



Compared digestibility of plant protein isolates by using the INFOGEST digestion protocol

Marta Santos-Hernández^a, Fabio Alfieri^b, Veronica Gallo^b, Beatriz Miralles^a, Paolo Masi^b, Annalisa Romano^b, Pasquale Ferranti^b, Isidra Recio^{a,*}

^a Instituto de Investigación en Ciencias de la Alimentación, CIAL (CSIC-UAM, CEI UAM+CSIC), Nicolás Cabrera, 9, 28049 Madrid, Spain

^b Department of Agricultural Sciences, Division of Food Science and Technology, University of Naples Federico II, Via Università 100, 80055 Portici, Naples, Italy

ARTICLE INFO

Keywords:

Protein isolate
Garden pea
Grass pea
Soybean
Lentil
Casein
Whey protein
Gastrointestinal digestion
Epitopes

ABSTRACT

The use of ingredients based on plant protein isolates is being promoted due to sustainability and health reasons. However, it is necessary to explore the behaviour of plant protein isolates during gastrointestinal digestion including the profile of released free amino acids and the characterization of resistant domains to gastrointestinal digestion. The aim of the present study was to monitor protein degradation of four legume protein isolates: garden pea, grass pea, soybean and lentil, using the harmonized Infogest *in vitro* digestion protocol. *In vitro* digests were characterized regarding protein, peptide and free amino acid content. Soybean was the protein isolate with the highest percentage of insoluble nitrogen at the end of the digestion (12%), being this fraction rich in hydrophobic amino acids. Free amino acids were mainly released during the intestinal digestion, comprising 21–24% of the total nitrogen content, while the percentage of nitrogen corresponding to peptides ranged from 66 to 76%. Legume globulins were resistant to gastric digestion whereas they were hydrolysed into peptides and amino acids during the intestinal phase. However, the molecular weight (MW) distribution demonstrated that all intestinal digests, except those from soybean, contained peptides with MW > 4 kDa at the end of gastrointestinal digestion. The profile of free amino acids released during digestion supports legume protein isolates as an excellent source of essential amino acids to be used in protein-rich food products. Peptides released during digestion matched with previously reported epitopes from the same plant species or others, explaining the ability to induce allergic reactions and cross-linked reactivity.

1. Introduction

Gastrointestinal digestion is a physical and chemical process where food is broken down by the action of different enzymes to release nutrients to be absorbed and used by the organism. In the case of digestible proteins, these are hydrolysed into a large variety of peptides and amino acids. It is critical to know the way food proteins are digested since this influences their nutritional value and, furthermore, protein degradation products may be physiologically relevant. *In vitro* digestion protocols are useful to predict the digestibility and to assess the structural changes of proteins under simulated physiological conditions (Bohn et al., 2018). They offer many advantages in comparison to *in vivo* models. In addition to the easy sampling and the high reproducibility under standardised conditions, *in vitro* models are rapid, less expensive than *in vivo* trials and have no ethical implications. A harmonised *in vitro* digestion protocol based on human physiological conditions was proposed by scientists

from more than 35 countries within the COST Action INFOGEST (Brodtkorb et al., 2019). The biological relevance of this protocol has been demonstrated for milk proteins by comparison with porcine and human digests (Egger et al., 2017; Sanchón et al., 2018).

Animal proteins-derived digestion products have been studied following this protocol due to the potential health implications of dietary peptides (El et al., 2015; Tagliazucchi, Martini, Shamsia, Helal, & Conte, 2018). Some plant foods are excellent sources of high-quality proteins for human and animal nutrition, such as soybean, peas or lentils, but there is less information regarding their digestion and the identification of their resistant domains to gastrointestinal digestion. It is generally accepted that ileal digestibility of plant proteins is usually lower than that of animal proteins, although this can be improved once plant cell wall constituents are removed (Tomé, 2013). Ultrafiltration or isoelectric protein precipitation from legumes has been explored in order to produce high purity protein fractions with improved

* Corresponding author at: Nicolás Cabrera, 9, 28049 Madrid, Spain.

E-mail address: i.recio@csic.es (I. Recio).

<https://doi.org/10.1016/j.foodres.2020.109708>

Received 5 June 2020; Received in revised form 21 August 2020; Accepted 6 September 2020

Available online 17 September 2020

0963-9969/© 2020 Elsevier Ltd. All rights reserved.

digestibility (Barbana & Boye, 2013). Porzucek, Larssonraznikiewicz, and Klepacka (1991) reported that pea protein isolate reached 82% digestibility, while whole pea seeds showed 64% (Bishnoi & Khetarpaul, 1994). Similarly, Aryee and Boye (2016) demonstrated the digestibility of raw lentil flour was 77% and it increased up to 85% in the case of lentil protein isolate. Nguyen, Bhandari, Cichero, and Prakash (2015) described a digestibility of 76% in soy protein isolate, slightly above the reported value of 72% for whole soybean seed (Han, Swanson, & Baik, 2007). It remains to be studied if this better digestibility of legume protein isolates may have health implications, for instance, by reducing immunogenicity.

