Biochemical and physical–chemical characterisation of leaf proteins extracted from Cichorium endivia leaves

Maude Ducrocq, Marie-Hélène Morel, Marc Anton, Valérie Micard, Sylvain Guyot, Valérie Beaumal, Véronique Solé-Jamault, Adeline Boire

Biochemical and physical-chemical characterisation of leaf proteins extracted from *Cichorium endivia* leaves

Maude Ducrocq [1, 2], Marie-Hélène Morel [1], Marc Anton [2], Valérie Micard [1], Sylvain Guyot [2], Valérie Beaumal [2], Véronique Solé-Jamault [2], Adeline Boire [2,*]

[1] *Univ. Montpellier, INRAE, Institut Agro, UMR IATE, Montpellier, France*

[2] *INRAE, UR BIA, Rue Yvette Cauchois, F-44316, Nantes, France*

** Corresponding author, e-mail address: adeline.boire@inrae.fr*

Abstract

This study provides a detailed characterisation of a leaf protein concentrate (LPC) extracted from *Cichorium endivia* leaves using a pilot scale process. This concentrate contains 74.1% protein and is mainly composed of Ribulose-1,5-BISphosphate Carboxylase/Oxygenase (RuBisCO). We show that the experimentally determined extinction coefficient (around 5.0 cm^{-1} g^{-1} L depending on the pH) and refractive index increment (between 0.27 and 0.39 mL g^{-1}) are higher than the predicted ones (about 1.6 cm^{-1} g^{-1} L and 0.19 mL g^{-1}, respectively). In addition, the UV-visible absorption spectra show a maximum at 258 nm. These data suggest the presence of non-protein UV-absorbing species. Chromatographic separation of the concentrate components in denaturing conditions suggests that RuBisCO SC may be covalently bounded to few phenolic compounds. Besides, the solubility of LPC proteins is higher than 90% above pH 6. Such high solubility could make LPC a good candidate as a functional food ingredient.

Key-words:

Leaf protein concentrate, RuBisCO, UV-visible spectroscopy, Protein solubility

1. Introduction

Leaf proteins were first described by Rouelle in 1773. Researches on leaf proteins were stimulated in the early 1940s when Norman Pirie discussed the potential of leaf protein concentrates (LPC) as a source of proteins for human consumption to overcome wartime food shortages (Pirie, 1942). Proteins constitute between 15 to 20% of the leaf dry mass depending on the species. Leaf proteins are often classified according to their affinity with water. Water-insoluble proteins, also called "green proteins", are mainly embedded in the membrane of plant organelles and often associated with lipophilic pigments such as chlorophyll (Thornber, 1975). Water-soluble proteins, also called "white proteins", are mainly composed of chloroplastic and cytoplasmic enzymes. The first commercial LPC, produced from alfalfa leaves, was launched in 1967 and consisted in a sole green concentrate rich in proteins and xanthophylls intended for monogastric and poultry feeding (Knuckles, Spencer, Lazar, Bickoff, & Kohler, 1970). Once improved, the pilot-scale process allowed the fractionation of water-soluble and water-insoluble proteins, resulting in a green concentrate with a lower protein content and a white protein concentrate, suitable for human consumption (Edwards et al., 1975). This white LPC is mainly composed of a protein named Ribulose-1,5-BISphosphate Carboxylase/Oxygenase (RuBisCO).

RuBisCO is a key enzyme in photosynthetic carbon assimilation, catalysing the first step of CO_2 fixation in the Calvin cycle. RuBisCO is found in most autotrophic organisms from prokaryotes to eukaryotes. Despite its deficient specificity and its low carboxylase activity, RuBisCO fulfils its carbon fixation functions in plants thanks to its high level of expression in the growing leaf. The enzyme represents up to 30-50% of the soluble proteins in leaves. RuBisCO is therefore claimed to be the most abundant protein on Earth (Ellis, 1979). RuBisCO is an hexadecameric protein, consisting of 8 large chains (LC) and 8 small chains (SC) arranged around a four-fold axis. The quaternary structure of RuBisCO mainly relies on electrostatic interactions although it was shown that LC tends to cross-link by disulphide bridges in oxidative stress conditions (Mehta, Fawcett, Porath, & Mattoo, 1992). In addition to its abundance, it was mentioned from the early 1940s that RuBisCO was interesting for human nutrition because of its balanced amino acid profile (Pirie, 1942). Some authors even compare the nutritive value of RuBisCO to that of casein (Hood, Cheng, Koch, & Brunner, 1981).

Several LPC extraction processes have been described (Edwards et al., 1975; Knuckles et al., 1970). LPC have up to now not been used as an ingredient in human food due to the difficulty of combining the economic viability of the extraction process with the quality of the protein concentrate (reviewed in Chiesa & Gnansounou, 2011).

Recent advances in extraction process methods such as membranes technologies are encouraging research on the use of LPC as a functional ingredient in human food (Ducrocq, Boire, Anton, Micard, & Morel, 2020; Kiskini, 2017; Martin, Castellani, de Jong, Bovetto, & Schmitt, 2019). These studies agree that LPC show very interesting functional properties, especially gelling properties, for industrial applications. In addition, the production of LPC could contribute to the valorization of green agro-industrial waste. Leaves from crop residues, such as sugar beet leaves, or from plant processing discards, such as ready-to-eat salad leaves, may be use as raw materials for LPC production. This would contribute to an optimized biomass valorization while providing an alternative protein source. In this paper, we were interested in proteins extracted from Escarole (*Cichorium endivia)* leaves, a vegetable consumed as fresh or packaged Ready-To-Eat salad. Its chemical and nutritional composition depends greatly on the variety and on the cultivation conditions (Otalora et al., 2018). In average, the protein content is about 1-1.5% of the fresh leaves (Otalora et al., 2018). Therefore, significant amount of protein may be extracted from the by-product of ready-to-eat salad processing.

Numerous characterisation studies have been carried out on highly purified RuBisCO to understand its structure and its enzymatic activity to improve its carbon fixation rate. However, in the field of Food Science and Technology, the characterisation of the LPC is fragmented and, to our knowledge, no exhaustive study characterises in details the biochemical and physical-chemical properties of the LPC as a food ingredient. The present paper describes in detail the biochemical and physical-chemical characterisation of a leaf protein concentrate extracted from *Cichorium endivia* leaves using a pilot scale process. A combination of techniques based on biochemical assay, chromatography, and spectroscopy was used to investigate the composition of the LPC and the physical-chemical properties of RuBisCO, its major protein. The experimentally measured parameters are compared with values from the literature as well as with theoretical values computed from known RuBisCO sequences.

2. Material and methods

2.1 Leaf protein concentrate extraction and storage

Leaf protein concentrate (LPC) was provided by Florette (Lessay, France). LPC was extracted from *Cichorium endivia* leaves using the extraction conditions of the WO 2014/104880 patent but excluding the hydrophobic column adsorption step. This process is based on the classical process of protein extraction from leaves which has been studied for many years and widely described in the literature (Edwards et al., 1975; Knuckles et al.,

1970). Indeed, the process we use includes the classical steps of the leaf protein extraction process: i) separation of the fibres by pressing, ii) heat precipitation of the coloured proteins iii) concentration of the proteins of interest and removal of residual contaminants. Briefly, the juice was first extracted from raw material by pressing using a twin-screw press with addition of a buffering solution (1 M MES buffer pH 6.5) and a reducing agent (e.g. sodium metabisulphite) to avoid phenolic oxidation and cross-linking. Membrane proteins and pigments were then coagulated upon heating at 55°C for 20 minutes. After cooling and decantation of the coloured pellet, a cross-flow microfiltration step on a 0.2 µm membrane allows the sterilisation of the supernatant. The protein-rich supernatant was then concentrated by ultrafiltration using a 100 kDa molecular weight cut-off and subsequently diafiltrated to remove salts, phenolics, and other impurities. The liquid protein concentrate was then freeze-dried. At least 14 days before every experiment, the protein powder was placed in a humidity-controlled chamber with a K_2CO_3 saturated salt solution at 20 °C to maintain 43.2% relative humidity to ensure a constant moisture content. All experiments were performed with Milli- 100 Q water (Millipore Systems, Guyancourt, Molsheim, France) and all chemicals used were of analytical grade and were purchased from Sigma (Bornem, Belgium), VWR 99 international (Leuven, Belgium) or Carlo Erba (Peypin, France).

