rpm for 60 min at 20 °C), and the supernatant with SDS-soluble proteins was recovered after centrifugation (20,817*g*, 5 min, 20 °C). An additional extraction was performed from the pellet to recover SDS-insoluble proteins linked by disulfide bonds using the previous solvent added with a thiol reducing agent, dithioerythritol (DTE) at 20 mM. The SE-HPLC apparatus (Schimadzu) was equipped with an analytical column, TSK G4000-SWXL (Tosoh bioscience) (7.8 × 300 mm), and a TSK gel SW XL precolumn (4 cm × 6 mm) (Merck, Darmstadt, Germany). A volume of 20 μL was loaded and eluted at 25 °C with the 0.1 M sodium phosphate buffer at pH 6.8 and SDS 0.1% with a flow rate of 0.7 mL/min. The absorbance spectra were recorded from 190 to 800 nm. Apparent molecular weights (MWs) were assessed with column calibration with protein standards with known MWs. The composition of SDS-soluble and SDS-DTE-soluble proteins was also investigated by SDS-PAGE, described in Section 2.6.4.

2.3. Extractability of Pigeon Pea Proteins. About 60 mg of flour was dispersed in 3 mL of each of the following buffer solutions (concentration = 20 mM): acetate buffer (pH 4.0, pK_a = 4.76), citrate buffer (pH 6.0, pK_a = 6.39), phosphate buffer (pH 8.0, pK_a = 7.2), and carbonate buffer (pH 10.0, pK_a = 10.25). Each buffer solution was evaluated for low (50 mM) and high (500 mM) ionic strength using sodium chloride if necessary to achieve the desired ionic strength. For the pH 6.0 buffer solution at low ionic strength, MES free acid (pK_a = 6.21) was used instead of citric acid due to the high contribution of citric acid to the ionic strength. The suspensions were agitated for 1 h at 20 rpm. After agitation, the tubes were centrifuged (10,000*g*, 10 min, 20 °C), and the supernatant was collected. The nitrogen content in the protein extracts was determined by Dumas analysis; protein composition was assessed by SDS-PAGE, UV−visible absorbance spectra, intrinsic fluorescence, and dynamic light scattering (DLS) (methodology presented in Section 2.6).

2.4. Laboratory-Scale Purification of Pigeon Pea 7S Globulin. *2.4.1. Removal of Small Molecules from the Protein Extract.* To remove small molecules from protein extracts, two approaches were used: membrane ultrafiltration and injection into a desalting column. First, 100 mg of pigeon pea flour was dispersed in 5 mL of 50 mM Tris buffer at pH 8.0. The mixture was stirred for 1 h and centrifuged (10,000*g*, 10 min, 20 °C) with supernatant collection. A 5 mL aliquot of buffer was added to the pellet, and a new agitation and centrifugation were carried out followed by collection of the supernatant. Finally, the two supernatants were pooled.

For ultrafiltration, 5 mL of protein extract was transferred to a membrane ultrafiltration tube (Amicon Ultra Centrifugal Filters, 3K, Regenerated Cellulose, 3000 NMWL). The tube was centrifuged (5000*g*, 5 min, 20 °C), and the filtrate was collected. The retentate was added with the buffer (up to 5 mL), and two additional centrifugations were performed with subsequent collections of the filtrates and the retentate. The three filtrates were pooled. At the end of the process, the retentate and filtrate were evaluated by fluorescence measurements (as described in Section 2.6.7).

For the desalting step, a desalting column (HiPrep 26/10, Cytiva) equilibrated in Tris 50 mM at pH 8.0 was used at 12 mL/min. The injection was performed in two 10 mL runs, and the volume of the collected fraction was 15 mL. The separated peaks were collected and evaluated by fluorescence measurements in the same way as for the ultrafiltered samples.

2.4.2. Optimization of Fractionation Conditions by Anion Exchange and Size Exclusion Chromatography. For separation by anion exchange chromatography (AEC), the ultrafiltered and desalted extracts were injected into a Q Sepharose FF column (HiTrap IEX Selection kit, GE HealthCare) coupled to an Akta Avant 25 chromatography system (GE HealthCare). Samples were injected at 1 mL/min, and elution was conducted on a NaCl gradient (step 1 = 0 to 0.75 M in 20 mL and then step 2 = 0.5 to 1 M in 5 min). The absorbance of the eluted proteins was recorded at 280 nm. The collected fractions were evaluated by polyacrylamide gel electrophoresis for identification of the 7S globulin fraction. Fractions with the same band composition of 7S globulin were pooled. To further improve separation, two approaches were used for the injection of the extracts in the Q Sepharose column: (i) increase of the NaCl elution gradient (step 1: 0 to 0.75 M in 40 mL and step 2: 0.5 to 1 M in 5 min) and (ii) addition of 1 M NaCl solution in the extract to obtain a final concentration of around 100 mM before the injection + increase of the elution gradient of NaCl. For further purification by size exclusion chromatography (SEC), fractions were injected (100 μL) into a Superose 6 column (Cytiva) equilibrated in Tris 50 mM at pH 8.0 and NaCl 150 mM at 0.5 mL/min connected to an HPLC system (1260 Infinity, Agilent). The ultrafiltered and desalted extracts were also evaluated by gel filtration for comparison with the purified 7S globulin.

2.5. Large-Scale Purification of Pigeon Pea 7S Globulin. The purification was conducted on a fast protein liquid chromatography (FPLC) system (Akta pure 150, GE Healthcare). Figure 1 shows the purification process that was developed. First, pigeon pea flour (200 g) was dispersed in 1 L of 50 mM Tris buffer at pH 8.0 and EDTA 5

https://doi.org/10.1021/acs.jafc.5c01681
J. Agric. Food Chem. 2025, 73, 11288−11302

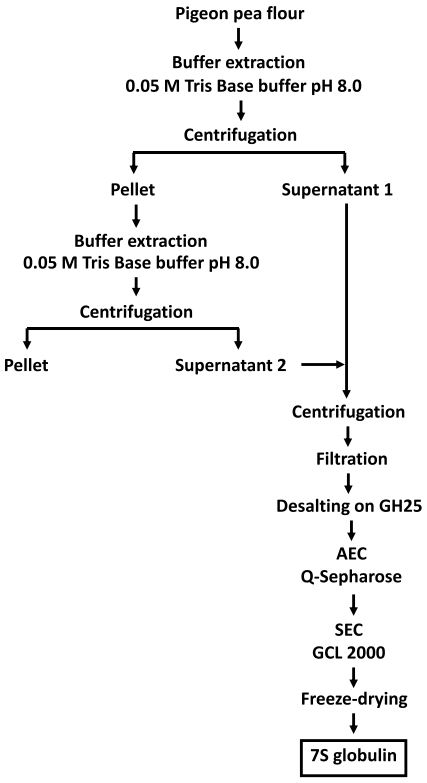


Figure 1. Schematic of the large-scale purification processes of pigeon pea 7S globulin.

mM, and the suspension were stirred for 1 h in a mechanical stirrer (RZR 2041, Heidolph) at 225 rpm. Then, the suspension was centrifuged (10,000*g*, 10 min, 20 °C), and the supernatant was collected. Another 1 L of the buffer was added to the pellet with new agitation and centrifugation with collection of the supernatant. The two supernatants were pooled, centrifuged again (10,000*g*, 40 min, 20 °C) and filtered (8 μm, mixed cellulose ester (MCE) membrane, Millipore). The extract was injected into a desalting column (100 × 887 mm) filled with the Cellufine GH25 cellulose gel (Amicon-Millipore) equilibrated in 50 mM Tris buffer at pH 8.0 to eliminate pigments and minor compounds. The injection was conducted at a flow rate of 50 mL/min. The collected volume of desalted extract was 2.8 L. After collection, 280 mL of 50 mM Tris buffer at pH 8.0 + 1 M NaCl solution were added to the extract.

Then, the desalted extract (around 3.1 L) was injected in a column (50 × 150 mm) filled with Q-Sepharose Fast Flow resin (GE Healthcare Bio-Sciences) at a flow rate of 13 mL/min. The elution was conducted in two steps: 35% of 50 mM Tris buffer at pH 8.0 + 1 M NaCl solution for a time corresponding to an elution volume of 750 mL and 100% of 50 mM Tris buffer at pH 8 + 1 M NaCl solution for a time corresponding to an elution volume of 750 mL. The absorbance of the eluted proteins was recorded at 280 nm, and fractions were collected in 50 mL tubes and evaluated with SDS-PAGE.

For further purification by size exclusion chromatography, protein fractions were injected directly (flow rate = 12 mL/min) onto the Cellufine GCL 2000 gel filtration column equilibrated with ammonium carbonate buffer at pH 8 (1 g/L). Fractions were collected in 50 mL tubes, pooled, and freeze-dried.

Powdered purified fractions were stored in a desiccator containing a K_2CO_3 saturated solution (RH = 43.2%) at 20 °C for equilibration before further analysis.

2.6. Analytical Characterization of the Purified Fraction. *2.6.1. Process Yield and Protein Recovery.* The purification process yield (%) was calculated by the mass ratio between the obtained powdered purified fraction and the initial flour, whereas protein recovery (%) was calculated by the mass of protein recovered in the freeze-dried purified fraction and the mass of protein in the initial flour. The recovery of 7S globulin was calculated by the ratio of the mass of 7S globulin in the purified fraction and in the initial flour based on data from the SE-HPLC analysis (Section 2.2).

2.6.2. Composition. The moisture content was determined using a gravimetric method under a nitrogen atmosphere (TGA 2050, TA Instruments). Samples (20 mg) were heated from 20 to 105 °C (3 K/min) and held at 105 °C until a constant weight.

The protein content was determined by the Dumas method using an elemental analyzer for nitrogen (Vario Micro Cube, Elementar, Frankfurt, Germany) with a nitrogen-to-protein conversion factor of 5.7.[21]

Neutral sugars were determined by gas−liquid chromatography analysis after hydrolysis of the samples to produce alditol acetate derivatives, according to the method described by Le Gall et al. (2016).[22]

The trypsin inhibitor activity (TIA) was determined using the AOCS method Ba 12a-2020 that determines the presence of tryptic inhibitors by a colorimetric reaction in the presence of trypsin and a substrate (Nα-benzoyl-DL-arginine 4-nitroanilide hydrochloride, BAPNA).[23,24]

2.6.3. Proteomic Analysis. Proteomic analysis was performed on the pigeon pea flour and total protein extract (after desalting) and on SDS-PAGE bands of the purified 7S globulin fraction.