There is an increasing interest in exploring legume digestibility and derived products to promote their consumption as alternative to animal protein sources due to sustainability and health reasons. The application of legume protein isolates in the food industry is a growing area, being used as meat replacers and analogues, in plant-based milk substitutes and in gluten-free bakery products. For instance, soy and pea protein isolates are blended with wheat gluten to obtain materials with fibrous morphology and similar strength than chicken meat (Schreuders et al., 2019). Nevertheless, the knowledge about the behavior of plant protein isolates during gastrointestinal digestion is still limited, as well as the peptidomic characterization of the resistant peptides to gastrointestinal digestion. Four plant-based food were selected: soybean (*Glycine max*), commonly used as ingredient in different food as soy milk or tofu (Reynaud, Lopez, Riaublanc, Souchon, & Dupont, 2020); garden pea (*Pisum sativum L.*), an important plant source of proteins and a potential alternative to soybean in Europe; lentils (*Lens culinaris*), that is considered the fourth most important crop in the world (Longobardi et al., 2017); and grass pea (*Lathyrus sativus*), a limited source in the human diet due to the presence of the neurotoxic non-protein amino acid β -N-oxalyl-L- α , β -diaminopropionic acid (β -ODAP) (Yan et al., 2006; Romano et al., 2019a) although it is rich in proteins and easily digested in both *in vitro* adult and elderly models (Romano et al., 2019b).

In the present study, the main objective was to monitor protein degradation and to identify the digestion products of four plant protein sources: soybean, garden pea, grass pea and lentil, using the harmonized INFOGEST *in vitro* protocol. For this purpose, after protein extraction, simulated gastrointestinal digests were prepared and protein degradation during digestion was evaluated by SDS-PAGE, aided by MALDI-TOF/TOF for protein identification. Peptides resistant to gastrointestinal digestion were characterized by HPLC-tandem mass spectrometry (HPLC-MS/MS). These peptides were matched with reported linear food epitopes in order to detect the presence of digestion-resistant peptides with potential immunogenicity. Moreover, free and total amino acid contents were also analysed to determine the distribution of the nitrogen fractions after gastric and gastrointestinal digestion of these plant protein isolates. Casein and whey protein digests were used as reference to compare the distribution of the nitrogen fraction and free amino acid profile.

2. Materials and methods

2.1. Samples

Grass pea (*Lathyrus Sativus*) seeds were obtained from Granoro (Italy). Garden pea seeds (*Pisum Sativum* var. *Gloton*) were provided by Plant Genetic Resources Center, CRF (INIA). Red lentil flour (*Lens Culinaris*) was commercial and derived from The Lentil of Altamura P. G. I. (Puglia, Italy). Yellow soybean seeds were purchased in a local market in Madrid (Spain). Casein and whey protein gastric and intestinal digests were used as standards in the distribution of nitrogen content and in the profile of free amino acids.

2.2. Protein extraction

Protein isoelectric precipitation was performed by following Barac

et al. (2010) with minor modifications. Briefly, garden pea, grass pea and soybean seeds were powdered in a multi-functional grinder with a particle size range between 200 and 800 μ m. Lentil flour was purchased in a local Italian market. For protein extraction, 50 g of the flour were suspended in 500 mL of milli-Q water and stirred for 1 h at room temperature after adjusting pH to 9 with 10 M NaOH. The mixture was centrifuged at $4,000 \times g$ for 15 min in order to remove insoluble components as fiber. The supernatant was separated and the insoluble part was re-extracted stirring 30 min at pH 9 followed by centrifugation under the same conditions. Supernatants of two extractions were merged and precipitated at pH 4.5 with 6 M HCl. The mixture was stored at 4 °C for 2 h and centrifuged at $4,000 \times g$ for 15 min. The precipitate was re-dissolved at pH 9 for 15 min and precipitated again at pH 4.5. After storage at 4 °C for 2 h, it was centrifuged and the insoluble fraction, which represents the protein isolate, was collected and re-dissolved at pH 7. Samples were freeze-dried and kept at -20 °C until analysis.

After the protein extraction procedure, the activity of the protease inhibitors was measured in the freeze-dried products. Protein isolates from garden pea, grass pea, soybean and lentil were dissolved to 40 mg/mL in 50 mM HCl for 2 h at 4 °C, and were centrifuged at $15,000 \times g$ for 5 min. Then, the supernatant was collected to measure the inhibition of trypsin activity following Domoney and Welham (2008), and chymotrypsin, following Clemente, MacKenzie, Jeenes, and Domoney (2004).

Protein content was determined by elemental analysis using nitrogen to protein conversion factor 5.5 for soybean and 5.4 for the others (Mariotti, Tomé, and Mirand, 2008). The *in vitro* digestion was performed on a protein basis.