2.2 Amino acids composition and nitrogen to protein conversion factor determination

Amino acid analysis was carried out with an amino acid analyser (L-8900, Hitachi, Tokyo, Japan) as described in Margier et al. (2018). Briefly, four types of hydrolysis were performed before the analysis. Acid hydrolysis with 6 N HCl was performed for 24 hours at 110 °C to determine the amount of most amino acids. Oxidation with performic acid was performed before the hydrolysis to assay sulphur-amino acids. Leucine, Isoleucine and valine were quantified after an acid hydrolysis with 6 N HCl for 48 hours at 110 °C. The amount of tryptophan was determined after basic hydrolysis with 4 N $Ba(OH)_2$ for 16 hours at 110 °C. For each hydrolysis, norleucine was used as internal standard. The relative standard deviation for the concentration of each amino acid standard is comprised between 0.6 and 2.5 % with this method and equipment. From the amino acid composition, the chemical score of essential amino acids (EAA) was calculated as the ratio between the EAA content in the concentrate over its content in the reference amino acid profile established by the French Food Safety Agency (AFSSA, 2007). As white LPC is generally mainly composed of RuBisCO, the LPC amino acid composition was compared with the amino acid composition of RuBisCO from *lactuca sativa* was computed from the sequences of its subunits (accession numbers of LC and SC in UniProt database: P48706 and Q40250). The whole RuBisCO

sequence was obtained by adding the sequences of 8 large chains and 8 small chains, excluding the signal peptides.

The Nitrogen to Protein conversion factor (N:P factor) was computed from the ratio of total anhydrous mass of amino acids to the total mass of nitrogen, as described in Sosulski & Imafidon (1990). The anhydrous molecular weight of an amino acid corresponds to its molecular weight minus the molecular weight of a molecule of water (18 $g.mol^{-1}$). The Aspagarine (Asn) and Glutamine (Gln) were assayed in their acidic form. Therefore, they could not be distinguished from Asparagic acid (Asp) and Glutamic acid (Glu) content in the amino acid profile analysis. We estimated the content in Asn and Gln in LPC by taking the proportion of Asn/Asp and Gln/Glu from lettuce RuBisCO amino acid sequence.

2.3 Composition of the leaf protein concentrate

Water and ash contents of LPC powder were determined by thermogravimetric analyses (TGA 2050, TA instruments, New Castle, England) under nitrogen atmosphere. About 10 mg of powder was heated at 3 °C/min until 130 °C for 30 min. The measured mass loss was attributed to water content. The powder was further heated at 600 °C for 180 min to determine ash content. Measurements were done in triplicates. The total nitrogen content of LPC powder was determined using the Kjeldahl procedure (NF V 03–050, 1970). Protein content was obtained using a N:P factor determined from LPC amino acid composition.

Uronic acid was determined by an automated m-hydroxybiphenyl method (Thibault, 1979). Neutral sugars were analysed as to their alditol acetate derivatives by gas-liquid chromatography after hydrolysis (Englyst & Cummings, 1988). The total free phenolic compounds were determined using the Folin-Ciocalteu reagent. Polyphenols were extracted in acetone (80%) with a liquid:solid ratio of 40. The extraction was performed at 4°C in an ultrasounds bath for 45 min. The mixture was centrifuged (9000 x g, 15 min) and the extraction was repeated once. Both supernatants were combined and evaporated under nitrogen stream. Dry extracts were then diluted in ultrapure water so as to obtain an absorbance within the linearity range of the test. In a microplate, 25µL of sample was mixed with 125µL of Folin Ciocalteau (10%) and 100µL Na_2CO_3 (75 $g.L^{-1}$). The microplate was then incubated at 40 °C for 5 min before absorbance reading at 735 nm. Gallic acid was used as standard and results were expressed as gallic acid equivalent. Measurements were done in triplicates.

2.4 SDS-PAGE and Western Blot

SDS-PAGE was performed in non-reducing and reducing conditions. LPC was solubilised at 2 $g.L^{-1}$ in sodium phosphate buffer (0.1 M, pH 8, 1 mM EDTA) and diluted twice in Laemmli buffer. β-Mercaptoethanol at 50 $\mu g.mL^{-1}$ was added to the Laemmli buffer for electrophoresis under reducing conditions. Both unreduced and reduced samples were then heated at 100 °C for five min. Gels (4-12 % Bolt Bis-tris-Plus, 8 x 8 cm^2, Novex) were run in MES running buffer at a constant voltage of 200 V for 40 min in the Mini Gel Tank (ref A25977, ThermoFisher scientific). For the non-specific labelling of proteins, electrophoresis gel was incubated one hour in Coomassie blue stain (InstantBlue, Expedeon, San Diego, CA, USA), rinsed several times in distilled water and scanned.

A Western-Blot was performed using rabbit anti-RuBisCO large chain (anti-Rbcl, Agrisera) as described in O'Donnelly et al. (2014) with some modifications. Briefly, the gel was equilibrated in transfer buffer (25mM Tris, 192mM Glycine, SDS 0.1% (w/v), Ethanol 20% (v/v)) for 15 min after SDS-PAGE running. Polypeptides fractionated on SDS-PAGE gels were transferred to a nitrocellulose membrane (162-0112, Bio Rad) at 250 mA for 90 min (mini Trans-Blot cell, BIO RAD). The membrane was then washed with the saturating solution (5% (w/v) milk powder in PBS - 0.1% (v/v) Tween20) for 30 min at room temperature, and rinsed three times 10 min in PBS-tween buffer. The membrane incubation with primary antibody solution (rabbit anti-RbcL dissolved in PBS, 2% (w/v) milk powder at a ratio 1/10000) was performed for 1 hour. The membrane was then rinsed 3 times 10 min in PBS-Tween. The secondary Antibody Solution (AntiRabbit HRP, A-8025, sigma) was dissolved in PBS 1X, 2% milk powder at a ratio 1/25000. The membrane was incubated with the secondary Antibody Solution for 1 hour before being rinsed twice in PBS-tween and once in PBS. Western BrightTM Quantum (Advansta) reagent was mixed with the membrane for 2 min; the membrane was then imaged with the Fujifilm LAS 3000 camera.

2.5 UV-visible spectra

Determination of the specific extinction coefficient

The specific extinction coefficients of LPC dispersion were determined at pH 5, pH 7 and pH 8. To obtain the same ionic strength for all buffers, specific ratio of acid and base were used to achieve the desired pH. The molecular species used were as follows: acetate buffer (CH_3COOH/CH_3COO^-) 0.025 M pH 5, phosphate buffer ($H_2PO_4^-/HPO_4^{2-}$) 0.01 M pH 7, phosphate buffer ($H_2PO_4^-/HPO_4^{2-}$) 0.01 M pH 8. When needed, minimal amount of NaCl was added to reach an ionic strength of 25 mM. LPC was dispersed in buffer overnight at room temperature at 3 $mg.mL^{-1}$; dispersions were centrifuged the day after for 30 min at 39191 x g. The total nitrogen content of

supernatant was determined using the Kjeldahl procedure (NF V 03–050, 1970). Protein content was obtained using the N:P factor determined from the LPC amino acid composition.

The UV-visible spectra were recorded on the supernatant and after six dilutions (30, 12, 6, 4, 3 and 1.5 times) using a UV-compatible microplate and a plate reader (Spark®, Tecan Trading, Ltd**.,** Switzerland). The liquid pathlength in each well was determined as described in Lampinen, Raitio, Perälä, Oranen, & Harinen (2012).