For the flour and total protein extract, 5 mg of powder was suspended on 100 μL of the solubilization buffer (50 mM ammonium bicarbonate, ProteaseMAX 0.2% v/v) for 20 min at room temperature followed by an ultrasound bath for 10 min. Then, 2 μL was picked up and diluted in 84 μL of 25 mM ammonium bicarbonate. Protein reduction was performed after addition of 1 μL of 0.5 M DTT solution and a 20 min incubation at 56 °C. Proteins were then alkylated by addition of 2.7 μL of a 0.55 M iodoacetamide solution and incubation in the dark for 15 min. Finally, tryptic digestion was performed by addition of 10 μL of trypsin solution at 200 ng/μL and 1 μL of ProteaseMAX 1% v/v (Promega, Madison, United States) and incubation for 3 h at 37 °C. Digests were filtered on an EMPORE C8 disk (3M, Fischer Scientific, F67403 Illkirch Cedex, France) and stored at −20 °C. For SDS-PAGE bands of the purified fraction, thin bands were manually excised from the SDS−PAGE gel lanes (corresponding to bands of approximately 64 and 49 kDa) and subjected to in-gel tryptic digestion as previously described.[25]

Tryptic peptides were measured on an LC−MS/MS setup composed of a nanoscale liquid chromatography (Ultimate U3000 RSLC system, Thermo-Fisher Scientific, Bremen, Germany) combined with a hybrid Quadrupole-Orbitrap mass spectrometer (Q Exactive HF, Thermo-Fisher Scientific, Bremen, Germany).

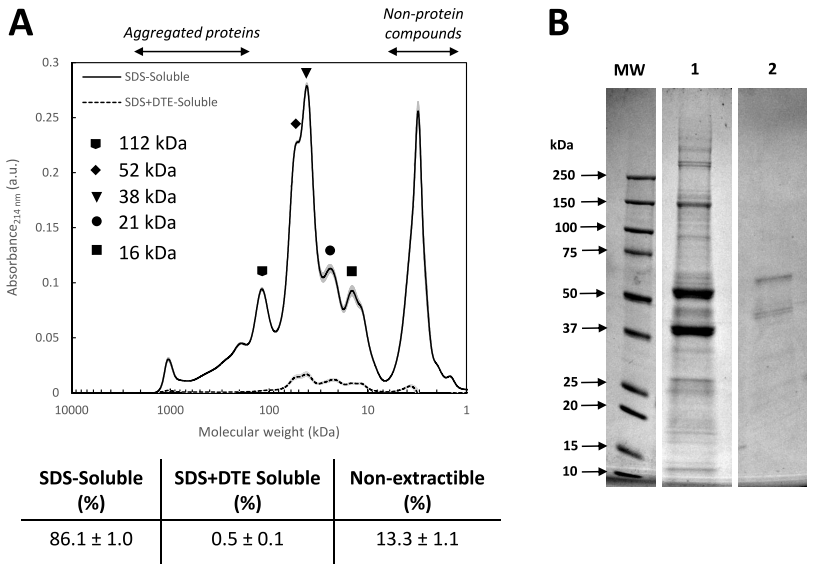
Then, 4 μL of tryptic digests of each sample was injected on the system and eluted on a reversed-phase capillary column (Acclaim PepMap C18, 2 μm, 100 Å, 75 μm i.d. × 25 cm long, Thermo-Fisher Scientific, Bremen, Germany) following a gradient of 50 min and a flow rate of 0.3 μL/min. The data were acquired using a "Top15" data-dependent acquisition method, and an exclusion delay of 5 s was applied. Full MS scans (m/z 400−2000) were acquired at 60K resolution, and the 15 most intense ions (with charges of 2−6) were fragmented in the HCD cell (NCE = 26).

Raw data were converted in the mgf format using msConvert (https://proteowizard.sourceforge.io/tools/msconvert.html) prior to running databank searches with the i2MassChroQ engine (http://pappso.inrae.fr/bioinfo/i2masschroq/).

Protein identification was completed against the Uniprot library restricted to taxon 3803 (Fabaceae; approximately 1200k sequences). The mass tolerance for the precursor and fragment ions was set to 5 ppm. Carbamidomethylation of Cys residue was set as a fixed modification, while oxidation of Met was set as a variable modification. Trypsin enzymatic cleavage was specified for the search, with allowance of up to three missed cleavages. Only proteins identified with at least two specific peptides with an e value below 10^{-3} were conserved for data interpretation.

2.6.4. Polyacrylamide Gel Electrophoresis in Sodium Dodecyl Sulfate (SDS-PAGE). First, samples were solubilized or diluted in the

https://doi.org/10.1021/acs.jafc.5c01681
J. Agric. Food Chem. 2025, 73, 11288−11302



SDS-Soluble (%)	SDS+DTE Soluble (%)	Non-extractible (%)
86.1 ± 1.0	0.5 ± 0.1	13.3 ± 1.1

Figure 2. (A) SE-HPLC in denaturing conditions at 214 nm (upper) and protein extraction yields (lower) panels. (B) SDS-PAGE of pigeon pea flour. 1: SDS-soluble proteins, 2: SDS + DTE soluble proteins. MW: molecular weight marker (Biorad) loaded in 4−20% gradient gels. The gray area on the SE-HPLC profile stands for the standard deviation of the three repeats.

Laemmli buffer (1 mg/mL). For reducing conditions, 2-β-mercaptoethanol (5%, v/v) was added to the sample solutions. Then, all samples were heated at 95 °C for 5 min. Aliquots of 15 μL were added to each well of the polyacrylamide gel (4−12 or 4−20% Bis−Tris Plus Gels, Bolt, Invitrogen). The molecular weight marker (Precision Prestained Protein Standard, MW range of 250−10 kDa, Bio-Rad) was also added to the gel. Electrophoresis was conducted (200 V/100 mA) until complete migration of the samples in the MES SDS running buffer (Novex). After complete migration, the gel was transferred to a container with a fixation solution (2% phosphoric acid, 50% ethanol, v/v) and kept for 1 h under agitation. Then, the gel was rinsed (2% phosphoric acid solution, v/v) for 40 min under agitation, transferred to a sensitization solution (17% ethanol, 2% orthophosphoric acid, and 15% ammonium sulfate, v/v), and stained with Coomassie Brilliant Blue G250 after 20 min and kept overnight. After staining, the gel was rinsed with water and scanned.

2.6.5. UV−Visible Spectroscopy. The UV−visible absorbance spectrum of samples was determined in a UV−visible microplate reader (Varioskan Lux, Thermo Scientific) using a microplate (μDrop, A2-3 model, Thermo Scientific). Each well of the plate was filled with a drop of the supernatant (3 μL). The samples were read in an absorbance range of 210−550 nm with a sampling interval of 2 nm.

2.6.6. Dynamic Light Scattering. The particle size distribution of the samples was determined by dynamic light scattering (DLS, Zetasizer Nano Series, Malvern Instruments). About 0.6 mL of each supernatant was filtered (pore size 0.45 μm, Minisart), placed in a cuvette, and placed in the equipment. Each sample was analyzed at 173° in triplicate at 25 °C, and the refraction index used was 1.45 for protein dispersions. For each reading, 10 runs were performed with the equipment. The CONTIN model was used with the software for particle size calculations.

2.6.7. Intrinsic Fluorescence Measurements. Three-dimensional fluorescence measurements (excitation−emission fluorescence matrices, 3D-EEFM) were performed by using a spectrofluorimeter (FP-8550, Jasco, Japan). Measurements were performed at excitation wavelengths ranging from 240 to 480 nm and emission wavelengths ranging from 280 to 520 nm. A data step of 1 nm and bandwidths of 2.5 nm for excitation and emission were set. For dry state measurements, samples were evaluated using a powder cell and performing the measurements in front-surface mode; in the liquid state, a quartz cuvette was used, and the measurements were performed in right-angle mode. The liquid dispersions were prepared by dispersing the powders in buffer solutions (20 mg/mL of buffer) and centrifuging (10,000g, 10 min, 20 °C). Samples were diluted to reach an absorbance at 280 nm of around 0.075.

2.7. Statistical Analysis. Data were analyzed using the OriginPro 2016 software. Variance analysis (ANOVA) followed by Tukey post hoc test was applied to determine the difference between samples at a 5% significance level when relevant.

3. RESULTS AND DISCUSSION

3.1. Protein Composition of Pigeon Pea Flour. Pigeon pea flour was characterized based on its protein composition using SE-HPLC under denaturing conditions, SDS-PAGE, and proteomic analyses. Following sequential extraction, around 87% of pigeon pea proteins were soluble in SDS, then 0.5% of the proteins in SDS with DTE and finally 13% corresponded to nonextractable proteins.

The SE-HPLC profile in denaturing conditions (Figure 2A) showed a broad range of molecular weights: most likely aggregated proteins with molecular weights higher than 100 kDa; groups of proteins with molecular weights around 112, 52, 38, 21, and 16 kDa; and low-molecular-weight nonprotein compounds (<10 kDa) (Figure S1). Proteins eluting with molecular weights of 52 and 38 kDa correspond to the most abundant proteins in pigeon pea flour. The SDS-PAGE analysis (Figure 2B) of SDS-soluble proteins showed two main subunits: one between 50 and 75 kDa and another between 37 and 50 kDa. These values are quite different from the 72 and 57 kDa subunits identified by Krishna and Bratia[3] for pigeon pea 7S globulin. In other studies,[5,26] subunits of 65, 52, and 47 kDa were also reported to be characteristic of pigeon pea 7S globulin. Also, the protein bands were not sensitive to the reducing agent in our study, being compatible with 7S globulins. The same protein bands were visualized at lower intensity in SDS-insoluble proteins.

Therefore, the characterization of pigeon pea flour showed that approximately 48% of soluble proteins correspond to 7S β-conglycinin, which is expected to be a heterotrimer.[4] The ratio between the two main protein bands identified at 52 and 38 kg/mol was 0.59 (Figure 2A).