2.3. *In vitro* simulated gastrointestinal digestion

Garden pea, grass pea, soybean and lentil protein isolate powders were digested according to the Infogest *in vitro* digestion protocol (Brodkorb et al., 2019) but the concentration of bile salts in the intestinal phase was reduced to use the resultant digests in future cell culture studies. Protein isolates were dispersed in milli-Q water at 60 mg of protein per mL. The mix was diluted with simulated salivary fluid (1:1; v:v) containing human salivary α -amylase (final amylase concentration 75 U/mL of oral digest, EC 3.2.1.1. Sigma-Aldrich, St Louis, MO, USA) and was incubated for 2 min at pH 7. The oral bolus was then diluted 1:1 (v:v) with simulated gastric fluid containing pepsin from porcine gastric mucosa (2000 U/mL of digest, EC 3.4.23.1, Sigma-Aldrich). The reaction was stopped after 120 min at 37 °C increasing the pH at 7 with 1 M NaOH. The intestinal phase was carried out by mixing the end point of the gastric phase with the same volume of simulated intestinal fluid containing pancreatin from porcine pancreas (100 U trypsin activity/mL of final mixture, Sigma-Aldrich) and porcine bile extract (2.5 mM in the final mixture, Sigma-Aldrich). The intestinal phase was stopped by heating at 85 °C for 15 min. Gastric and intestinal digests were centrifuged at $5,000 \times g$ for 20 min to separate soluble and insoluble material prior to freeze-drying. Samples were stored at -20 °C until analysis.

2.4. Distribution of the nitrogen fraction

Soluble and insoluble material of gastric and intestinal digests were weighted after freeze-drying. The protein content in each fraction was measured by elemental analysis and by determination of the total amino acid content. To measure total amino acids, digests were hydrolyzed with 6 N HCl at 110 °C for 24 h. The analysis was performed in a Biochrom 30 series Amino Acid Analyzer (Biochrom Ltd, Cambridge, UK) with a cation-exchange column followed by ninhydrin derivatization. Free amino acids were also determined in the soluble phase, as previously described (Santos-Hernández, Tomé, Gaudichon, & Recio, 2018). The difference between total and free amino acid content in the soluble phase was ascribed to soluble proteins and peptides.

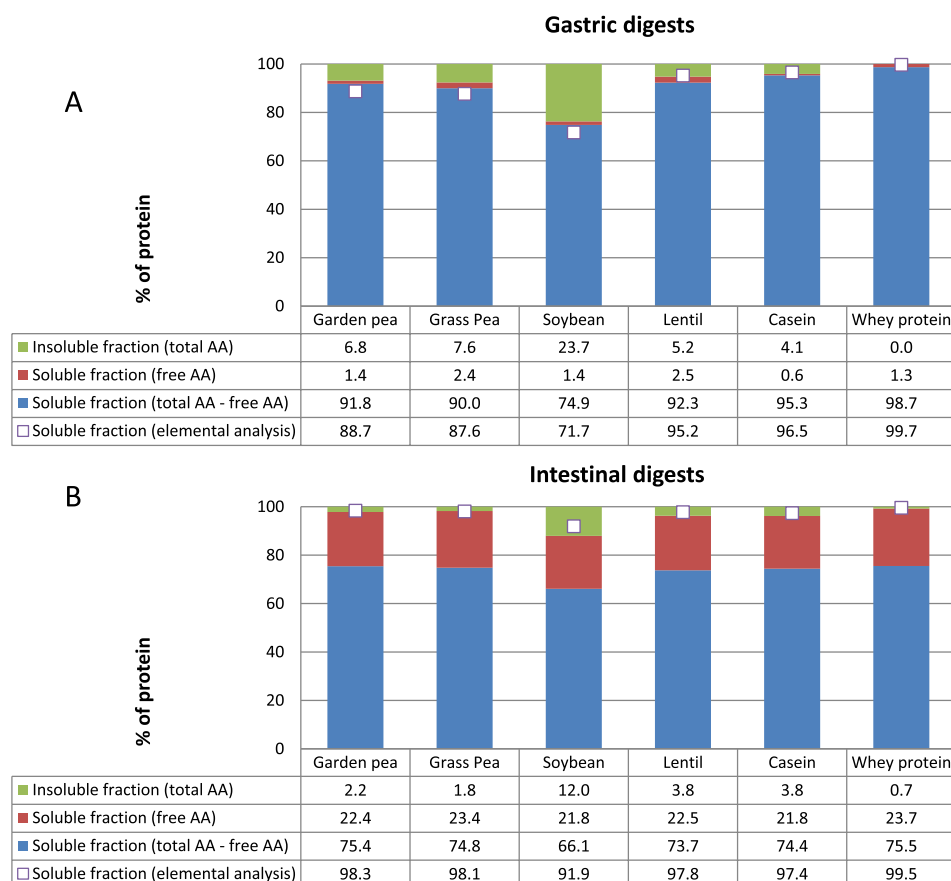


Fig. 1. Distribution of protein content between soluble and insoluble fraction after gastric (A) and intestinal (B) digestion of garden pea, grass pea, soybean, lentil, casein and whey protein. Digests were centrifuged at $5,000 \times g$ for 20 min. Supernatant and precipitate were freeze-dried and weighted; the protein content in each fraction was measured by elemental and by total amino acids analysis. The difference between total and free amino acids in the soluble fraction is ascribed to proteins and peptides.