Calculation of the theoretical UV-visible spectrum

The theoretical UV-visible spectrum of the LPC proteins was calculated from LPC amino acid profile. First, the concentration of tyrosine, tryptophan and phenylalanine was calculated for 1 $g.L^{-1}$ LPC dispersion. Then, the theoretical UV-visible spectrum was calculated as a linear combination of individual spectrum of each AA extracted from the photochemcad database (http://photochemcad.com/). Tyrosine and tryptophan spectra were measured in phosphate buffer (0.1 M, pH 7) and phenylalanine in water.

2.6 Refractive index increment

The dn/dc was determined at 589.3 nm using a multi-wavelength Abbe refractometer (Anton Paar GmbH, Graz/AUSTRIA) at 20 °C for LPC dispersions, similarly to the specific extinction coefficients, at pH 5 (25 mM acetate buffer), pH 7 (10 mM phosphate buffer) and pH 8 (10 mM phosphate buffer). Theoretical dn/dc was calculated from the amino acid composition of LPC according to Zhao, Brown, & Schuck (2011).

2.7 Protein solubility

LPC was solubilised at 5 g.L-1 in water for 30 min at room temperature. The pH was adjusted to the desired value with 0.1 or 0.5 M NaOH or HCl. Dispersions were mixed on a rotary shaker for 2 hours; the pH was checked and adjusted if needed at the end. Dispersions were then centrifuged (10 000 x g, 20 °C, 15 min) to recover soluble proteins. The nitrogen content of total dispersion and of the supernatant was analysed according to the Dumas method using a rapid MAX N exceed (Elementar, Langenselbold, Germany). About 1 mL of protein dispersion was placed in a steel crucible and injected into an oven under a stream of oxygen and helium. The combustion was operated at 900°C with an oxygen flow rate of 100 mL/min for 120 s. The gas mixture was then reduced in a reduction tube and separated on selective trap columns. Glycine was used as a standard for nitrogen content. Measurements were done in triplicates.

2.8 Electrophoretic mobility

To determine LPC isoelectric point, electrophoretic mobility was measured on LPC dispersions at pH ranging from 3 to 11. LPC was solubilised at 5 $g.L^{-1}$ in water for two hours before being centrifuged at 12 000 xg for 20 min. One milliliter of supernatant was adjusted to the desired pH using 0.1 M HCl or 0.1 M NaOH. Electrophoretic mobility was measured in triplicate at each pH using a zetasizer Nano Series (Nano-ZS, Malvern instrument, Germany). The conductivity was about 0.7 +/- 0.3 $mS.cm^{-1}$. The whole experiment was performed twice. The theoretical net charge of lettuce RuBisCO as a function of pH was computed using the webserver pdb2pqr (http://server.poissonboltzmann.org/pdb2pqr), for unfolded and folded spinach RuBisCO (ProteinDataBank entry for spinach RuBisCO: 1aus). The computation was performed on spinach RuBisCO because the 3D-structure of lettuce RuBisCO was not available. Note that the theoretical net charges of unfolded spinach RuBisCO and unfolded lettuce RuBisCO were similar.

2.9 Dynamic light scattering

Dynamic light scattering (DLS) measurements were performed using a zetasizer Nano Series (Nano-ZS, Malvern instrument, Germany). The measurements were performed at 20°C in manual mode: 10 runs of 30 sec were collected and repeated twice for each sample. From the correlation function, the size distribution was calculated using the "general purpose" analysis, provided by the commercial software. It consists in a distribution analysis with a sum of ideal exponential decays. All LPC dispersions were performed overnight, at room temperature before being centrifuged at 12 000 x g for 20 min. The protein concentration was determined after centrifugation using UV absorbance. To study the effect of ionic strength, LPC was dispersed in 0.01 M phosphate buffer pH 7 to which was added NaCl to achieve given ionic strengths: 25 mM, 50 mM, 100 mM or 150 mM. To determine the effect of pH, LPC was dispersed in MES buffer 0.03 M pH 6, phosphate buffer 0.01 M pH 7 or phosphate buffer 0.01 M pH 8. When needed, minimal amount of NaCl was added to reach a ionic strength of 25 mM. The effect of protein concentration was measured on dispersions at 1.5 $g.L^{-1}$, 3 $g.L^{-1}$ and 9 $g.L^{-1}$ in phosphate buffer 0.01 M pH 7. For each physical-chemical condition, three samples were prepared and measured.

2.10 Size-exclusion chromatography analysis

LPC proteins were extracted as described in Ducrocq et al. (2020) with some modifications. Proteins were first extracted in 0.1 M sodium phosphate buffer pH 6.8 with 1% SDS (v/v) at a solid to liquid ratio of 8 $g.L^{-1}$. Extraction was performed on a rotary shaker set at 60 rpm at 60 °C for 80 min. The supernatant containing SDS-soluble proteins was recovered after centrifugation (39 191 x g, 30 min, 20 °C). The SE-HPLC apparatus (Waters model

LC Module1 plus) was equipped with an analytical column, TSK G4000-SW (7.5 x 300 mm) and a guard column, TSK G3000-SW (7.5 x 75 mm) (both from Merck, Darmstadt, Germany). The columns were eluted at 25°C with 0.1 M sodium phosphate buffer pH 6.8 containing 0.1% SDS (v/v). The flow rate was 0.7 $mL.min^{-1}$ and absorbance was recorded from 210 nm to 700 nm.

3. Results and discussion

3.1 Composition of the leaf protein concentrate

The composition of the LPC was investigated through the analysis of the amino acid composition and of the nitrogen, ash and sugar contents. Results are expressed on a dry matter basis. Amino acid composition of LPC is detailed in Table 1 and is compared to the theoretical amino acid composition of lettuce RuBisCO. Experimental data obtained on the LPC are very similar to the calculated amino acid composition of lettuce RuBisCO, except for the cysteine content. A 2.5 fold higher amount of cysteine than the amount calculated for lettuce RuBisCO is found in LPC. According to the literature, this difference in cysteine amount is not related to species (Gerloff, Lima, & Stahmann, 1965). The LPC amino acid composition was used to compute the chemical score (Supplementary Table S1). Chemical scores of all essential amino acids are above 100%, meaning that LPC includes all essential amino acids needed for human diet. The composition in amino acids allows the calculation of the N:P factor, which is essential for calculating the protein content from a nitrogen assay. From the amino acid composition described in Table 1, we obtained a nitrogen-to-protein conversion factor of 5.79. This value is within the range of nitrogen-to-protein conversion factor estimated for the leaf proteins of 90 plant species, which varies from 5.15 to 5.93 (Yeoh & Wee, 1994). The nitrogen content of LPC is of 12.8% according to Kjeldahl method, which leads to a protein content of 74.1%. LPC contains about 8.2% neutral sugars and 1.2% uronic acid (Supplementary Table S2). The significant content of sugars, detected in the LPC, probably results from leaf polysaccharides. Besides neutral sugars, the presence of uronic acid suggests the presence of pectin (Kiskini, 2017). The LPC contains less than 5 mg of free phenolics (in gallic acid equivalents) in 100 g of powder. Ashes represent 7.8% of the LPC mass and 8.7% of LPC dry mass is of unknown origin.

3.2 Evaluation of protein purity and subunit molecular weights

LPC dispersions were analysed by SDS-PAGE under reduced and non-reduced conditions to evaluate the protein purity and determine the molecular weight of the major proteins in denaturing conditions. Resulting gels are shown in Figure 1A. Three major bands are observed on the SDS-PAGE gels in both non-reducing and reducing

conditions at 50, 40 and 14 kg.mol^{-1}. Bands at 50 and 14 kg.mol^{-1} may respectively correspond to the large-chain (LC) and the small-chain (SC) of RuBisCO. Both bands at 50 and 40 kg.mol^{-1} are revealed by anti-Rbcl (RuBisCO LC) antibody (Figure 1B) during the Western Blot analysis. This confirms the presence of the RuBisCO LC at 50 kg.mol^{-1} and shows that the band at 40 kg.mol^{-1} also contains RuBisCO LC. It may correspond to a hydrolysed form of RuBisCO LC, as previously suggested by Hood et al. (1981). Several small bands of medium intensity are observed in non-reducing conditions around 100 kg.mol^{-1}. Their intensity decreases upon reduction and they react with the antibody against RuBisCO LC. These bands may correspond to disulphide bonded dimers of LC. Disulphide linked dimers of RuBisCO LC have already been identified *in vivo* and *in vitro* (Mehta et al., 1992; Rintamaki, 1989) and related to an oxidative stress *in planta* and/or during the extraction process. A low intensity band is observed at ~35 kg.mol^{-1}. Anti-Rbcl antibodies reveal this band and the intensity of the band increases after reduction. This band may also correspond to a hydrolysed form of LC, which would be part of the disulphide bonded aggregates in non-reducing conditions. Apart from identified RuBisCO bands, less intense bands were observed between 10 and 55 kg.mol^{-1}. It is not possible to precisely quantify the RuBisCO purity of the LPC based on SDS-PAGE patterns due to the protein-type dependency of Coomassie staining (Fountoulakis, Juranville, & Manneberg, 1992). However, the SDS-PAGE analysis highlights that RuBisCO is the major protein of LPC.