Proteomic analysis of pigeon pea flour showed the presence of around 125 protein groups in the Uniprot databank

https://doi.org/10.1021/acs.jafc.5c01681
J. Agric. Food Chem. 2025, 73, 11288−11302

Table 1. Proteins Identified on Pigeon Pea (*Cajanus cajan*) Flour by LC−MS/MS analysis

group	protein name	Uniprot accession	log(*E*value)	MW	spectra	peptides
1a	beta-conglycinin, alpha chain	A0A151U6Q2	−341	64269	661	64
1b	beta-conglycinin, beta chain	A0A151S2A5	−251	48585	388	45
2	sucrose-binding protein	A0A151RL88	−286	54728	205	48
3	glycinin G3	A0A151TUL0	−308	50145	252	43
4	glycinin	A0A151RNW3	−270	55550	266	43
5	alpha-1,4 glucan phosphorylase	A0A151T017	−302	109838	63	33
6	uncharacterized protein	A0A151RC24	−161	24641	128	30
7	low-temperature-induced 65 kDa protein	A0A151RHC5	−193	71752	45	26
8	protein disulfide-isomerase	A0A151SQ13	−204	58231	36	24
9	embryonic protein DC-8	A0A151U337	−139	40489	48	21
10	urease	A0A151SDR3	−173	89113	76	19
11	basic 7S globulin	A0A151TW66	−117	21274	31	19
12	trypsin inhibitor 1A	A0A151SYZ1	−129	21513	57	18
13	late embryogenesis abundant protein	A0A151SSD0	−132	46859	50	18
14	seed biotin-containing protein SBP65	A0A151SEU4	−34	53463	52	17
15	vicilin-like antimicrobial peptides 2-2	A0A151T151	−123	49306	44	17
16	vicilin-like antimicrobial peptides 2-2	A0A151TN63	−138	52448	42	17
17	peroxiredoxin	A0A151S664	−128	24441	42	17
18	glyceraldehyde-3-phosphate dehydrogenase	A0A151TR34	−135	36602	32	16
19	formate dehydrogenase, mitochondrial	A0A151R9F0	−140	42561	41	16
20	peroxidase	A0A151 RKT1	−128	70075	24	15
21	UTP-glucose-1-phosphate uridylyltransferase	A0A151S0L7	−89	52103	19	13
22	glucose and ribitol dehydrogenase	A0A151STP5	−68	31614	32	13
23	poly[ADP-ribose] polymerase	A0A151TNV1	−125	91077	17	12
24	phosphopyruvate hydratase	A0A151TV65	−132	48666	21	12
25	glucose and ribitol dehydrogenase	A0A151U4W9	−68	30240	21	11
26	CobW/HypB/UreG nucleotide-binding domain-containing protein	A0A151S4D2	−133	30493	34	11
27	uncharacterized protein	A0A151SXY6	−86	16106	30	11
28	fructose-bisphosphate aldolase	A0A151R1P9	−88	38255	18	11
29	embryonic protein	A0A151U337	−68	20548	31	11
30	18.2 kDa class I heat shock protein	A0A151SX67	−90	17405	29	10
31	annexin	A0A151TPX1	−82	34445	17	10
32	EMB-1 protein	A0A151S2P2	−62	10913	26	10
33	18.2 kDa class I heat shock protein	A0A151S1P1	−90	17156	24	10
34	nucleoside diphosphate kinase	A0A151RIQ5	−65	16488	11	9
35	lipoxygenase	A0A151SUB6	−55	97316	8	8
36	oil body-associated protein 2B	A0A151RE06	−63	27919	14	8
37	uncharacterized protein	A0A151U0F1	−49	75605	10	8
38	Kunitz-type trypsin inhibitor KTI1	A0A151SZ51	−52	16773	18	8
39	luminal-binding protein 4	A0A151RID0	−68	73319	15	7
40	peptidyl-prolyl cis−trans isomerase	A0A151RAJ7	−62	18185	9	7
41	heat shock cognate protein 80	A0A151S4L4	−58	80833	9	7
42	alcohol dehydrogenase 1	A0A151TSB7	−41	43053	7	7
43	5-methyltetrahydropteroyltriglutamate-homocysteine *S*-methyltransferase	A0A151SX55	−72	84830	14	7
44	cytochrome P450 71A1	A0A151TR99	−65	94943	11	7
45	elongation factor 1-alpha	A0A151RR52	−45	49331	9	7
46	phosphoglycerate kinase	A0A151QRU3	−54	42549	10	7
47	Bowman−Birk type proteinase inhibitor 2	A0A151RI37	−24	12006	10	6
48	uncharacterized protein ECU03-1610	A0A151RZ44	−44	24329	8	6
49	desiccation protectant protein Lea14 isogeny	A0A151UCY8	−40	16784	6	6
50	cysteine proteinase inhibitor	A0A151RC35	−50	26727	11	6
51	triosephosphate isomerase, cytosolic	A0A151SUT3	−50	27196	7	6
52	cell division cycle protein 48 isogeny	A0A151SGI9	−51	84474	7	6
53	elongation factor 2	A0A151S2W2	−52	92537	30	6
54	protein IN2-1 isogeny B	A0A151TUS7	−46	29406	29	6
55	oleosin 18.5 kDa	A0A151T0 × 9	−42	14622	10	6
56	calmodulin	P17928	−36	16851	17	5
57	14-3-3-like protein E	A0A151TDK8	−32	29760	10	5
58	heat shock protein 101 family	A0A151SPH8	−51	100920	9	5
59	defensin-like protein 4	A0A151S2H7	−33	8151	8	5
60	uncharacterized protein	A0A151RHY5	−33	14045	8	5

https://doi.org/10.1021/acs.jafc.5c01681
J. Agric. Food Chem. 2025, 73, 11288−11302

Table 1. continued

group	protein name	Uniprot accession	log(*E*value)	MW	spectra	peptides
61	UPF0098 protein MTH-273	A0A151RGP8	−43	18729	6	5
62	2S albumin	A0A151QQ22	−32	16695	6	5
63	annexin	A0A151SY78	−42	36738	6	5
64	malate dehydrogenase	A0A151T1Y5	−51	35331	6	5
65	aminopeptidase	A0A151R7C2	−54	98738	6	5
66	aquaporin TIP-type alpha	A0A151RKJ8	−39	27113	5	5
67	uncharacterized protein	A0A151QU84	−32	8160	5	5
68	superoxide dismutase	A0A151UDT1	−37	15136	15	5
69	glutaredoxin-dependent peroxiredoxin	A0A151TIT0	−50	17287	13	5
70	reticulon-like protein	A0A151SV55	−37	28677	10	5
71	peroxygenase	A0A151T132	−34	27336	7	5
72	thioredoxin H-type 1	A0A151TC84	−40	12998	6	4
73	calreticulin	A0A151RF02	−36	50444	6	4
74	selenium-binding protein	A0A151SQ99	−24	51415	7	4
75	translationally controlled tumor protein isogeny	A0A151T5 × 4	−41	18963	6	4
76	dehydrin COR47	A0A151TD44	−25	24274	5	4
77	UDP-arabinopyranose mutase	A0A151TRX5	−40	41314	4	4
78	outer envelope pore protein 16-2, chloroplastic	A0A151QR91	−26	20244	4	4
79	UPF0098 protein MTH-273	A0A151RH57	−25	17897	4	4
80	EC protein isogeny 1	A0A151S071	−24	8624	4	4
81	subtilisin inhibitor 1	A0A151RL22	−29	10859	9	4
82	histone H4	A0A151RKC5	−17	11402	7	3
83	eukaryotic translation initiation factor 5A	A0A151SL27	−17	17475	5	3
84	gamma-interferon-inducible lysosomal thiol reductase	A0A151SXY3	−24	28642	4	3
85	uncharacterized protein	A0A151TCQ9	−17	9277	13	3
86	lactoylglutathione lyase	A0A151T6U3	−23	32390	4	3
87	laminin-like protein epi-1	A0A151TM15	−19	98292	5	3
88	seed biotin-containing protein SBP65	A0A151TPG0	−34	34956	5	3
89	18 kDa seed maturation protein	A0A151TUD4	−15	15538	5	3
90	alcohol dehydrogenase 1	A0A151UBN7	−25	37135	4	3
91	pectinesterase	A0A151UCA3	−26	54510	4	3
92	caffeoyl-CoA *O*-methyltransferase At4g26220 family	A0A151RGJ5	−30	26707	4	3
93	SMP domain-containing protein	A0A151RL90	−32	26783	4	3
94	leucine aminopeptidase 1	A0A151RXM3	−26	59642	4	3
95	gibberellin receptor GID1L3	A0A151S0I2	−36	53233	4	3
96	40S ribosomal protein S8	A0A151S1 × 3	−15	24956	3	3
97	seed maturation protein PM36	A0A151ST20	−17	26112	3	3
98	superoxide dismutase	A0A151TQZ9	−13	26656	3	3
99	ATP synthase subunit beta	A0A151U754	−30	59870	3	3
100	uncharacterized protein	A0A151PJF4	−20	16147	3	3
101	ribulose bisphosphate carboxylase large chain	A4GG89	−7	52358	3	2
102	ferritin	A0A0L9UVI2	−6	28416	3	2
103	argininosuccinate synthase	A0A0R0JQZ9	−3	48248	3	2
104	PPM-type phosphatase domain-containing protein	A0A151RJE8	−1	30945	3	2
105	universal stress protein A-like protein	A0A151RYQ8	−1	19950	3	2
106	40S ribosomal protein 5S	A0A151S1 × 3	−1	36765	8	2
107	60S acidic ribosomal protein P2B	A0A151R6 V4	−3	11493	8	2
108	22.7 kDa class IV heat shock protein	A0A151SIV6	−2	22237	5	2
109	aspartic proteinase	A0A151SU18	−9	55422	5	2
110	bifunctional dihydrofolate reductase-thymidylate synthase	A0A151R969	−9	56959	4	2
111	nonspecific lipid-transfer protein	A0A151R3S3	−1	12328	4	2
112	late embryogenesis abundant protein	A0A151R986	−2	13367	4	2
113	glycine-rich protein 2b	A0A151RNE5	−2	18710	4	2
114	calnexin isogeny	A0A151SY44	−8	61873	3	2
115	protein disulfide-isomerase	A0A151SQ85	−7	55628	3	2
116	serine hydroxymethyltransferase	A0A151SQI9	−2	51828	3	2
117	endo-1,3,1,4-beta-D-glucanase	A0A151STR4	−1	25513	2	2
118	mitochondrial outer membrane protein porin of 36 kDa	A0A151SX94	−1	45043	2	2
119	calnexin isogeny	A0A151SY44	−8	62338	2	2
120	superoxide dismutase	A0A151T157	−2	28324	2	2
121	elongation factor 1-gamma	A0A151T893	−2	44107	2	2

https://doi.org/10.1021/acs.jafc.5c01681
J. Agric. Food Chem. 2025, 73, 11288−11302

Table 1. continued

group	protein name	Uniprot accession	log(*E*value)	MW	spectra	peptides
122	uncharacterized protein	A0A151TIA9	−1	32116	2	2
123	uncharacterized protein	A0A151TR40	−8	54150	2	2
124	glutathione *S*-transferase DHAR2	A0A151U811	−2	23172	2	2
125	intracellular protease 1	A0A151U9 × 1	−3	41196	2	2