2.5. SDS-PAGE

SDS-PAGE was performed on gastric and intestinal digests before the centrifugation step. Gastric and intestinal digests were diluted in sample buffer (1:3 and 1:1; v:v, respectively), as described by Sanchón et al. (2018). After heating at 95 °C for 5 min, the mix was loaded on the 12% Bis-Tris polyacrylamide gels (Criterion XT, Bio-Rad, Hercules, CA, USA) placed into the Criterion cell (Bio-Rad) with XT MES running buffer (Bio-Rad). Electrophoretic separation was run at 100 V for 5 min and at 150 V. The gels were stained with Coomassie Blue (Instant Blue, Expedon, Swavesey, UK) and images were taken with a Molecular Imager VersaDoc™ MP 5,000 system (Bio-Rad, Hercules, CA, USA).

2.6. Identification of electrophoretic bands by MALDI-TOF/TOF.

Coomassie Blue-stained bands were manually excised and decolorized with 100 µL of 25 mM NH_4HCO_3 and acetonitrile (1:1; v:v), as previously described (Santos-Hernández, Amigo, & Recio, 2020). Afterwards, the gel bands were reduced with 10 mM 1,4-dithiothreitol (DTT, Sigma-Aldrich) in 25 mM NH_4HCO_3 for 1 h at 56 °C prior to 30 min incubation with 55 mM iodoacetamide at room temperature. Gel pieces were washed (x2) with 25 mM NH_4HCO_3 : acetonitrile (1:1; v:v) and then dehydrated in 100% acetonitrile. The in-gel digestion was carried out at 37 °C, overnight, with 15 µL of trypsin (V5117, Promega Biotech Ibérica S. L., Spain) at a final concentration of 12.5 ng/µL in 25 mM NH_4HCO_3 . After incubation, soluble peptides were recovered by centrifugation at $7,000 \times g$ for 1 min. Subsequently, the gel pieces were washed twice with acetonitrile: 10% formic acid (1:1; v:v) and centrifuged; the supernatants were merged and dried using a Speed Vac prior to reconstitution in acetonitrile 33% containing trifluoroacetic acid 0.1%. Reconstituted samples were spotted into a MALDI target plate with α -cyano-4-hydroxycinnamic acid (HCCA) matrix. The mass spectrometer

used was an Autoflex Speed™ (Bruker Daltonics, Bremen, Germany). Ions were detected in positive reflection mode at a mass range of m/z 600–3,500, and the laser pulses accumulated were 1,000 on average. Protein calibration Standard I, Bruker Daltonics was employed for external calibration of spectra. Monoisotopic peaks selected using the FlexAnalysis 3.3 (Bruker Daltonics) software were fragmented using the lift method for MS/MS analysis. For peptide identification the MASCOT v2.4 software (Matrix Science) Server 2.1 and Biotoools version 2.1 (Bruker Daltonics) were used. The parameters used were as follows: 0.1% peptide mass tolerance, 0.5 Da MS/MS tolerance, peptide charge of + 1, trypsin as digesting enzyme with two missed cleavages allowed; carbamidomethylation of cysteine; oxidation of methionine. A home-made database of plant proteins was used. Soybean database included a total of 32 protein sequences from *Glycine max*; pea database, 26 proteins belonging to *Pisum sativum*; lentil database, 71 proteins from eight different species, among others, *Lens culinaris*, *Pisum sativum*, *Lathyrus sativus*, *Vicia faba*, and *Glycine max*; and Grass pea database contained 114 proteins from *Lathyrus sativum* and *Pisum sativum*.

2.7. Molecular weight distribution of peptides by MALDI-TOF

Gastric and intestinal digests were diluted at 0.1 mg of protein/mL in 33% acetonitrile with 0.1% trifluoroacetic acid. Analysis were carried out in a MALDI target plate with HCCA as the matrix. Ions were detected in positive linear mode at a mass range of m/z 500–20,000 using the Autoflex Speed™ and were collected from the sum of 1,000 on average laser shots. Two protein standards (Protein Calibration Standard I and Protein Calibration Standard II, Bruker Daltonics) were employed for external calibration of spectra. The monoisotopic peak lists were generated by the FlexAnalysis 3.3 software and were represented in a molecular weight (MW) distribution range.