3.3 Determination of the UV-visible spectroscopic and refractometric parameters

In the wavelength range of 235-300 nm, the UV-visible spectrum of a protein is mostly driven by its content in UV-absorbing amino acids namely tyrosine, tryptophan, and phenylalanine. The UV-visible spectrum of a protein also varies with the environment of the aromatic amino acids and their solvent exposure but this variation is expected to be small (<5%) (Pace, Vajdos, Fee, Grimsley, & Gray, 1995). Therefore, from the proportions of the three UV-absorbing amino acids in LPC and their respective individual spectra, a theoretical UV-visible spectrum of the protein was calculated for a 1 g.L^{-1} LPC dispersion in aqueous condition (Figure 2A). LPC theoretical spectrum shows a maximum of absorbance at 275 nm due to a major contribution in tryptophan and tyrosine. Experimental UV-visible spectra of LPC dispersion in water at several protein concentrations are represented in Figure 2B. Above 245 nm, the absorbance reaches a maximum at 258 nm and displays residual absorption between 300 and 420 nm. Other spectra obtained at pH 5, pH 7 and pH 8 present the same profile and maximum absorbance around 258 nm (Supplementary Figure S1). This suggests the presence of non-protein compounds that absorb around 260 nm and 300-450 nm. Phenolic compounds are common non-protein contaminants in

LPC (Pedone, Selvaggini, & Fantozzi, 1995). Among them, hydroxycinnamates and flavonols are well-known phenolic compounds previously identified in *Cichorium endivia* leaves exhibiting maximum absorbance in the 300-330 nm and in the 340-370 nm ranges, respectively (Mascherpa, Carazzone, Marrubini, Gazzani, & Papetti, 2012). Nucleic acids absorb in the range 230-300 nm with a maximum of absorbance around 260 nm (Groves, W.E., Davis, F.C., & Sells, B.H., 1968). They could contribute to the UV-visible spectra of LPC dispersions below 300 nm. Absorbance at higher wavelength (400-420 nm) may also correspond to brown-yellow phenolic oxidation products that would have been generated during the preparation of the LPC. Some of them, such as rutin, have high absorption around 260 nm and 400 nm and could correspond to the non-protein UV-absorbing species (Mirgorod, Borodina, & Borsch, 2013).

These results point to a significant contribution of non-protein species to the LPC absorbance at 280 nm. Therefore, the use of extinction coefficients values found in literature or calculated from the amino acid sequence is not adapted to LPC dispersions. The specific extinction coefficient at 280 nm was determined for LPC in buffers at pH 5, pH 7 and pH 8. The linearity of Beer Lambert's law was checked in triplicate for each physical-chemical condition (inset in Figure 2B). The specific extinction coefficients are of 4.96 cm^{-1} g^{-1} L at pH 5, 5.82 cm^{-1} g^{-1} L at pH 7 and 4.58 cm^{-1} g^{-1} L at pH 8. Several studies carried out on purified RuBisCO report specific extinction coefficients in the range of 1.41 to 1.82 cm^{-1} g^{-1} L depending on the species and the method of extraction of the proteins (reviewed in Douillard & de Mathan, 1994). Moreover, the specific extinction coefficient calculated from the LPC amino acid profile is 0.94 $cm^{-1}.g^{-1}$.L. This latter value is quite lower than the previously reported values, which may be due to the difficulty of assaying tryptophan in protein concentrates (Oelshlegel, Schroeder, & Stahmann, 1970). These values are much lower than those measured in our study, supporting the presence of non-protein UV-absorbing species in the LPC.

We computed the dn/dc from the amino acid composition of LPC and obtained 0.189 $mL.g^{-1}$ at 589.3 nm and 25°C. This value is close to the consensus value of 0.185 $mL.g^{-1}$ generally accepted for proteins (Zhao et al., 2011). The dn/dc values measured on LPC dispersions are of 0.27 at pH 8, 0.31 at pH 7 and 0.39 at pH 5. These values are high as compared to calculated values, suggesting the presence of non protein species. Polysaccharides and nucleic acids have dn/dc values of 0.15 and 0.17-0.19 (Theisen, Johann, Deacon, & Harding, 2000). Tannins are expected to have dn/dc around 0.25-0.26 mL/g at 620 nm (Vernhet A., Dubascoux S., Cabane B., Fulcrand H., Dubreucq E., Poncet-Legrand C., 2011). The presence of polysaccharides, nucleic acids or tannins seems

298 therefore not to explain such high dn/dc values. To our knowledge, no experimental dn/dc are available for
299 flavonol and hydroxycinnamates that are compatible with the UV-visible spectra (e.g. caffeic acid).

300 *3.4 Insight into the interactions between UV-absorbing species and RuBisCO*

301 To identify possible covalent bound between RuBisCO and the UV-absorbing species, LPC proteins were
302 extracted in a denaturing buffer containing SDS and analysed using size-exclusion HPLC. Signal was recorded
303 from 200 to 700 nm. The elution profile obtained at 214 nm is represented on Figure 3A. Three major peaks are
304 visible. The main fraction was eluted with an apparent molecular weight of about 15 kg.mol^{-1} and was attributed
305 to RuBisCO SC. The peak eluted at 14.3 min (~45 kg.mol^{-1}) was attributed to the RuBisCO LC and the peak eluted
306 at 13.15 min (~100 kg.mol^{-1}) corresponds to the LC dimer. A shoulder is visible between the LC and the SC peaks,
307 corresponding to a specie of about 28 kg.mol^{-1}. Smaller peaks are detected below 10 kg.mol^{-1}, they correspond
308 to small peptides or non-protein species.

309 UV-visible spectra recorded for these four main peaks are represented on Figure 3B. Spectra associated with
310 RuBisCO LC and LC dimer have the same profile with a maximum absorbance at 276 nm. This wavelength is close
311 to the wavelength of the maximum absorbance of the theoretical spectrum of LPC (Figure 2A). This result
312 suggests that RuBisCO LC and LC-dimer are not covalently linked with any UV-absorbing species. In contrast, the
313 absorption spectrum associated with the RuBisCO-SC displays a maximum at 265 nm. Despite the absence of
314 molecular weight change, few phenolic compounds may be covalently-bound to RuBisCO-SC. In addition,
315 molecules eluted before and after the SC have high absorbance at 260 nm resulting in high $A_{260\ nm}/A_{214\ nm}$ ratio
316 (0.37 and 0.79 as compared to 0.16 for SC peak). The distortion of the RuBisCO SC spectrum may also be due to
317 the absorbance of these two adjacent species. The absorption spectra associated with the peak detected around
318 28 kg.mol^{-1} displays a maximum at 259 nm. A similar 28 kg.mol^{-1} molecule is also detected on a commercial
319 RuBisCO (Sigma-Aldrich, St. Louis, MO). Its maximum absorbance is measured at 277 nm, supporting a proteic
320 nature (Supplementary Figure S2).

321 All these data suggest that RuBisCO LC is not covalently bonded with a UV-absorbing species. In contrast, RuBisCO
322 SC and a protein specie of about 28 kg.mol^{-1} may be covalently bounded to few phenolic compounds.