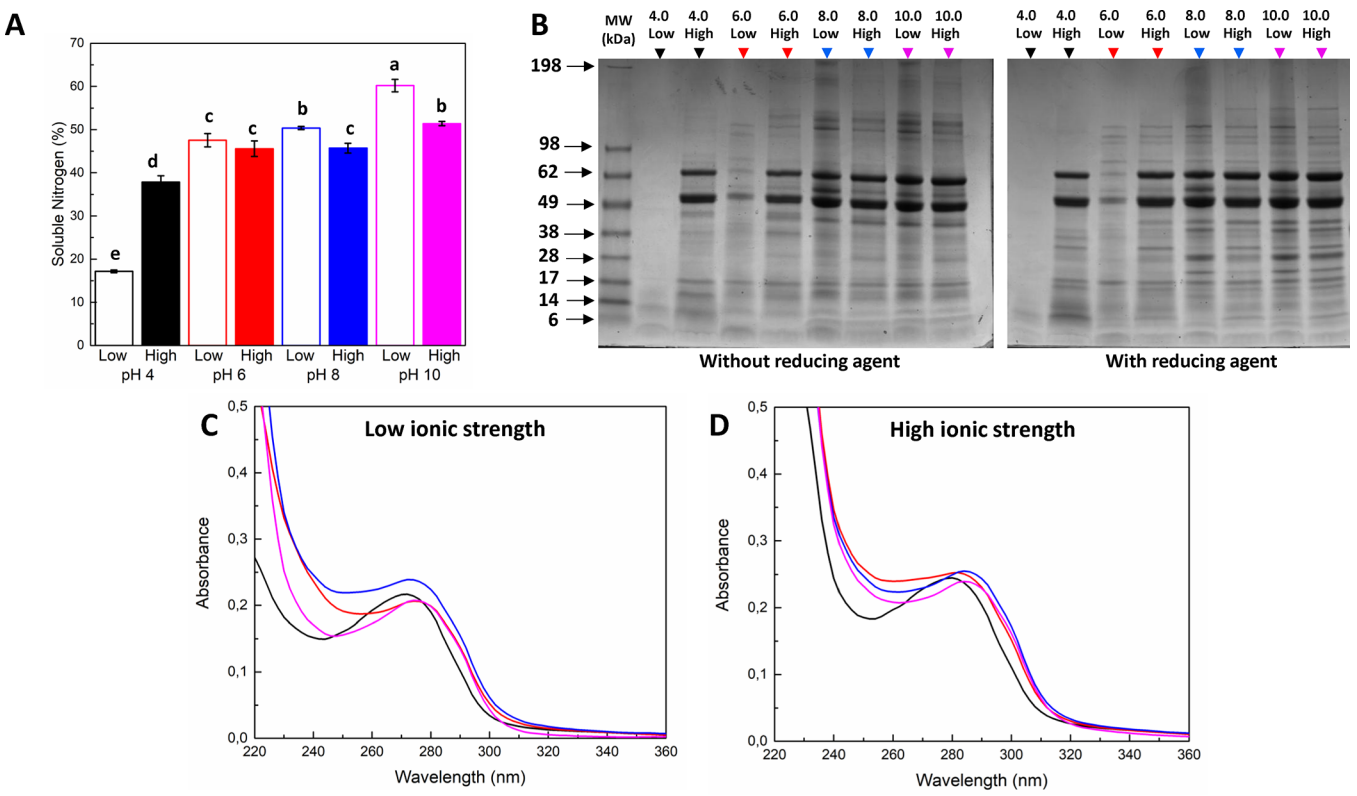


Figure 3. Extractability of pigeon pea proteins with flour suspensions (20 mg/mL) in buffers at different pHs and ionic strength. (A) Nitrogen solubility in the resulting supernatants (mean values with standard deviation bars, $n = 3$). Different letters indicate statistical difference ($p < 0.05$). (B) SDS-PAGE with and without reducing agent. (C, D) UV−visible absorption spectra of supernatants. The lines and markers in black, red, blue, and pink correspond to pHs 4.0, 6.0, 8.0, and 10.0, respectively. Low and high refer to low ionic strength (50 mM) and high ionic strength (500 mM), respectively.

restricted to Fabaceae species (Table 1). Note that only proteins identified with at least two individual peptides were considered valid and are displayed in Table 1. The protein representing the group (i.e., the one identified with the highest number of peptides) is shown.

The main protein identified (on the basis of the number of spectra assigned) was the α-chain of 7S β-conglycinin (64 kDa, group 1a in Table 1) and the β-chain of 7S β-conglycinin (48 kDa, group 1b). Basic 7S globulin (21 kDa, group 11) was also identified, in agreement with a study by Krishnan et al.[26] A sucrose-binding protein was identified in the flour, probably due the fact that pigeon pea 7S globulins are classified as glycoproteins.[4] The identification of glycinin (groups 3 and 4) also indicates the presence of 11S globulins in pigeon pea flour, although in a smaller proportion when compared to the 7S fraction (205 and 256 spectra for groups 3 and 4, respectively, compared to 388 and 661 spectra for groups 1b and 1a, respectively).

Among the oleosins, an oleosin of 18.5 kDa (group 55) was identified, in agreement with predictions using pigeon pea amino acid sequences reported by Locali-Pereira et al.[27] An oil body-associated protein (group 36) of approximately 28 kDa was also identified. Oleosins are proteinaceous components of the lipid storage bodies of plant seeds called oil bodies. Oil bodies are made up of triacylglycerol molecules covered by phospholipids/oleosin annulus. Regarding their function, oleosins have a structural role during embryo development by stabilizing and preventing the coalescence of oil from the oil bodies.[28,29]

Various protease inhibitors, such as cysteine proteinase inhibitor (27 kDa, group 50), trypsin inhibitor 1A (21 kDa, group 12), Kunitz-type trypsin inhibitor KTI1 (16 kDa, group 38), Bowman−Birk type proteinase inhibitor (12 kDa, group 47), and subtilisin inhibitor (11 kDa, group 87), were also identified. These protease inhibitors are commonly found in legume seeds and have a protective action against consumption by predators, like insects, while at the same time existing as seed storage proteins themselves. Its inhibitory action on digestive enzymes can reduce protein digestibility. Bowman−Birk type proteinase inhibitors, for example, interact with the enzymes they inhibit through an exposed surface loop, forming a noncovalent complex that results in the inactivation of the proteinase.[30,31]

https://doi.org/10.1021/acs.jafc.5c01681
J. Agric. Food Chem. 2025, 73, 11288−11302

Other different enzymes, such as lipoxygenase (group 35), peroxidase (group 20), peroxygenase (group 71), and aspartic proteinase (group 109), were also identified. While lipoxygenase, peroxidase, and peroxygenase are involved in oxidation processes of organic compounds, such as esterified fatty acid residues, aspartic proteinase is responsible for the catalysis of peptide substrates.[32,33]

Finally, in the albumin group, only 2S albumin (16 kDa, group 62) was identified. This accession may be related to the molecular weight group of 16 kDa identified in the SE-HPLC analysis (Figure 2A) and to protein bands between 25 and 10 kDa identified in the SDS-PAGE analysis (Figure 2B).

3.2. Extractability of Pigeon Pea Proteins. The extractability of pigeon pea proteins was investigated as a first step in the purification process. To find the right conditions of extraction, we used different pHs (4.0, 6.0, 8.0, and 10.0) and low (50 mM) and high (500 mM) ionic strength. Protein extracts obtained using the different solvents were evaluated for their nitrogen content, SDS-PAGE, UV−visible absorbance measurements, intrinsic fluorescence measurements, and size distribution by DLS.

Protein extraction was affected by the pH and ionic strength of the buffer used. At low ionic strength, protein extraction increased with pH, being approximately 17, 48, 50 and 60% for pHs 4.0, 6.0, 8.0, and 10.0, respectively. At high ionic strength, the amount extracted was approximately 39, 45, 45 and 52% for pHs 4.0, 6.0, 8.0 and 10.0, respectively (Figure 3A).

SDS-PAGE analysis showed two main bands in the protein extracts: one around 64 kDa and the other around 49 kDa (Figure 3B). No bands were observed in the acidic condition (pH 4.0) and at low ionic strength. However, at high ionic strength, the same two bands as identified for the other extracts were visible. A similar result was observed for pH 6.0, with only the band close to 49 kDa being identified yet with the recovery of the two main bands at high ionic strength. For these conditions, the extraction of proteins with subunits of approximately 64 and 49 kDa showed high dependence on ionic strength. We identified that a minimum ionic strength of 100 mM is required to recover both bands under these conditions (data not shown).

The UV−visible absorption spectra also showed intriguing results for the extract at pH 4.0 and low ionic strength (Figure 3C,D). Despite the low nitrogen content of this condition in relation to the other pHs evaluated, the UV−visible spectra were similar for all conditions evaluated. This result suggests that the tryptophan concentration is the same for different nitrogen contents, which may indicate a different extracted protein group. It is also surprising that the UV−visible spectrum is similar to those obtained in the other conditions since no protein bands were identified in the SDS-PAGE analysis for the extract at pH 4.0 and low ionic strength. It may correspond to smaller peptides not revealed by SDS-PAGE.

The emission−excitation fluorescence matrices of the protein extracts measured in right-angle mode showed two distinct regions: one referring to the emission of proteins, mainly to tryptophan/tyrosine emission (λ_{Ex} = 270−290 nm, λ_{Em} = 300−360 nm), and the other referring to the fluorescence of unknown compounds (λ_{Ex} = 300−340 nm, λ_{Em} = 440−480 nm) (Figure S2). For pHs of 4.0 and 6.0 at low ionic strength, a lower fluorescence intensity was observed in relation to the other protein extracts, whereas the absorbance at 280 nm was similar. These results may be due to the different compositions of the proteins extracted under these conditions (potential decrease in tryptophan content) and/or the change in their conformation. Such pH-induced conformational changes have been documented with bovine serum albumin, where acidic conditions lead to structural transitions affecting fluorescence characteristics.[34] For pH 10.0, the fluorescence intensity was the highest among the protein extracts evaluated. It is important to highlight that the absorbance at 280 nm was normalized for all extracts, and since the protein composition seems similar for pH 6.0−10.0, this normalization may correspond to a constant protein content in these extracts. The higher fluorescence intensity may be due to the effect of the pH on the quantum yield of tyrosine/tryptophan. While no changes in the quantum yield of tyrosine are expected in this pH range, tryptophan may show an increase in its quantum yield for pH 8.0 and 10.0.[35,36] Therefore, the increase in fluorescence observed for 10.0 nm may be related to changes in tryptophan quantum yield.

In relation to the fluorescence region of phenolic compounds, a lower fluorescence intensity was observed for pH values of 8.0 and 10.0 regardless of the ionic strength. These results may indicate a lower extraction of these compounds at these pHs, but studies were conducted by extracting the proteins at pH 10.0 and acidifying the extract to pH 8.0, 6.0, and 4.0 with intrinsic fluorescence measurements at each step, and it was found that this is a pH-dependent response to the fluorescence of these compounds (data not shown).

The extracts were also investigated for size distribution by DLS (Figure S3). The size distribution of the samples showed two populations: a small population around 15−20 nm and another larger population around 100−130 nm. Ruiz-Henestrosa et al.[37] reported a size distribution of 10.1−11.7 nm for soybean 7S globulin. The first population may therefore be attributed to the 7S globulin trimer. The larger population (≈100 nm) suggests the presence of aggregates in the protein extracts. For extracts at low ionic strength, the contribution of the population around 15−20 nm to the signal was lower than that in extracts at high ionic strength. At high ionic strength, the aggregation state of pigeon pea proteins was reduced.