Intestinal digests

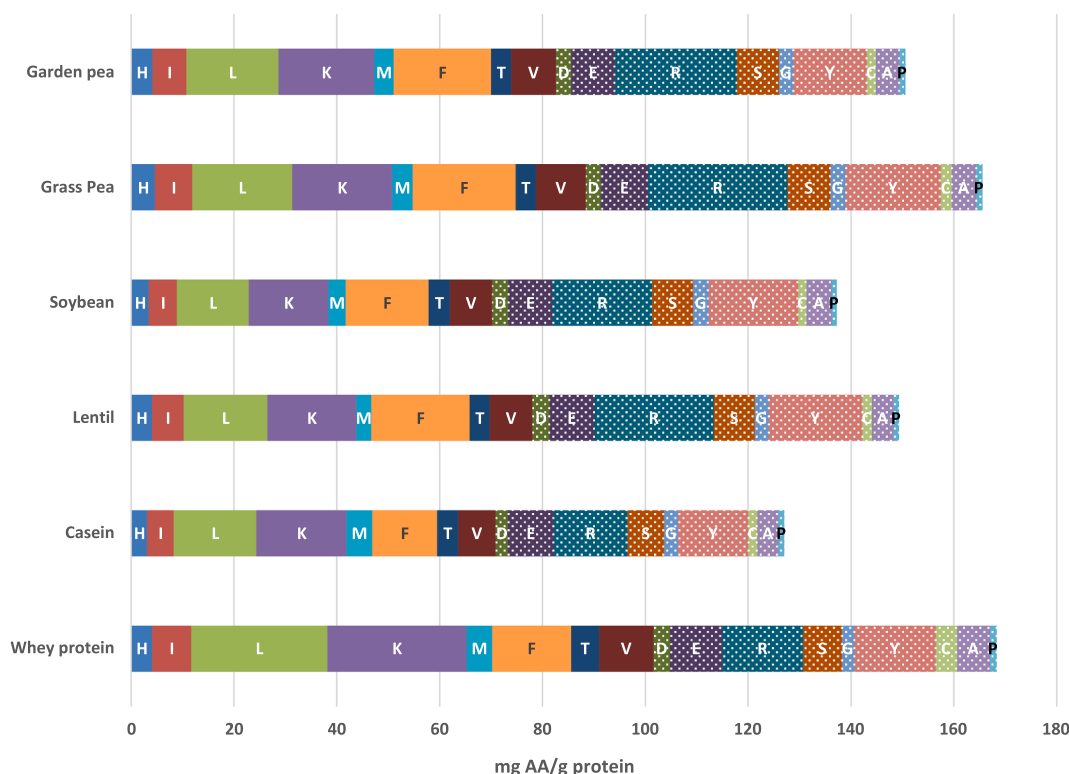


Fig. 2. Free-amino acid profile found at the end of intestinal digestion of garden pea, grass pea, soybean and lentil, casein and whey protein isolates. The content of each amino acid is expressed in mg/g of protein in the freeze-dried digest. Solid colors correspond to essential amino acids and frame-colored bars correspond to non-essential amino acids.

2.8. Peptide identification by HPLC-tandem mass spectrometry (HPLC-MS/MS) analysis

Intestinal digests were analysed by HPLC-MS/MS using an Agilent 1100 HPLC system (Agilent Technologies, Waldbronn, Germany), connected to an Esquire 3000 linear ion trap mass spectrometer (Bruker Daltonics GmbH, Bremen, Germany) supplied with an electrospray ionization source, as described by Santos-Hernández et al. (2018). The column was a reversed phase Mediterranea Sea₁₈ column (150 × 2.1 mm, Teknokroma, Barcelona, Spain). Peptides were eluted with a linear from 0 to 45% of solvent B (acetonitrile : formic acid, 100 : 0.1, v : v) in solvent A (water : formic acid, 100 : 0.1, v : v) in 150 min. The injection volume was 50 µL and the flow was set at 0.2 mL/min. Nitrogen was used as the nebulizing and drying gas and the ion source parameters were: nebulizer pressure, 60 psi; drying gas, 8 L/min and temperature, 350 °C. The capillary was held at 4 kV and the helium pressure was set at 5×10^{-3} bar. The signal threshold to perform auto MS (n) analyses was 10,000 and the precursor ions were isolated within a range of 4.0 m/z and fragmented with a voltage ramp going from 0.3 to 1.5 V. Spectra were recorded over the mass/charge (m/z) ranges of 100–600, 100–1700, and 100–2000, and target mass was set at 500, 750, and 1200, respectively.