323 *3.5 Effect of pH on protein solubility and surface charge*

The solubility of LPC proteins was determined as a function of pH. For pH > 6, the solubility was higher than 90% as shown in Figure 4A. A minimum of solubility was obtained for pH around 4. The overall features of the solubility curve are consistent with literature data on LPC proteins (Kiskini, 2017). The minimum of solubility of the LPC proteins is consistent with the electrophoretic mobility measurements as presented in Figure 4B (empty symbols). The isoelectric point (IEP) of LPC dispersion is around pH 3.8 and it has a negative charge for pH > 4.0. The IEP of the LPC is in the range of sugar beet leaves protein concentrate as determined by zeta potential measurement (Kiskini, 2017). The IEP of the LPC is however lower than other experimental values obtained using isoelectric focusing: between 5 and 5.5 for spinach RuBisCO (Iwanij, Chua, & Siekevitz, 1974; Matsumoto, Sugiyama, & Akazawa, 1969), around 7 for lettuce RuBisCO (Rabinowitz, Reisfeld, Sagher, & Edelman, 1975), and 6 for alfalfa RuBisCO (reviewed in Douillard & de Mathan, 1994). The isoelectric point of maize RuBisCO was reported around 4.6 as obtained by isoelectric focusing (Reger, Ku, Pottert, & Evans, 1983). For comparison, the pH-dependence of the calculated total net charge of spinach RuBisCO is reported (full symbols). For unfolded spinach RuBisCO, an equal number of positive and negative charges, the isoionic point (IIP), is expected at pH 6.6. The IIP of spinach RuBisCO computed with PROPKA software (version 3.0) is 6.0 instead of 6.6, a shift related to RuBisCO conformation. The experimental IEP of LPC is much lower than the theoretical IIP of RuBisCO. Such discrepancy may be ascribed to the presence of the absorbed species previously highlighted.

In a previous work, it was suggested that the discrepancy between experimental IEP and theoretical IIP on LPC dispersions was related to the presence of pectin (Kiskini, 2017). Pectin that have been highlighted in our LPC may indeed reduce the IIP of RuBisCO. Moreover, several studies report an acidification of proteins after their derivatisation by phenolic compounds. Higher degrees of derivatisation induced lower IEP of soybean glycinins (Rawel, Czajka, Rohn, & Kroll, 2002) and of lysozyme (Rawel, Kroll, & Rohn, 2001).

3.6 Impact of pH on the size distribution of proteins

DLS measurements were performed to highlight the effect of pH on LPC dispersion hydrodynamic properties. Representative correlation functions are shown in Figure 5A. Two major peaks are found with a mean hydrodynamic radius (R_h) of 7.5 and 45 nm, respectively (Figure 5B). The smallest size population could correspond to RuBisCO hexadecamer since the R_h is close the Stockes radius of 6.4 nm estimated for RuBisCO from citrus leaves (Penarrubia & Moreno, 1988) and of 7.4 nm from oat (Steer, M.W., Gunning, B.E., Graham, T.A., & Carr, D.J., 1968). This is also in the range of the one computed for lettuce RuBisCO using the Hullrad server

(6.4-6.5 nm). An increase in pH from 6 to 8 favors the proportion of smaller particles. Electrostatic repulsions at basic pH, as suggested by the electrophoretic mobility results, may prevent protein aggregation. At the other pH, part of the RuBisCO may no longer be in the form of hexadecamer. The protein concentration, up to 9.5 $g.L^{-1}$, and the ionic strength, up to 150 mM, do not affect the size distribution of the LPC dispersion (Supplementary Figure S4).

4. Conclusion and perspectives

We showed that the LPC has a protein content higher than 74% and mainly consists of RuBisCO. The presence of non-protein compounds was evidenced, mostly pectin and UV-absorbing species. These compounds affect the extinction coefficient, the refractive index increment and the surface charge of RuBisCO. Despite these changes, a solubility higher than 90% was observed for pH > 6.0. Such a high solubility may provide interesting techno-functionalities for the use of LPC as a food ingredient. Beyond the characterisation of a leaf protein concentrate, this study raises the effect of minor compounds on the spectroscopic and physical-chemical properties of protein ingredients. This is especially true for plant-based protein concentrate that are increasingly investigated to replace proteins of animal origins. Since such minor compounds may affect the techno-functional properties of protein ingredients, a thorough characterisation of such ingredients would help in better understanding and controlling their use in food applications.

Acknowledgments

This study was conducted in the framework of the EU funded GreenProtein BBI-JU project (Grant Agreement No 720728). The authors would like to thank Joëlle Bonicel (IATE, INRAE) for her help in performing SE-HPLC analyses, Marie-Jeanne Crépeau (BIA, INRAE) for performing totals sugar assay, Nathalie Geneix for her help in performing Western-Blot and Noureddine Hafnaoui (UNH, INRAE) and Didier Remond (UNH, INRAE) for performing amino acid composition determination.

Funding sources

This work was supported by the Institut national de la recherche agronomique (National Institute of Agricultural Research) in the framework of the EU funded GreenProtein BBI-JU project (Grant Agreement No 720728).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Maude Ducrocq: Conceptualization, Methodology, Investigation, Writing - original draft, Writing - review & editing. Marie-Hélène Morel: Conceptualization, Methodology, Writing - review & editing. Marc Anton: Project administration, Funding acquisition. Valérie Micard: Writing - review & editing, Project administration, Funding acquisition. Sylvain Guyot: Conceptualization, Writing - review & editing. Valérie Beaumal: Methodology, Investigation. Véronique Solé-Jamault: Investigation, Resources. Adeline Boire: Conceptualization, Methodology, Investigation, Writing - original draft, Writing - review & editing, Supervision.

References

AFSSA. (2007). Apport en protéines : consommation, qualité, besoins et recommandations.

Chiesa, S., & Gnansounou, E. (2011). Protein extraction from biomass in a bioethanol refinery - Possible dietary applications: Use as animal feed and potential extension to human consumption. Bioresource Technology, 102(2), 427–436. https://doi.org/10.1016/j.biortech.2010.07.125

D'alvise, N., Lesueur-Lambert, C., Fertin, B., Dhulster, P., & Guillochon, D. (2007). Removal of Polyphenols and Recovery of Proteins from Alfalfa White Protein Concentrate by Ultrafiltration and Adsorbent Resin Separations. Separation Science and Technology, 35(15), 2453–2472. https://doi.org/10.1081/

Douillard, R., & de Mathan, O. (1994). Leaf protein for food use: potential of Rubisco. In B. J. F. Hudson (Ed.), New and Developping Sources of Food Proteins (pp. 307–342). Springer, Boston, MA.

Ducrocq, M., Boire, A., Anton, M., Micard, V., & Morel, M.-H. (2020). Rubisco: A promising plant protein to enrich wheat-based food without impairing dough viscoelasticity and protein polymerisation. Food Hydrocolloids, 109(June). https://doi.org/10.1016/j.foodhyd.2020.106101

Edwards, R. H., Miller, R. E., de Fremery, D., Knuckles, B. E., Bickoff, E. M., & Kohler, G. O. (1975). Pilot Plant Production of an Edible White Fraction Leaf Protein Concentrate from Alfalfa. Journal of Agricultural and Food Chemistry, 23(4), 620–626. https://doi.org/10.1021/jf60200a046

Ellis, R. J. (1979). The most abundant protein in the world. Trends in Biochemical Sciences, 4(11), 241–244. https://doi.org/10.1016/0968-0004(79)90212-3

Englyst, H. N., & Cummings, J. H. (1988). Improved method for measurement of dietary fiber as non-starch polysaccharides in plant foods. Journal - Association of Official Analytical Chemists, 71(4), 808–814. https://doi.org/10.1093/jaoac/71.4.808

Fountoulakis, M., Juranville, J. F., & Manneberg, M. (1992). Comparison of the Coomassie brilliant blue, bicinchoninic acid and Lowry quantitation assays, using non-glycosylated and glycosylated proteins. Journal of Biochemical and Biophysical Methods, 24(3–4), 265–274. https://doi.org/10.1016/0165-022X(94)90078-7

411 Gerloff, E. D., Lima, I. H., & Stahmann, M. A. (1965). Leal Proteins as Foodstuffs - Amino Acid Composition of Leaf
412 Protein Concentrates. Journal of Agricultural and Food Chemistry, 13(2), 139–143.
413 https://doi.org/10.1021/jf60138a012

414 Groves, W.E., Davis, F.C., & Sells, B.H. (1968) Spectrophotometric Determination of Microgram Quantities of
415 Protein without Nucleic Acid Interference. Analytical Biochemistry, 22, 195-210.