Altogether, the results showed that all extracts, except extracts at pH 4.0 and 6.0, presented the same protein profile. Therefore, pHs above 6.0 were shown to be suitable for the extraction of 7S globulin. The low effect of ionic strength on the extractability is suitable for subsequent separation using ion exchange chromatography. Also, all extracts showed the presence of aggregates and minor compounds regardless of pH or added salt. However, the results suggest that increasing the salt after ion exchange chromatography can reduce protein aggregation.

3.3. Laboratory-Scale Purification. *3.3.1. Desalting vs UF.* The elimination of small molecules contained in protein extracts was investigated by using two approaches: membrane ultrafiltration and injection on a desalting column. Samples were evaluated by emission−excitation fluorescence matrices to monitor the presence of the compounds of interest (tryptophan/tyrosine fluorescence emission) and the elimination of unknown compounds (small molecules).

The use of membrane ultrafiltration proved to be an efficient technique for separating small molecules, which were concentrated in the extract filtrate, whereas proteins were concentrated in the retentate (Figure S4). Only a low intensity of fluorescence could be seen in the region of unknown

https://doi.org/10.1021/acs.jafc.5c01681
J. Agric. Food Chem. 2025, 73, 11288−11302

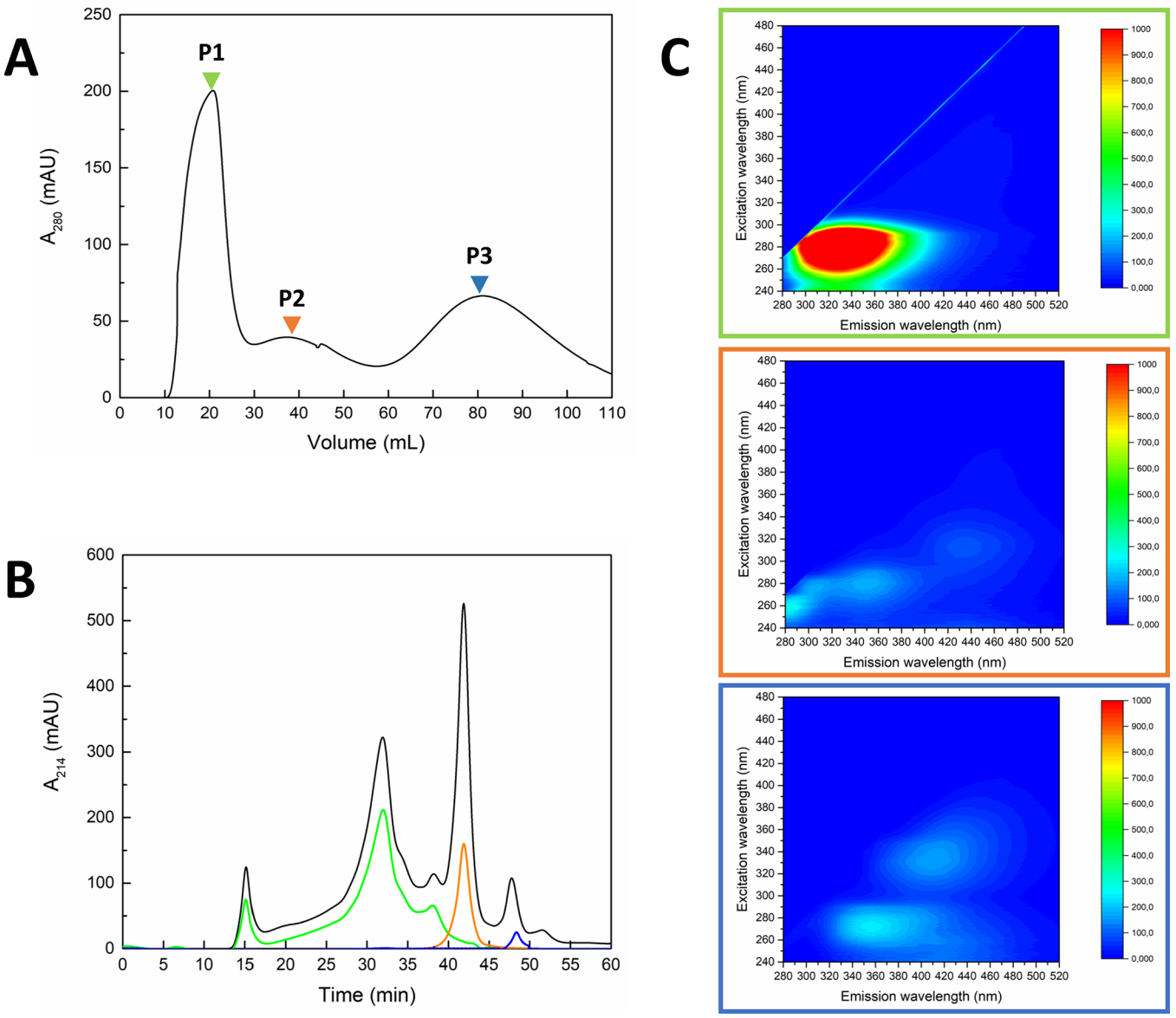


Figure 4. Injection of pigeon pea protein extract into a desalting column to remove small molecules. (A) Elution profile on the HiTrap desalting column. (B) Size exclusion chromatogram of the pigeon pea protein extract and the fractions separated in the desalting column. (C) 3D emission−excitation fluorescence matrices of the fractions separated in the desalting column. Black, green, orange, and blue lines refer to the initial protein extract and peaks P1, P2, and P3, respectively.

compounds in the retentate, suggesting that some low-molecular-weight compounds were still present in the extract.

In parallel, the desalting of the protein extract in a HiPrep desalting column showed three chromatographic peaks (Figure 4A). The size exclusion chromatograms (SECs) of the protein extract and of the fractions separated by the desalting column show the predominance of the target protein in the first peak and smaller molecules in peaks 2 and 3 (Figure 4B). Also, the fluorescence signal in the region attributed to proteins was observed only in the first peak, while the other fractions showed low fluorescence in the protein region (Figure 4C).

The ultrafiltered and desalted extracts were also compared by analytical size exclusion chromatography (Figure 5). This showed that desalting contains less aggregates (eluted around 15 min) and less small molecules (eluted between 40 and 55 min), evidenced by the peak intensity ratio (30/42 min), which was 0.6 for the initial protein extract and 19.6 for the desalted protein extract. On the other hand, the peak intensity ratio (30/15 min) was 2.6 for the initial protein extract and 2.8 for the desalted protein extract. This difference between both techniques can be explained by a higher concentration on the membrane caused by ultrafiltration, which can lead to an irreversible aggregation state. Despite this, membrane ultrafiltration was shown to be a fast and efficient technique for concentrating the minor compounds present in the protein extract, which can be interesting for studies that aim to separate them for identification and quantification. For protein purification purposes, desalting was shown to be the most suitable technique and was used in the protocol described in this work for the purification of pigeon pea 7S globulin.

3.3.2. Separation by Anion Exchange and Size Exclusion Chromatography. A prescreening of different chromatography columns showed that anion exchange columns performed better in retaining pigeon pea proteins compared to cation

https://doi.org/10.1021/acs.jafc.5c01681
J. Agric. Food Chem. 2025, 73, 11288−11302

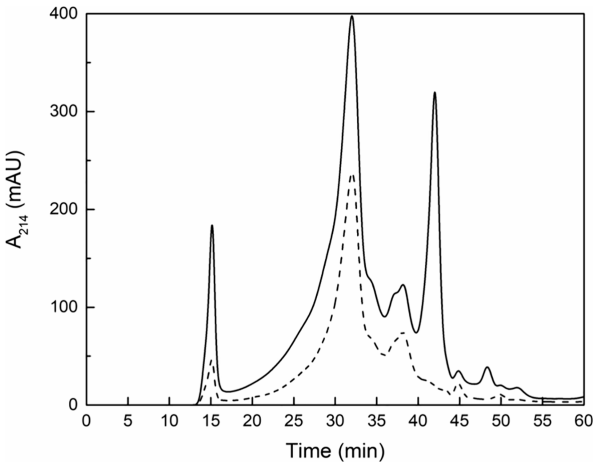


Figure 5. Size exclusion chromatography of pigeon pea protein extract after membrane ultrafiltration (solid line) and injection in a desalting column (dashed line).

exchange columns. Pigeon pea protein extracts at pH 8.0 and injected into the Q column showed better results in protein separation (Figure S5), so these conditions were used to compare both protein extracts.

To further optimize the separation using the desalted extract, two procedures were investigated for anion exchange chromatography: (i) increasing the duration of the NaCl elution gradient from 0 to 0.75 M NaCl and (ii) increasing the NaCl concentration at 0.1 M in the extract to avoid fixation of weakly charged proteins onto the column. In the first procedure, the AEC separation showed three peaks (Figure 6A). SDS-PAGE analysis of the fractions showed the presence of weak protein bands characteristic of 7S globulin in the fraction that was not retained in the column. The first separated peak showed a mixture of different proteins, the second peak showed the characteristic bands of the 7S globulin, and the third peak showed no protein bands. The second procedure used showed a better separation of the target protein (Figure 6B), eliminating the first peak and enabling better separation of the 7S globulin peak. SDS-PAGE analysis of the non-retained fraction from this procedure also showed that most nonretained proteins were depleted with the second procedure. The 7S globulin fractions separated by AEC were collected, pooled, and further purified by SEC (Figure 6C), which showed a single major peak of the 7S fraction and small molecule peaks.

3.4. Large-Scale Purification. The large-scale purification process of pigeon pea 7S globulins was carried out using the same extraction procedures as the optimized process on an analytical scale. The extract desalting step on the GH25 column presented results similar to those obtained on the analytical scale. Three peaks were observed, the first peak referring to the protein-rich fraction and the two subsequent peaks referring to minor compounds such as phenolic compounds and salts (Figure 7A).

The desalted extract was then injected onto the Q-Sepharose column to be fractionated by AEC with an elution step (Figure 7B). Three peaks were observed, with peak 1 referring to the fraction not retained on the column, whereas peaks 2 and 3 corresponded to proteins bound on the column eluting at 0.35 and 1 M NaCl, respectively. SDS-PAGE analysis showed that the fraction that was not retained on the column is composed of protein bands different from those of 7S globulin, whereas fraction 2 presented two main bands close to the region of pigeon pea 7S globulin subunits (around 47 and 64 kDa), and peak 3 presented the same bands as peak 2 but less intense. Proteins not retained in the Q-Sepharose column were also analyzed for their composition. The nonretained fraction was diluted and reinjected into the Q-Sepharose and GLC 2000 column for further separation (Figure S6). The main protein band was identified as trypsin inhibitor 1A (Uniprot A0A151SYZ1_CAJCA) by proteomic analysis (Table S1). This shows that part of the trypsin inhibitors is potentially eliminated at this stage of the process.