Digests were freeze-dried and reconstituted in 25 mM ammonium bicarbonate at pH 7.0. The samples were reduced with DTT (Sigma-Aldrich) at a final concentration of 70 mM and pH 7, for 1 h at 37 °C, to improve the characterization of disulphide-linked fragments. Finally the samples were diluted 1:1 v : v in solvent A (pH 2.6) to obtain a final concentration of 1 mg of protein per mL, and centrifuged at $11,000 \times g$, 10 min prior to analysis. Results were processed by using Data Analysis (version 4.0 Bruker Daltonics) and Biotoools version 3.2 and MASCOT v2.4 software (Matrix Science). A homemade database of various

legume proteins, was used for peptide sequencing in MASCOT, as detailed above. Peptide mass tolerance was set to 0.1% and 0.5 Da for MS and MS/MS analysis, respectively. Identified peptides were represented by using the web application Peptigram (Manguy et al., 2017).

3. Results and discussion

3.1. Distribution of the nitrogen fraction after protein digestion

The protein content of the freeze-dried garden pea, grass pea, soybean and lentil protein isolate were 77%, 82%, 63%, 75%, respectively. Protein isolates did not show trypsin or chymotrypsin inhibitory activities, and were submitted to the simulated digestion protocol. After gastric and gastrointestinal *in vitro* digestion of garden pea, grass pea, soybean and lentil protein isolates, soluble and insoluble material were separated by centrifugation and freeze-dried. Casein and whey proteins, used as references, were digested under the same conditions. The distribution of the nitrogen fraction after simulated gastric (Fig. 1A) and intestinal digestion (Fig. 1B) was assessed by elemental analysis and total amino acid analysis. Both techniques gave similar values of nitrogen content in each fraction. Free amino acids were also determined in the soluble part, being the difference between total and free amino acids assigned to proteins and peptides in the gastric and intestinal digests. At the end of the gastric phase, all plant protein isolates, except soybean, the soluble part of the digest contained 88–95% of the nitrogen, measured by elemental analysis or by total amino acid analysis (Fig. 1A). For the soybean gastric digest, the soluble nitrogen represented 72% or 74% determined by elemental analysis or amino acid analysis, respectively (Fig. 1A). This lower value for soybean isolate ($P < 0.01$) could be attributed to the lower solubility of the gastric digestion products due to the 22 disulfide bonds of glycinin, the main soybean protein, being less

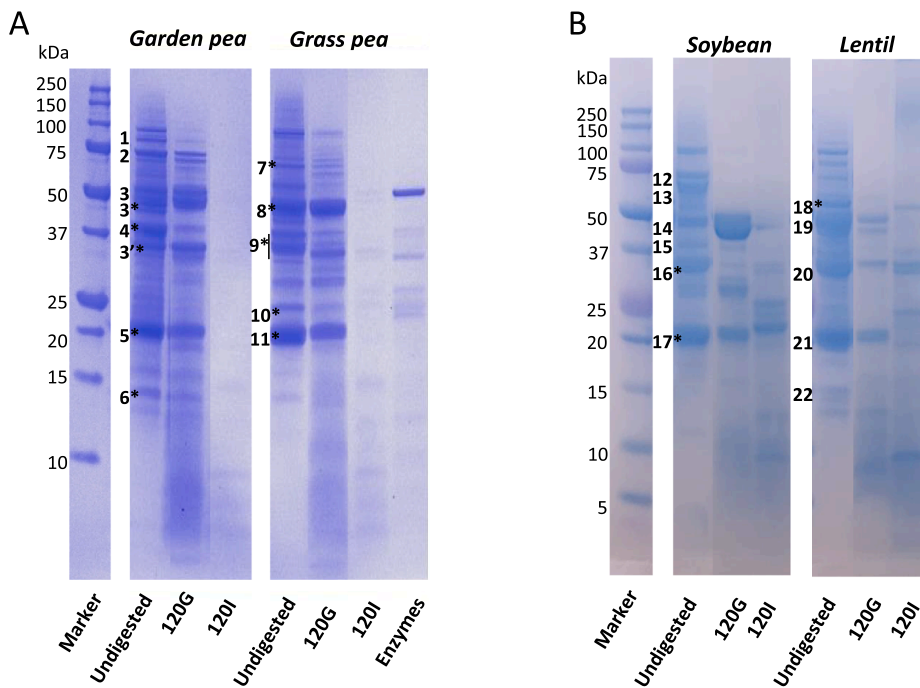


Fig. 3. SDS-PAGE protein profiles of garden pea, grass pea, soybean and lentil protein isolates at different times of *in vitro* gastrointestinal digestion. MW, molecular weight marker; Undigested; undigested protein; 120G, 120 min gastric digestion; 120I, 120 min intestinal digestion. The numbers correspond to an identified band. 1. Lipoygenase; 2. Convicilin; 3 and 3*, Vicilin; 4. Legumin A α -chain; 5. Legumin B β -chain; 6. Vicilin fragment; 7. Convicilin *Lathyrus sativus*; 8. Vicilin; 9. Legumin; 10. Albumin-2 *Lathyrus*; 11. Legumin; 12. β -conglycinin α' subunit; 13. β -conglycinin α subunit; 14. β -conglycinin β subunit; 15. glycinin acid polypeptide 16. glycinin acid polypeptide; 17. glycinin basic polypeptide; 18. Convicilin; 19. Vicilin; 20. 11S globulin acidic subunit; 21. 11S globulin basic subunit; 22. Vicilin 14 kDa component. The identity of the electrophoretic bands with * was confirmed by MALDI-TOF/TOF after hydrolysis with trypsin.