416 Hood, L. L., Cheng, S. G., Koch, U., & Brunner, J. R. (1981). Alfalfa Proteins: Isolation and Partial Characterization
417 of the Major Component — Fraction I Protein. Journal of Food Science, 46(6), 1843–1850.
418 https://doi.org/10.1111/j.1365-2621.1981.tb04501.x

419 Iwanij, V., Chua, N., & Siekevitz, P. (1974). The purification and some properties of ribuloseBisphosphate
420 Carboxylase and of its subunits from the green alga chlamydomonas reinhardtii. Biochimica et Biophysica Acta,
421 358, 329–340.

422 Kiskini, A. (2017). Sugar Beet Leaves: from biorefinery to techno-function. Wageningen University.
423 https://doi.org/10.18174/421994

424 Knuckles, B. E., Spencer, R. R., Lazar, M. E., Bickoff, E. M., & Kohler, G. O. (1970). PRO-XAN Process: Incorporation
425 and Evaluation of Sugar Cane Rolls in Wet Fractionation of Alfalfa. Journal of Agricultural and Food Chemistry,
426 18(6), 1086–1089. https://doi.org/10.1021/jf60172a001

427 Lampinen, J., Raitio, M., Perälä, A., Oranen, H., & Harinen, R. (2012). Correction Method for Photometric DNA
428 Quantifi cation Assay. Thermo Fisher Scientific Aplication Notes, 1–7.

429 Margier, M., Georgé, S., Hafnaoui, N., Remond, D., Nowicki, M., Du Chaffaut, L., … Reboul, E. (2018). Nutritional
430 composition and bioactive content of legumes: Characterization of pulses frequently consumed in France and
431 effect of the cooking method. Nutrients, 10(11), 1–12. https://doi.org/10.3390/nu10111668

432 Martin, A. H., Castellani, O., de Jong, G. A. H., Bovetto, L., & Schmitt, C. (2019). Comparison of the functional
433 properties of RuBisCO protein isolate extracted from sugar beet leaves with commercial whey protein and soy
434 protein isolates. Journal of the Science of Food and Agriculture, 99(4), 1568–1576.
435 https://doi.org/10.1002/jsfa.9335

436 Mascherpa, D., Carazzone, C., Marrubini, G., Gazzani, G., & Papetti, A. (2012). Identification of phenolic
437 constituents in Cichorium endivia var. crispum and var. latifolium salads by high-performance liquid
438 chromatography with diode array detection and electrospray ioniziation tandem mass spectrometry. Journal of
439 Agricultural and Food Chemistry, Dec 12(60(49)).

440 Matsumoto, C., Sugiyama, T., & Akazawa, T. (1969). Structure and Function of Chloroplast Proteins IX . Further
441 Comparative Studies on Chlorella and Spinach Leaf Ribulose-1.5-Diphosphate Carboxylase. Archives of
442 Biochemistry and Biophysics, 135(1).

443 Mehta, R. A., Fawcett, T. W., Porath, D., & Mattoo, A. K. (1992). Oxidative stress causes rapid membrane
444 translocation and in vivo degradation of ribulose-1,5-bisphosphate carboxylase/oxygenase. Journal of Biological
445 Chemistry, 267(4), 2810–2816.

446 Mirgorod, Y. A., Borodina, V. G., & Borsch, N. A. (2013). Investigation of interaction between silver ions and rutin
447 in water by physical methods. Biophysics (Russian Federation), 58(6), 743–747.
448 https://doi.org/10.1134/S0006350913060146

449 O'Donnelly, K., Zhao, G., Patel, P., Butt, M. S., Mak, L. H., Kretschmer, S., … Barter, L. M. C. (2014). Isolation and
450 kinetic characterisation of hydrophobically distinct populations of form I Rubisco. Plant Methods, 10(1), 1–11.
451 https://doi.org/10.1186/1746-4811-10-17

452 Otálora, G., Piñero, M. C., López-Marín, J., Varó, P., & del Amor, F. M. (2018). Effects of foliar nitrogen fertilization
453 on the phenolic, mineral, and amino acid composition of escarole (Cichorium endivia L. var. latifolium). Scientia
454 Horticulturae, 239(January), 87–92. https://doi.org/10.1016/j.scienta.2018.05.031

455 Pace, C. N., Vajdos, F., Fee, L., Grimsley, G., & Gray, T. (1995). How to measure and predict the molar absorption
456 coefficient of a protein. Protein Science, 4(11), 2411–2423. https://doi.org/10.1002/pro.5560041120

457 Pedone, S., Selvaggini, R., & Fantozzi, P. (1995). Leaf protein availability in food: Significance of the binding of
458 phenolic compounds to ribulose-1,5-diphosphate carboxylase. LWT - Food Science and Technology, 28(6), 625–
459 634. https://doi.org/10.1016/0023-6438(95)90012-8

460 Penarrubia, L., & Moreno, J. (1988). Ribulose 1,5-bisphosphate carboxylase/oxygenase from citrus leaves.
461 Phytochemistry, 27(7), 1999-2004. https://doi.org/10.1016/0031-9422(88)80085-2.

462 Pirie, N. W. (1942). Green leaves as a source of proteins and other nutrients. Nature, 149(3774), 251.
463 https://doi.org/10.1038/149251a0

464 Rabinowitz, H., Reisfeld, A., Sagher, D., & Edelman, M. (1975). Ribulose diphosphate carboxylase from
465 autotrophic microorganisms. Plant Physiology, 56, 345–350. https://doi.org/10.1128/jb.110.2.633-642.1972

466 Rawel, H. M., Czajka, D., Rohn, S., & Kroll, J. (2002). Interactions of different phenolic acids and flavonoids with
467 soy proteins. International Journal of Biological Macromolecules, 30(3–4), 137–150.
468 https://doi.org/10.1016/S0141-8130(02)00016-8

469 Rawel, H. M., Kroll, J., & Rohn, S. (2001). Reactions of phenolic substances with lysozyme - Physicalchemical
470 characterisation and proteolytic digestion of the derivatives. Food Chemistry, 72(1), 59–71.
471 https://doi.org/10.1016/S0308-8146(00)00206-5

472 Reger, B. J., Ku, M. S. B., Pottert, J. W., & Evans, J. J. (1983). Purification and characterization of maize ribulose-
473 1,5-bisphosphate carboxylase. Phytochemistry, 22(5), 1127–1132. https://doi.org/10.1016/0031-
474 9422(83)80205-2

475 Rintamaki, E. (1989). Formation of disulphide cross-linked aggregates of large subunit from higher plant ribulose-
476 1, 5-Bisphosphate carboxylase-oxygenase. Journal of Experimental Botany, 40(12), 1305–1313.
477 https://doi.org/10.1093/jxb/40.12.1305

478 Sosulski, F. W., & Imafidon, G. I. (1990). Amino acid composition and nitrogen-to-protein conversion factors for
479 animal and plant foods. J.Agric.Food Chem, 38(6), 1351–1356. https://doi.org/10.1021/jf00096a011

480 Steer, M.W., Gunning, B.E., Graham, T.A., & Carr, D.J. (1968) Isolation, properties, and structure of fraction I
481 protein from Avena sativa L. Planta, 79(3),254-67. https://doi.org/10.1007/BF00396032

482 Theisen, A., Johann, C., Deacon, M. P., & Harding, S. E. (2000). Refractive Increment Data-Book for Polymer and
483 Biomolecular scientists (Nottingham). Nottigham UK.

484 Thibault, J. F. (1979). Automatisation du Dosage des Substances Pectiques par la Méthode au Méta-
485 hydroxydiphenyl. Lebensmittel-Wissenschaft Und-Technologie, 12, 247–251.