After fractionation by AEC, further fractionation of the fraction eluted at 0.35 M NaCl was performed by SEC. The strategy of equilibrating the GLC 2000 column with ammonium carbonate buffer at pH 8.0 allows for circumventing the dialysis step. In this case, salt elimination was conducted at the same time as protein size separation during SEC.

The fractionation by SEC on the GLC 2000 column resulted in the elution of the 7S globulins between 1000 and 1500 mL of volume and small peaks eluted between 1500 and 2500 mL of volume (Figure 7C), referring to small compounds. SDS-PAGE analysis showed protein bands around 49 and 64 kDa, characteristic of pigeon pea 7S globulin in the native state.

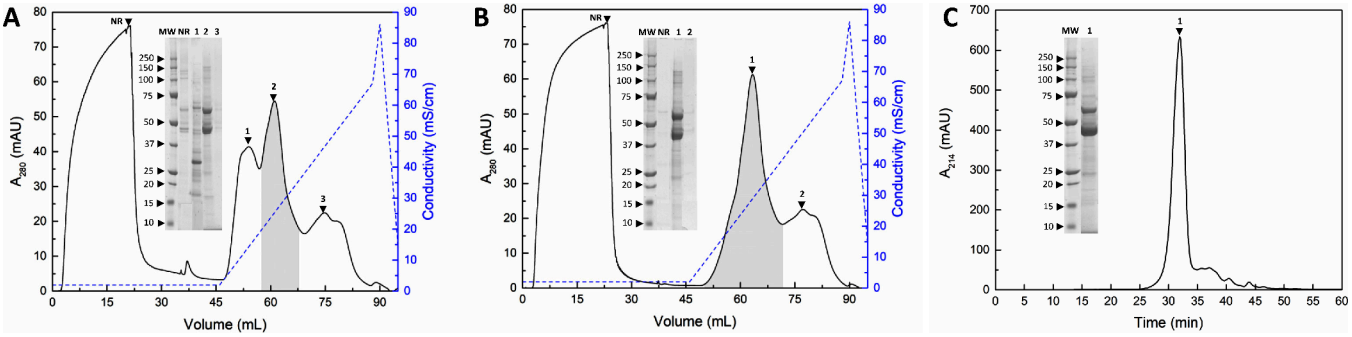


Figure 6. Analytical-scale purification of pigeon pea 7S globulin. (A) Ion exchange chromatography with increasing NaCl elution gradient (step 1 = 0 to 0.75 M in 40 mL, step 2 = 0.5 to 1 M in 5 mL) and (B) with the combination of increasing NaCl elution gradient and 1 M NaCl solution addition in the protein extract before injection. (C) Size exclusion chromatography of the pigeon pea 7S globulin fraction. NR: not-retained fraction; 1, 2, and 3: fraction peaks. The inset shows the SDS-PAGE gels of the separated peaks. Gray areas indicate pooled fractions for subsequent SEC analysis.

https://doi.org/10.1021/acs.jafc.5c01681
J. Agric. Food Chem. 2025, 73, 11288−11302

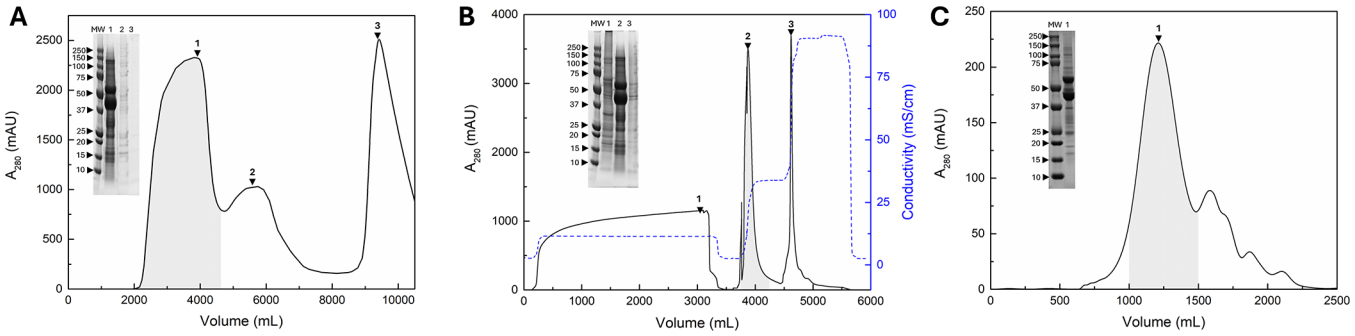


Figure 7. Large-scale purification of pigeon pea 7S globulin. (A) Elution profile in the desalting column. (B) Anion exchange chromatography of the desalted protein extract. (C) Size exclusion chromatogram of pigeon pea 7S globulin. The inset shows the SDS-PAGE analysis of numbered fractions. The gray filled peak corresponds to the 7S globulin fraction.

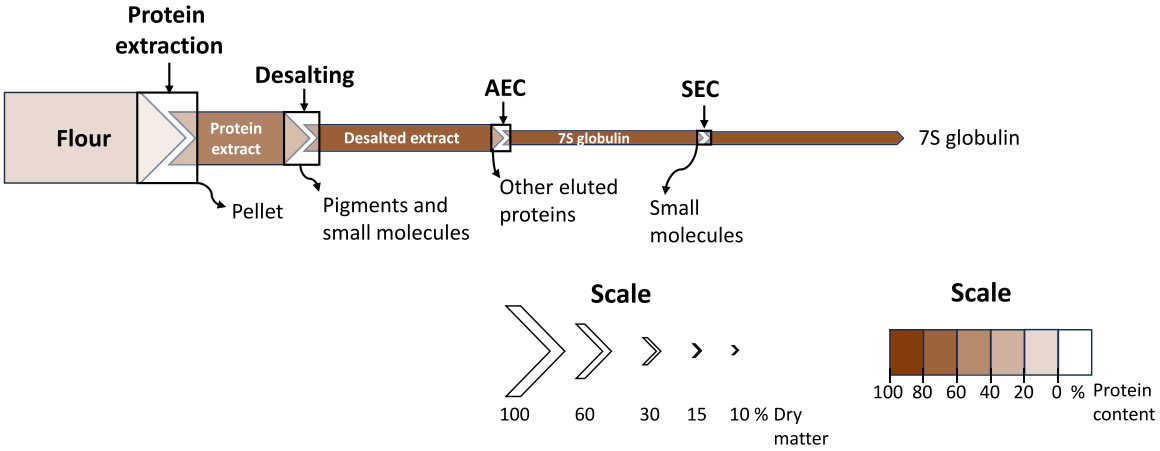


Figure 8. Distribution of dry matter and protein content during large-scale purification of pigeon pea 7S globulin. AEC: anion exchange chromatography. SEC: size exclusion chromatography.

3.5. Analytical Characterization of the Pigeon Pea 7S Globulin.

3.5.1. Yield and Purity. The process yield is not usually reported by authors in pulse protein purification studies, but it is expected to be around 10 mg of 7S globulin/1 g of flour (1%).[7] Figure 8 shows the distribution of dry matter and protein content throughout the purification process of pigeon pea 7S globulin.

Around 200 g of flour was used in both processes, with a total soluble protein content of approximately 33.6 g. The protein content of the purified 7S globulin fraction was about 89% (w/w) dry matter (DM). The dry matter obtained was 4.4 g. Considering that all purified proteins are 7S globulins and that the initial content of 7S globulins in the flour was 16.3 g, based on SE-HPLC analysis, the recovery of 7S globulins was 23.3%, being above the 1−19% range reported by Gravel and Doyen[7] for the recovery of 11S and 7S globulin fractions. Therefore, our protocol was shown to be efficient for recovering a reasonable amount of the 7S globulin.

3.5.2. Protein Composition. The protein composition of the purified pigeon pea 7S globulin fraction was characterized by SDS-PAGE and proteomic analysis. SDS-PAGE analysis revealed two main bands in the purified fraction: one between 75 and 50 kDa and two other bands between 50 and 37 kDa (Figure 9). Such bands correspond to the subunits of 7S globulin from pigeon pea, as described by Krishna and Bhatia.[3] As observed by the authors, the banding pattern did not show changes in the presence of the reducing agent 2-β-mercaptoethanol (subunits not linked by disulfide bonds).

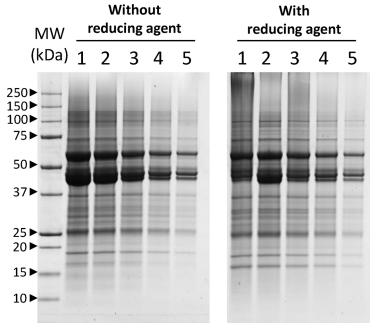


Figure 9. SDS-PAGE of pigeon pea 7S globulin in the presence or absence of the reducing agent (2-β-mercaptoethanol). Concentrations of 2.0, 1.5, 1.0, 0.5, and 0.25 mg/mL of proteins are indicated by the numbers 1, 2, 3, 4, and 5, respectively.

For the well with the highest concentration of proteins, it is possible to visualize other bands (minor contaminants), whereas the well with the lowest concentration of proteins shows the purity of the two main bands referring to pigeon pea 7S globulin. The protein bands identified in SDS-PAGE for the purified 7S globulin fraction (Table 2), as well as the crude protein extract (powder) (Table S2), were evaluated by proteomic analysis. The two main bands observed in 7S globulin of 64 and 49 kDa were identified as the α- and β-chains of β-conglycinin, respectively. Glycinin was also identified in the purified fraction, suggesting the presence of 11S globulin residues.

https://doi.org/10.1021/acs.jafc.5c01681
J. Agric. Food Chem. 2025, 73, 11288−11302

Table 2. Proteins Identified on Pigeon Pea 7S Globulin Purified Fraction by LC−MS/MS Analysis

group	protein name	Uniprot accession	log(*E*value)	MW	spectra	peptides
		band 64 kDa				
1	beta-conglycinin, alpha chain	A0A151U6Q2	−239	64269	659	38
2	glucose and ribitol dehydrogenase	A0A151U4W9	−27	31614	8	4
3	glycinin	A0A151RNW3	−21	55550	4	3
4	heat shock 70 kDa protein	P26413	−11	70835	2	2
group	**protein name**	**Uniprot accession**	**log(*E*value)**	**MW**	**spectra**	**peptides**
		band 49 kDa				
1	beta-conglycinin, beta chain	A0A151S2A5	−358	48585	883	60
2	guanosine nucleotide diphosphate dissociation inhibitor	A0A151UC16	−25	49691	6	4
3	ATP-dependent RNA helicase	A0A072URR3	−14	48555	2	2
4	glycinin	A0A151RNW3	−24	55550	2	2