susceptible to proteolysis (Faris, Wang, & Wang, 2008). As comparison, 96% and 100% of the nitrogen content was found in the soluble fraction after gastric digestion of milk casein and whey protein, respectively. Other authors had also reported a lower level of hydrolysis for soy protein than milk proteins under conditions simulating infant gastrointestinal digestion (Nguyen et al., 2015). Among the plant protein isolates, lentil showed a similar percentage of soluble nitrogen than milk proteins after gastric digestion, in line with previous reports regarding the digestibility of lentil protein isolates (Aryee et al., 2016). At the end of the intestinal digestion, the fraction of soluble nitrogen of all plant isolates increased to values ranging between 88% and 98% determined by both elemental analysis and total amino acid analysis (Fig. 1B). The percentage of nitrogen attributed to peptides ranged from 66% to 76% at the end of the simulated digestion. For casein and whey protein digests, the soluble fraction represented 97% and 100% of the total nitrogen content, respectively, and ca. 75% of the total nitrogen corresponded to peptides.

It is accepted that legume seeds present lower digestibility than animal proteins due to the presence of cell walls, protease inhibitors and tannins. However, once these components have been removed in plant protein isolates, their digestibility increases up to be comparable to that achieved for animal protein. Our results on the distribution of the nitrogen fraction confirms this increased digestibility, being the percentage of protein corresponding to the insoluble part after gastrointestinal digestion of garden pea, grass pea and lentil protein isolates comparable (2.2%, 1.8%, 3.5%, respectively) to that found in animal proteins (4% and 0% in casein and whey protein, respectively). The insoluble nitrogen content after gastrointestinal digestion of soy protein isolate was 12% (Fig. 1B). In agreement with our results, Kuipers et al. (2006) showed a decrease in solubility for hydrolysates of soy protein isolates with subtilisin compared with non-hydrolyzed isolate around neutral pH. Glycinin-derived peptides have a much stronger tendency to aggregate than β -conglycinin derived peptides via hydrophobic interactions (Kuipers et al., 2006). Our results on the amino acid analysis of the pellet after intestinal digestion of soybean isolate revealed an enrichment of the amino acids glycine and leucine compared to the soluble part, which confirms this point (results not shown).

As shown in Fig. 1, free amino acids were mainly released during intestinal digestion. Free amino acids accounted for 20–30% of the total

nitrogen content at the end of intestinal phase, whereas free amino acids comprised 1–2% at the end of the gastric phase. These results are in agreement with those reported by Adibi and Mercer (1973) with 30% of ingested nitrogen in the form of free amino acids in human jejunum. Le Gall, Gueguen, Seve, and Quillien (2005) showed that the degree of hydrolysis measured by the o-phthaldialdehyde method of garden pea flour after *in vitro* hydrolysis with pepsin and trypsin was around 25%, depending on the particle size of the flour. Karamac and Rybarczyk (2008) determined the digestibility of lentil protein isolate after 2 h hydrolysis with chymotrypsin by using the pH-stat method and it reached a degree of hydrolysis of 13%. In addition to these methods, the percentage of free amino acids released during intestinal digestion, which ranged between 20 and 30%, could also be used to estimate legume protein digestibility. As an advantage, the present analysis allows to obtain a portrait of the profile of free amino acids (Fig. 2). Leucine, lysine, and phenylalanine were the most abundant essential free amino acids in plant protein intestinal digests (around 14–19 mg/g of protein in the freeze-dried digest, Fig. 2), while whey protein intestinal digest was rich in leucine and lysine, reaching 27 mg/g of protein in the freeze-dried digest. Among the non-essential amino acids, arginine followed by tyrosine were the most abundant in all digests. The percentage of each free amino acid found in the soluble fraction at the end of the gastrointestinal digestion is also shown as percentage of the total content of each corresponding amino acid (Supplementary material Fig. S1). Expressed as percentage, methionine was the most abundant essential free amino acid in plant protein digests followed by phenylalanine, leucine, and lysine. Tyrosine, arginine and cysteine reached higher percentages than other non-essential amino acids. Rubio and Clemente (2009) measured the concentration of amino acids in portal plasma 150 min after feed ingestion of rats fed α -lactalbumin, casein or chickpea protein isolate. In good correspondence with our results (Fig. 2), leucine and lysine were found in lower concentration in rats fed casein or legume protein isolate than those fed α -lactalbumin.