Thornber, J. P. (1975). Chlorophyll-proteins: Light-harvesting and reaction center components of plants. Annual Review of Plant Physiology, 26, 127–158.

Vernhet A., Dubascoux S., Cabane B., Fulcrand H., Dubreucq E., Poncet-Legrand C. (2011) Characterization of Oxidized Tannins: Comparison of Depolymerization Methods, Asymmetric Flow Field-Flow Fractionation and Small-Angle X-Ray Scattering. Anal. Bioanal. Chem., 401 (5), 1559–1569. 10.1007/s00216-011-5076-2.

Yeoh, H. H., & Wee, Y. C. (1994). Leaf protein contents and nitrogen-to-protein conversion factors for 90 plant species. Food Chemistry, 49(3), 245–250. https://doi.org/10.1016/0308-8146(94)90167-8

Zhao, H., Brown, P. H., & Schuck, P. (2011). On the Distribution of Protein Refractive Index Increments. Biophysical Journal, 100(9), 2309–2317. https://doi.org/10.1016/j.bpj.2011.03.004

Table captions

Table 1 Amino acid composition determined on LPC in $mg.g^{-1}$ of protein as compared to calculated amino acid composition of lettuce RuBisCO (based on lettuce RuBisCO sequence found in UniProt database: *lactuca sativa*; accession numbers of LC and SC: P48706 and Q40250).

Figure captions

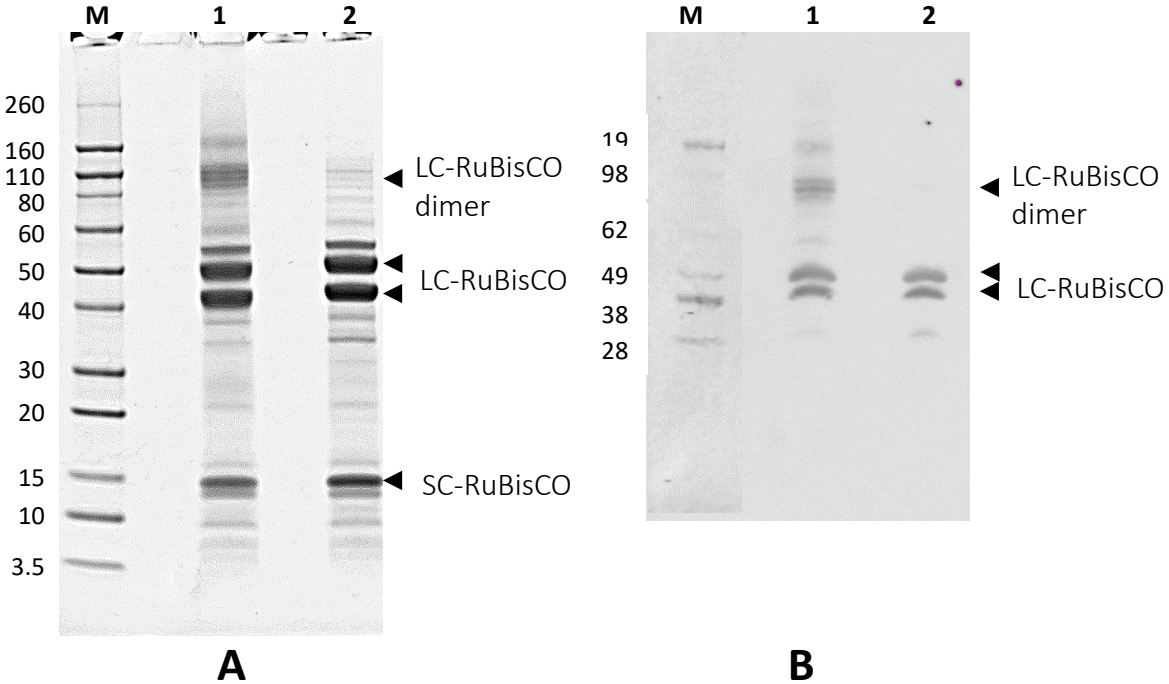
Figure 1 SDS-PAGE (**A**) and Western blot (**B**) patterns of LPC in non-reducing (lane 1) and reducing conditions (lane 2). Arrows indicate RuBisCO sub-units. LC-RuBisCO stands for RuBisCO large chain and SC-RuBisCO stands for RuBisCO small chain. Lane M corresponds to molecular weight standards in $kg.mol^{-1}$.

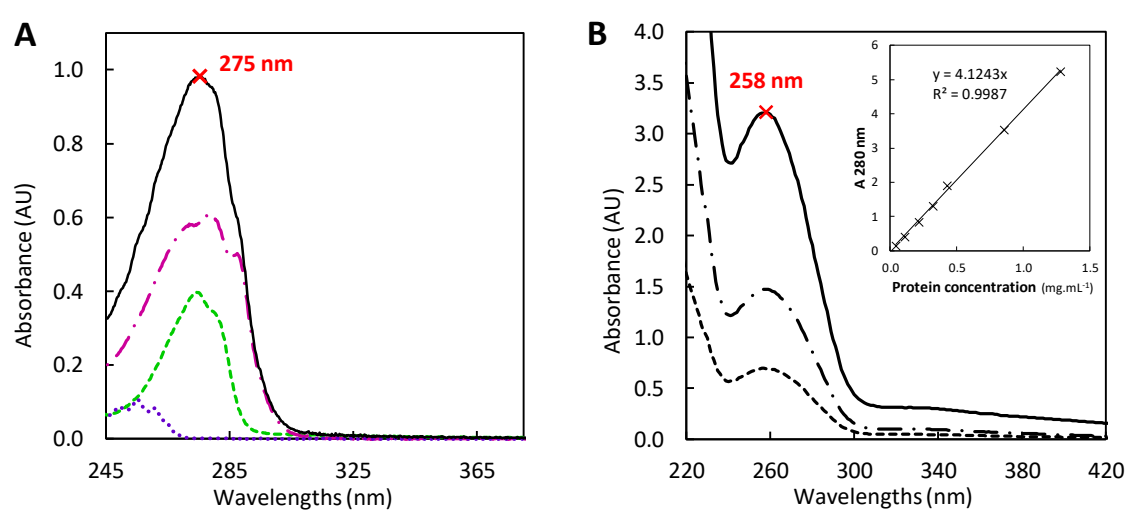
Figure 2 A. UV-visible spectra calculated for phenylalanine (dotted line), tyrosine (dashed line) and tryptophan (dashdotted line) for a 1 $g.L^{-1}$ LPC dispersion; and total calculated UV-visible spectra (solid line) for a 1 $g.L^{-1}$ LPC dispersion. Calculations are based on the amino acid composition of LPC. **B**. UV-visible spectra of LPC dispersion in water at several protein concentrations: 0.43 (full line), 0.21 (dashdotted line) and 0.11 (dashed line) $mg.mL^{-1}$. The inset represents the Beer-Lambert law established for RuBisCO in water.