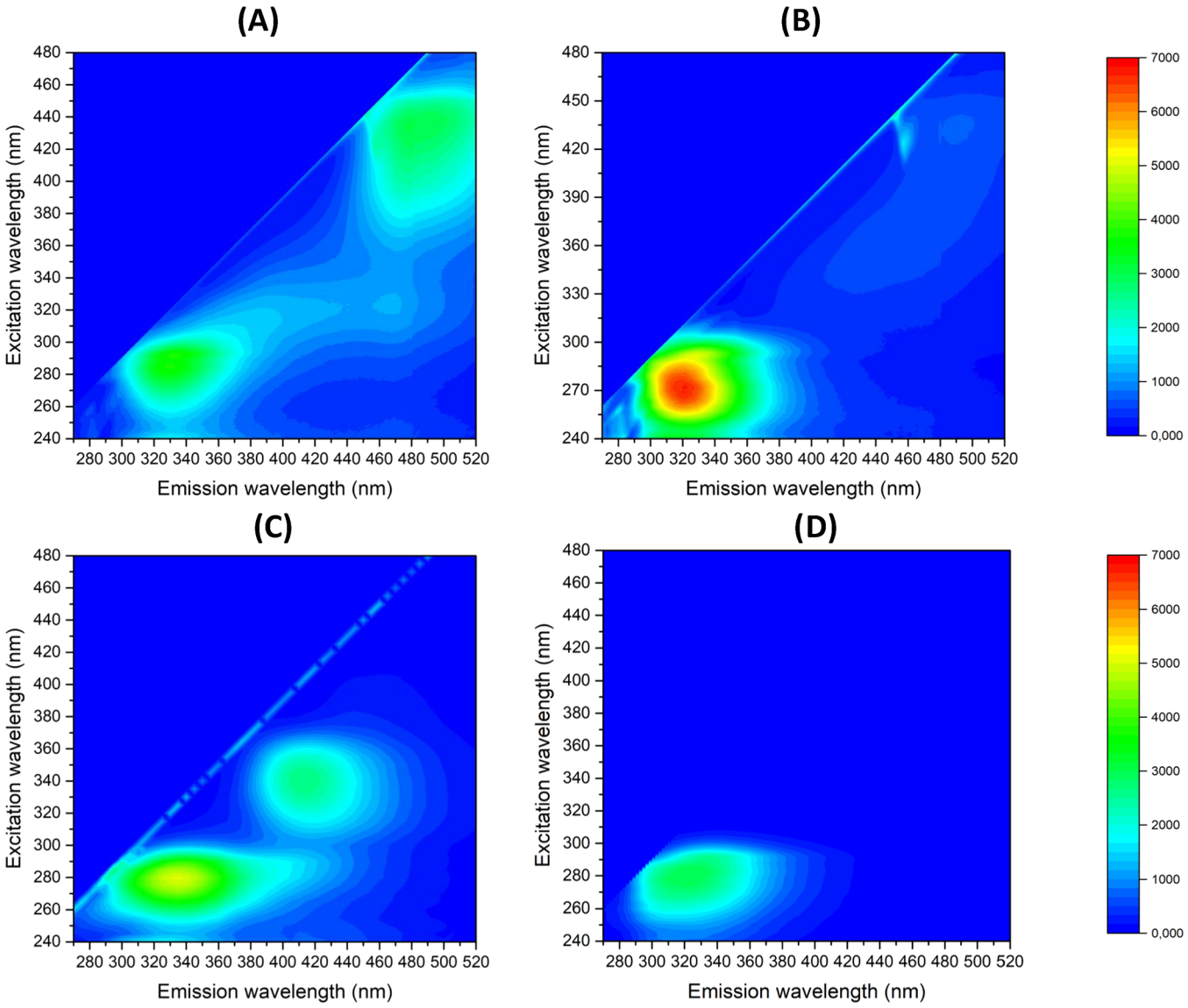


Figure 10. Emission−excitation fluorescence matrices of pigeon pea flour (A, C) and purified 7S globulin (B, D). (A, B) Measurements performed in front-surface mode (dry state) and (C, D) measurements performed in right-angle mode (liquid state).

3.5.3. Minor Components. The minor components present in the purified 7S globulin fraction were mainly small sugars. Around 3 and 2.3% of polysaccharides were identified in the purified 7S globulin (Table S3). Mannose and uronic acid were the main neutral sugars in the purified fraction with values of, respectively, 1.4 and 0.9 g/100 g. The high mannose content is surprising and may be linked to the glycosylation of pigeon pea 7S globulin.

As shown by proteomic analysis, trypsin inhibitors were present in the crude protein extract (Table S2), and the trypsin

https://doi.org/10.1021/acs.jafc.5c01681
J. Agric. Food Chem. 2025, 73, 11288−11302

inhibitor activity (TIA) was investigated in the purified fraction (Table S4). The trypsin inhibitor activity (TIA) was 3.6 μg TId/mg of sample for 7S globulin. The pigeon pea flour and protein concentrate presented TIA values of 8.3 and 11.5 μg TId/mg in a previous work by our research group.[8] Therefore, the purified fraction had a significant reduction in TIA in relation to these values, probably due to the elimination of trypsin inhibitors during the chromatographic separation steps.

The occurrence of minor compounds was also touched upon by monitoring the endogenous fluorophores present in the purified fraction. Fluorescence emission−excitation matrices were measured in front-surface and right-angle modes for the pigeon pea flour and purified 7S globulin fraction (Figure 10). While the flour presents two distinct regions of fluorescence emission (in both measurement methods), only the region referring to the fluorescence emission of proteins can be seen for the purified 7S globulin. In the latter samples, the emission intensity is also considerably higher. This is in line with the fair purity of the 7S globulin fraction, in addition to suggesting the complete elimination and/or degradation of the other endogenous fluorophores observed in the flour. The elimination of these compounds was demonstrated during the desalting step of the protein extract, and it probably corresponds to phenolic compounds.[38−40] A difference in the location of the fluorescence emission region of the unknown compounds was observed for the two measurement methods investigated. While in front-surface mode the fluorescence emission region of the unknown compounds was at $\lambda_{Ex} \approx 420$ nm and $\lambda_{Em} \approx 500$ nm, in right-angle mode the region was shifted to $\lambda_{Ex} \approx 340$ nm and $\lambda_{Em} \approx 410$ nm. The shape observed for this region also changed, being more circular for the measurement taken in the right-angle mode.

In this work, a large-scale purification method for pigeon pea 7S globulins was developed. The protocol used was first developed and optimized on a laboratory scale, showing reproducibility when transferred to a large scale. Desalting was shown to be a better technique than membrane ultrafiltration to eliminate aggregates and small molecules when applied before separation by anion exchange and size exclusion chromatography. The 7S globulin purified by large-scale chromatography was mainly composed of α- and β-chains of β-conglycinin, with subunits of molecular weight around 64 and 49 kDa. Carbohydrate residues, mainly mannose, were identified in the purified fraction, suggesting the presence of glycosylated proteins. Fluorescence measurements of the purified fraction showed the efficient removal of nonproteinaceous fluorophore compounds during the purification process. The developed processes are suitable for purifying pigeon pea 7S globulin in sufficient amounts for protein characterization and functional properties studies.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jafc.5c01681.

Protein composition of pigeon pea flour by SE-HPLC, pigeon pea protein extractability, fractionation optimization and large-scale purification of pigeon pea trypsin inhibitor, and analytical characterization (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Véronique Solé-Jamault − *INRAE, UR BIA, Nantes F-44316, France;* orcid.org/0000-0002-6221-7630; Email: veronique.sole@inrae.fr

Authors

Adilson Roberto Locali-Pereira − *Institute of Biosciences, Humanities and Exact Sciences, Department of Food Engineering and Technology, Unesp - São Paulo State University, São José do Rio Preto, SP 15054-000, Brazil;* orcid.org/0000-0003-0242-482X

Joëlle Davy − *INRAE, UR BIA, Nantes F-44316, France*

Mehdi Cherkaoui − *INRAE, UR BIA, Nantes F-44316, France; INRAE, PROBE/CALIS Research Infrastructures, BIBS Facility, Nantes F-44300, France*

Hélène Rogniaux − *INRAE, UR BIA, Nantes F-44316, France; INRAE, PROBE/CALIS Research Infrastructures, BIBS Facility, Nantes F-44300, France;* orcid.org/0000-0001-6083-2034

Hamza Mameri − *UMR IATE, Univ Montpellier, INRAE, Institut-Agro Montpellier, Montpellier F-34060, France*

Sophie Le Gall − *INRAE, UR BIA, Nantes F-44316, France; INRAE, PROBE/CALIS Research Infrastructures, BIBS Facility, Nantes F-44300, France;* orcid.org/0000-0001-7814-8185

Valérie Beaumal − *INRAE, UR BIA, Nantes F-44316, France*

Hanitra Rabesona − *INRAE, UR BIA, Nantes F-44316, France*

Sophie Laurent − *INRAE, UR BIA, Nantes F-44316, France*

Vânia Regina Nicoletti − *Institute of Biosciences, Humanities and Exact Sciences, Department of Food Engineering and Technology, Unesp - São Paulo State University, São José do Rio Preto, SP 15054-000, Brazil;* orcid.org/0000-0002-2553-4629

Claire Berton-Carabin − *INRAE, UR BIA, Nantes F-44316, France; Laboratory of Food Process Engineering, Wageningen University, Wageningen 6708 WG, The Netherlands*

Adeline Boire − *INRAE, UR BIA, Nantes F-44316, France*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jafc.5c01681

Funding

The authors thank the Brazilian Federal Agency for Support and Evaluation of Graduate Education (CAPES) and the Ministry of Europe and Foreign Affairs of France (EIFFEL Excellence Program) for the financial support of A.R.L.P.'s PhD grant.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

Carbohydrate analyses were performed on the BIBS instrumental platform (DOI: 10.15454/1.5572358121569739E12; http://www.bibs.inrae.fr/bibs_eng/, UR1268 BIA, IBiSA, Phenome-Emphasis-FR ANR-11-INBS- 0012, PROBE infrastructure, Biogenouest). The authors are also grateful to Joëlle Bonicel for her technical assistance during SE-HPLC analysis.

REFERENCES

(1) Fuller, D. Q.; Murphy, C.; Kingwell-Banham, E.; Castillo, C. C.; Naik, S. Cajanus cajan (L.) Millsp. origins and domestication: the South and Southeast Asian archaeobotanical evidence. *Genet. Resour. Crop Evol.* **2019**, *66* (6), 1175−1188.

(2) Krishna, T. G.; Bhatia, C. R. γ-Protein, a Sulphur Amino Acid Rich Protein from Pigeon Pea (Cajanus Cajan (L.) Millsp.). *J. Biosci* **1986**, *10* (1), 57−65.