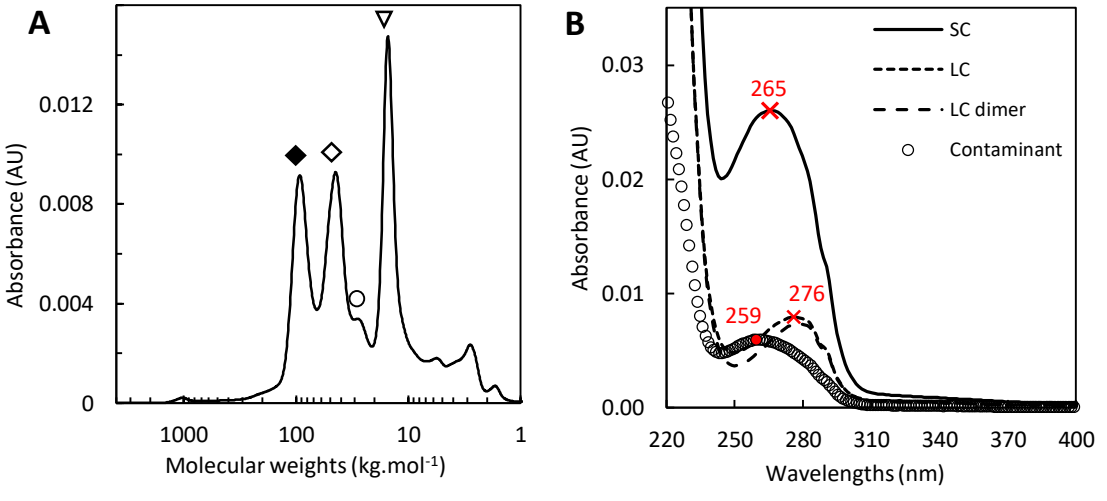
Figure 3 A. Elution profile, recorded at 214 nm, of SDS soluble proteins extracted from LPC. Symbols represent peaks of interest: RuBisCO small chain (SC; ▽), possible UV-absorbing species (o), RuBisCO large chain (LC; ◊) and RuBisCO large chain dimer (LC dimer; ♦). **B**. UV-visible spectra acquired at the elution times corresponding to the peaks at 214 nm of SC, contaminant, LC and LC dimer.

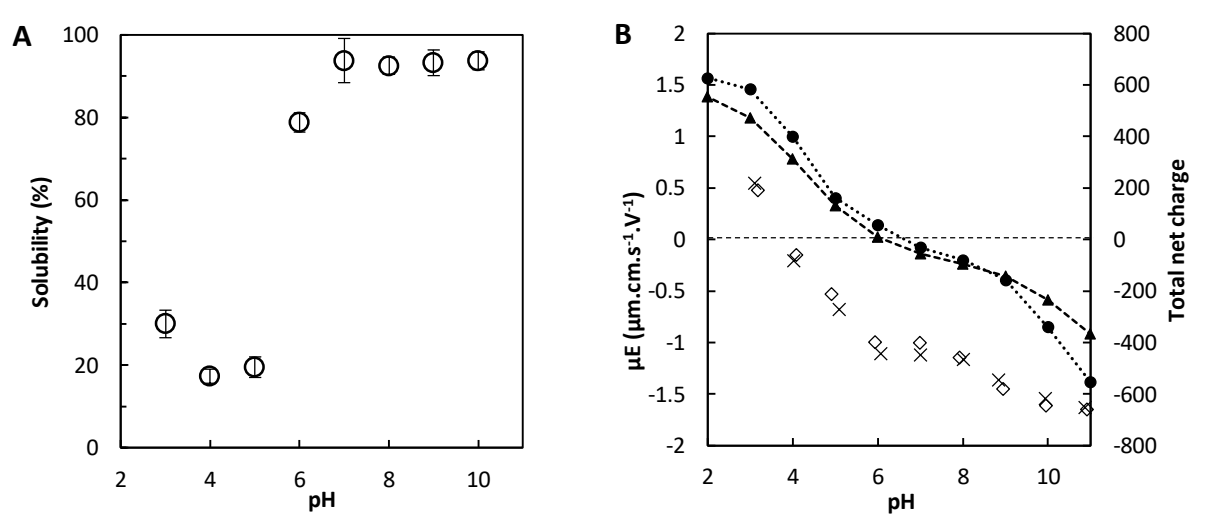
Figure 4 A. Evolution of the LPC protein solubility in water from pH 3 to pH 10. **B.** Experimental electrophoretic mobility (empty symbols) of LPC dispersion (at 2.9 $g.L^{-1}$ for diamonds and 2.8 $g.L^{-1}$ for crosses) according to pH and estimated total net charge (full symbols) of unfolded (circles, pointed line) and folded (triangles, dashed line) spinach RuBisCO. Spinach RuBisCO total net charge was estimated using PROPKA (3.0) with spinach RuBisCO structure from PDB (entry: 1AUS).

Figure 5 Correlation functions (**A**) and size distribution (**B**) obtained by dynamic light scattering according to pH. Correlation functions are representative of at least six observations. Size distribution curves are the mean of six observations. For readability reasons, the standard deviations are not shown; they suggest a significant difference between the mean curves.


M
1
2
260
160
110
80
60
50
40
30
20
15
10
3.5
LC-RuBisCO dimer
LC-RuBisCO
SC-RuBisCO
A
M
1
2
98
62
49
38
28
LC-RuBisCO dimer
LC-RuBisCO
B


A
275 nm
Absorbance (AU)
Wavelengths (nm)
B
258 nm
y = 4.1243x
R² = 0.9987
A 280 nm
Protein concentration (mg.mL-1)
Absorbance (AU)
Wavelengths (nm)


A
0.012
0.008
0.004
0
Absorbance (AU)
1000
100
10
1
Molecular weights (kg.mol-1)
B
0.03
0.02
0.01
0.00
Absorbance (AU)
SC
LC
LC dimer
Contaminant
265
259
276
220
250
280
310
340
370
400
Wavelengths (nm)


A
Solubility (%)
100
80
60
40
20
0
2
4
6
8
10
pH
B
μE (μm.cm.s-1.V-1)
Total net charge
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
800
600
400
200
0
-200
-400
-600
-800
2
4
6
8
10
pH

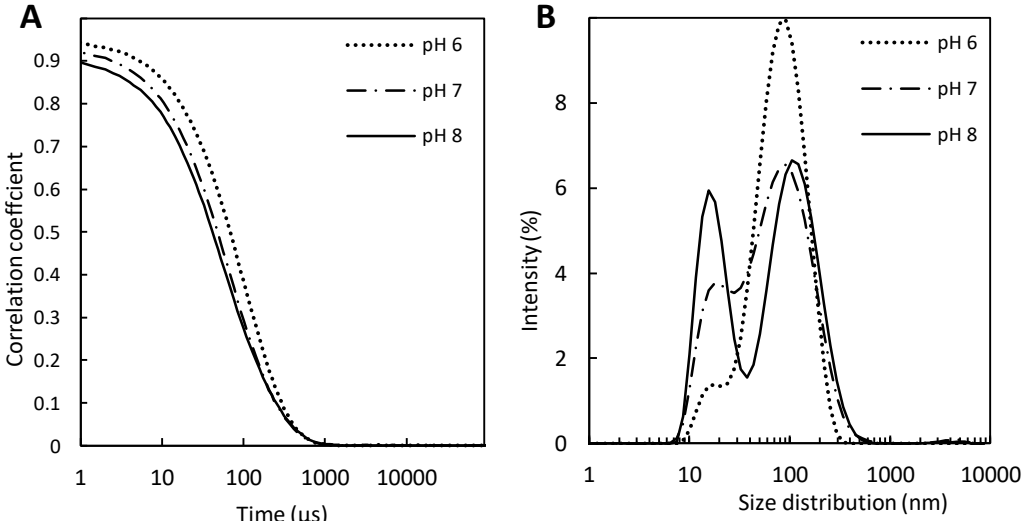

A
B
pH 6
pH 7
pH 8
Correlation coeffcient
Time (µs)
Intensity (%)
Size distribution (nm)
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
2
4
6
8
1
10
100
1000
10000

Table 1 Amino acid composition determined on LPC in $mg.g^{-1}$ of protein as compared to calculated amino acid composition of lettuce RuBisCO (based on lettuce RuBisCO sequence found in UniProt database: *lactuca sativa*; accession numbers of LC and SC: P48706 and Q40250).

Amino acids	**Amino acid residue content** *(mg.g^{-1} total amino acid residues)*	
	Leaf protein concentrate	**Lettuce RuBisCO**
Ala	55.6	49.7
Arg	56.7	76.7
Asp	82.3	83.6
Cys	45.0	18.4
Glu	116.6	116.9
Gly	49.9	46.7
His	33.7	36.7
Ile	45.4	53.9
Leu	91.1	84.2
Lys	62.9	64.8
Met	29.3	25.4
Phe	56.7	61.3
Pro	63.9	46.2
Ser	24.4	29.8
Thr	55.5	54.2
Trp	21.4	33.2
Tyr	48.4	60.7
Val	61.4	57.5