(3) Krishna, T. G.; Bhatia, C. R. Vicilin from Cajanus Cajan Seeds. *Phytochemistry* **1985**, *24* (10), 2201−2203.

(4) Gopala Krishna, T.; Mitra, R. K.; Bhatia, C. R. Seed Globulins of Cajanus Cajan. *Qual. Plant.* **1977**, *27* (3−4), 313−325.

(5) Fernández Sosa, E. I.; Chaves, M. G.; Quiroga, A. V.; Avanza, M. V. Comparative Study of Structural and Physicochemical Properties of Pigeon Pea (Cajanus Cajan L.) Protein Isolates and Its Major Protein Fractions. *Plant Foods for Human Nutrition* **2021**, *76* (1), 37−45.

(6) Bravo, R. K. D.; Angelia, M. R. N.; Uy, L. Y. C.; Garcia, R. N.; Torio, M. A. O.; et al. Isolation, purification and characterization of the antibacterial, antihypertensive and antioxidative properties of the bioactive peptides in the purified and proteolyzed major storage protein of pigeon pea (Cajanus cajan) seeds. *Food Chemistry: Molecular Sciences* **2022**, *4*, No. 100062.

(7) Gravel, A.; Doyen, A. Pulse Globulins 11S and 7S: Origins, Purification Methods, and Techno-Functional Properties. *J. Agric. Food Chem.* **2023**, *71*, 2704−2717.

(8) Locali-Pereira, A. R.; Caruso, Í. P.; Rabesona, H.; Laurent, S.; Meynier, A.; Kermarrec, A.; Birault, L.; Geairon, A.; Le Gall, S.; Thoulouze, L.; Solé-Jamault, V.; Berton-Carabin, C.; Boire, A.; Nicoletti, V. R. Pre-Treatment Effects on the Composition and Functionalities of Pigeon Pea Seed Ingredients. *Food Hydrocoll* **2024**, *152*, No. 109923.

(9) Pickardt, C.; Eisner, P.; Kammerer, D. R.; Carle, R. Pilot Plant Preparation of Light-Coloured Protein Isolates from de-Oiled Sunflower (Helianthus Annuus L.) Press Cake by Mild-Acidic Protein Extraction and Polyphenol Adsorption. *Food Hydrocoll* **2015**, *44*, 208−219.

(10) Ducrocq, M.; Morel, M.-H.; Anton, M.; Micard, V.; Guyot, S.; Beaumal, V.; Solé-Jamault, V.; Boire, A. Biochemical and Physical−Chemical Characterisation of Leaf Proteins Extracted from Cichorium Endivia Leaves. *Food Chem.* **2022**, *381*, No. 132254.

(11) Locali-Pereira, A. R.; Caruso, Í. P.; Rabesona, H.; Laurent, S.; Meynier, A.; Kermarrec, A.; Birault, L.; Geairon, A.; Le Gall, S.; Thoulouze, L.; Solé-Jamault, V.; Berton-Carabin, C.; Boire, A.; Nicoletti, V. R. Pre-Treatment Effects on the Composition and Functionalities of Pigeon Pea Seed Ingredients. *Food Hydrocoll* **2024**, *152*, No. 109923.

(12) Verfaillie, D.; Janssen, F.; Van Royen, G.; Wouters, A. G. B. A Systematic Study of the Impact of the Isoelectric Precipitation Process on the Physical Properties and Protein Composition of Soy Protein Isolates. *Food Research International* **2023**, *163*, No. 112177.

(13) Tang, C. H.; 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* (10), 6395−6402.

(14) Gueguen, J.; Vu, A. T.; Schaeffer, F. Large-scale Purification and Characterisation of Pea Globulins. *J. Sci. Food Agric* **1984**, *35* (9), 1024−1033.

(15) Dapremont, O. Use of Industrial Scale Chromatography in Pharmaceutical Manufacturing. In *Comprehensive Organic Synthesis II*; Elsevier: 2014; pp 181−206. .

(16) Larré, C.; Gueguen, J. Large Scale Purification of Pea Globulins. Comparasion between Six Anion Exchangers in Medium-Pressure Liquid Chromatography. *J. Chromatogr.* **1986**, *361*, 169.

(17) Crévieu, I.; Bérot, S.; Guéguen, J. Large Scale Procedure for Fractionation of Albumins and Globulins from Pea Seeds. *Nahrung - Food* **1996**, *40* (5), 237−244.

(18) Bérot, S.; Compoint, J. P.; Larré, C.; Malabat, C.; Guéguen, J. Large Scale Purification of Rapeseed Proteins (Brassica Napus L.). *J. Chromatogr. B: Anal. Technol. Biomed. Life Sci.* **2005**, *818*, 35−42.

(19) Suchkov, V. V.; Popello, I. A.; Grinberg, V. Ya.; Tolstoguzov, V. B. Isolation and Purification of 7S and 11S Globulins from Broad Beans and Peas. *J. Agric. Food Chem.* **1990**, *38* (1), 92−95.

(20) Morel, M.-H.; Dehlon, P.; Autran, J. C.; Leygue, J. P.; Bar-L'Helgouac'h, C. Effects of Temperature, Sonication Time, and Power Settings on Size Distribution and Extractability of Total Wheat Flour Proteins as Determined by Size-Exclusion High-Performance Liquid Chromatography. *Cereal Chem.* **2000**, *77* (5), 685−691.

(21) Alamanou, S.; Doxastakis, G. Effect of Wet Extraction Methods on the Emulsifying and Foaming Properties of Lupin Seed Protein Isolates (Lupinus Albus Ssp. Graecus). *Food Hydrocoll* **1997**, *11* (4), 409−413.

(22) Le Gall, S.; Even, S.; Lahaye, M. Fast Estimation of Dietary Fiber Content in Apple. *J. Agric. Food Chem.* **2016**, *64* (6), 1401−1405.

(23) AOCS *Official Method Ba 12a-2020. Assaying Trypsin Inhibitor Activity*. 2020.

(24) Liu, K. Trypsin Inhibitor Assay: Expressing, Calculating, and Standardizing Inhibitor Activity in Absolute Amounts of Trypsin Inhibited or Trypsin Inhibitors. *J. Am. Oil Chem. Soc.* **2021**, *98* (4), 355−373.

(25) Devouge, V.; Rogniaux, H.; Nési, N.; Tessier, D.; Guéguen, J.; Larré, C. Differential Proteomic Analysis of Four Near-Isogenic *Brassica Napus* Varieties Bred for Their Erucic Acid and Glucosinolate Contents. *J. Proteome Res.* **2007**, *6* (4), 1342−1353.

(26) Krishnan, H. B.; Natarajan, S. S.; Oehrle, N. W.; Garrett, W. M.; Darwish, O. Proteomic Analysis of Pigeonpea (Cajanus Cajan) Seeds Reveals the Accumulation of Numerous Stress-Related Proteins. *J. Agric. Food Chem.* **2017**, *65* (23), 4572−4581.

(27) Locali-Pereira, A. R.; Boire, A.; Berton-Carabin, C.; Taboga, S. R.; Solé-Jamault, V.; Nicoletti, V. R. Pigeon Pea, An Emerging Source of Plant-Based Proteins. *ACS Food Sci. Technol.* **2023**, *3*, 1777.

(28) Murphy, D. J.; Keen, J. N.; O'Sullivan, J. N.; Au, D. M. Y.; Edwards, E.-W.; Jackson, P. J.; Cummins, I.; Gibbons, T.; Shaw, C. H.; Ryan, A. J. A Class of Amphipathic Proteins Associated with Lipid Storage Bodies in Plants. Possible Similarities with Animal Serum Apolipoproteins. *Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression* **1991**, *1088* (1), 86−94.

(29) Tzen, J. T. C.; Lie, G. C.; Huang, A. H. C. Characterization of the Charged Components and Their Topology on the Surface of Plant Seed Oil Bodies. *J. Biol. Chem.* **1992**, *267* (22), 15626−15634.

(30) Aoki-Shioi, N.; Terada, S.; Hellinger, R.; Furuta, Y.; Gruber, C. W. Isolation and Functional Diversity of Bowman−Birk Type Serine Proteinase Inhibitors from *Hyacinthus Orientalis*. *Biochem. J.* **2021**, *478* (6), 1287−1301.

(31) Rawlings, N. D.; Tolle, D. P.; Barrett, A. J. Evolutionary Families of Peptidase Inhibitors. *Biochem. J.* **2004**, *378* (3), 705−716.

(32) Feussner, I.; Kühn, H.; Wasternack, C. Lipoxygenase-Dependent Degradation of Storage Lipids. *Trends Plant Sci.* **2001**, *6* (6), 268−273.

(33) Szecsi, P. B. The Aspartic Proteases. *Scand J. Clin Lab Invest* **1992**, *52*, 5−22.

(34) Bhattacharya, M.; Jain, N.; Bhasne, K.; Kumari, V.; Mukhopadhyay, S. PH-Induced Conformational Isomerization of Bovine Serum Albumin Studied by Extrinsic and Intrinsic Protein Fluorescence. *J. Fluoresc* **2011**, *21* (3), 1083−1090.

(35) De Lauder, W. B.; Wahl, P. PH Dependence of the Fluorescence Decay of Tryptophan. *Biochemistry* **1970**, *9* (13), 2750−2754.

(36) White, A. Effect of PH on Fluorescence of Tyrosine, Tryptophan and Related Compounds. *Biochem. J.* **1959**, *71* (2), 217−220.

(37) Pizones Ruiz-Henestrosa, V. M.; Martinez, M. J.; Patino, J. M. R.; Pilosof, A. M. R. A Dynamic Light Scattering Study on the Complex Assembly of Glycinin Soy Globulin in Aqueous Solutions. *J. Am. Oil Chem. Soc.* **2012**, *89* (7), 1183−1191.

(38) Wu, P.; Ben, T.; Zou, H.; Chen, Y. PARAFAC Modeling of Dandelion Phenolic Compound Fluorescence Relation to Antioxidant

https://doi.org/10.1021/acs.jafc.5c01681
J. Agric. Food Chem. 2025, 73, 11288−11302

Properties. *Journal of Food Measurement and Characterization* **2022**, *16* (4), 2811−2819.
(39) Jakubíková, M.; Sádecká, J.; Hroboňová, K. Determination of Total Phenolic Content and Selected Phenolic Compounds in Sweet Wines by Fluorescence Spectroscopy and Multivariate Calibration. *Microchemical Journal* **2022**, *181*, No. 107834.
(40) Murugesan, R. C.; Choudhury, M. T. A.; Rozhin, A. 2D Excitation-Emission Fluorescence Mapping Analysis of Plant Food Pigments. *Food Chem.* **2023**, *418*, No. 135875.

https://doi.org/10.1021/acs.jafc.5c01681
J. Agric. Food Chem. 2025, 73, 11288−11302