3.2. Molecular weight distribution of the protein digestion products

The MW distribution of proteins and peptides above 20 kDa was assessed by SDS-PAGE (Fig. 3), while peptides under 20 kDa were monitored by MALDI-TOF (Fig. 4). The identification of some

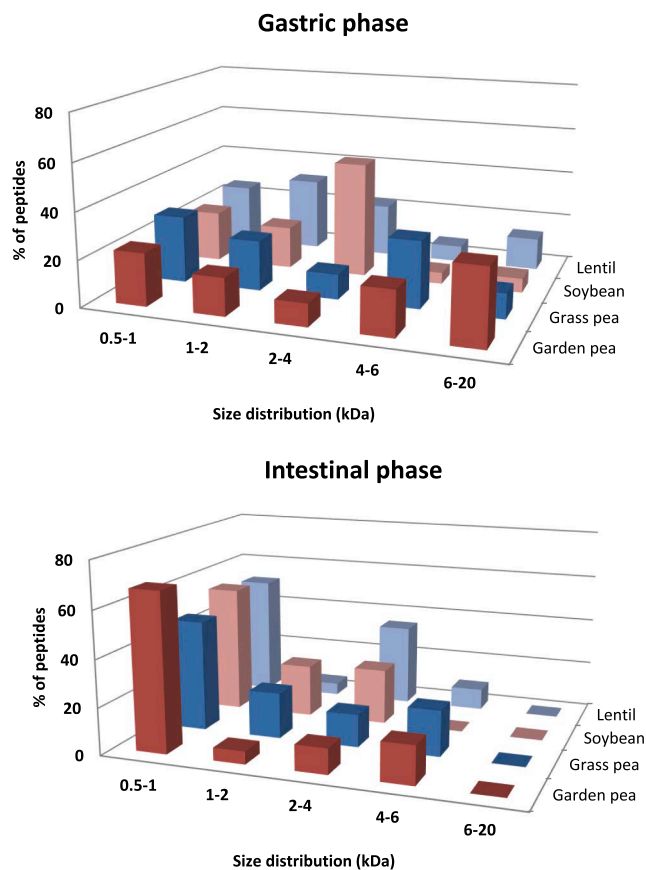


Fig. 4. Molecular weight distribution of garden pea, grass pea, soybean and lentil peptides found at the end of gastric and intestinal digestion using MALDI-TOF.

Table 1

Identification of electrophoretic bands from the plant protein isolates.

	Band n ^a	Protein name	Organism	Accession n ^b	Molecular mass (kDa) ^c	N ^o of matched peptides ^d	Mascot score ^e
Garden Pea	3	Vicilin	<i>Pisum sativum</i>	P13918	52.2	6	55
		Vicilin 47 k	<i>Pisum sativum</i>	Q43626*	47.0	4	49
	3'	Vicilin	<i>Pisum sativum</i>	P13918	52.2	7	120
		Provicilin	<i>Pisum sativum</i>	P02854	46.4	7	100
		Vicilin	<i>Pisum sativum</i>	Q43626*	47.0	8	100
	4	Legumin A (α-chain)	<i>Pisum sativum</i>	P02857	36.4	3	60
	5	Legumin B (β-chain)	<i>Pisum sativum</i>	P14594	19.5	2	50
	6	Vicilin fragment	<i>Pisum sativum</i>	P13918	25.0	6	74
Grass Pea	7	Convicilin	<i>Lathyrus sativus</i>	BOBCK4*	61.5	4	59
	8	Vicilin 47 K	<i>Pisum sativum</i>	D3VND9*	49.5	6	46
	9	Legumin	<i>Pisum sativum</i>	P15838	36.8	7	86
	10	Albumin-2	<i>Lathyrus sativus</i>	D4AEP7	25.6	3	43
	11	Legumin B (β-chain)	<i>Pisum sativum</i>	P15838	19.5	4	142
		Legumin (fragment)	<i>Pisum sativum</i>	O24294*	64.9	4	142
		Legumin A (β-chain)	<i>Pisum sativum</i>	P02857	20.2	3	88
		Legumin K (β-chain)	<i>Pisum sativum</i>	P05693	20.1	3	86
Soybean	16	Glycinin G1 (A1 subunit)	<i>Glycine max</i>	P04776	32.6	6	162
		Glycinin G2 (A2 subunit)	<i>Glycine max</i>	P04405	31.6	7	89
	17	Glycinin G1 (B subunit)	<i>Glycine max</i>	P04776	20.0	9	272
		Glycinin G2 (B subunit)	<i>Glycine max</i>	P04405	19.8	10	233
		Glycinin G3 (B subunit)	<i>Glycine max</i>	P11828	19.9	8	175
Lentils	18	Convicilin	<i>Pisum sativum</i>	P13915	67.0	9	142

^a) See SDS-PAGE in Fig. 2. ^b) Uniprot data base sequence entry. Asterisk denotes unreviewed sequences ^c) Theoretical molecular mass (recorded in UniprotKB databases) ^d) Number of matched peptides in the database search ^e) MASCOT protein score

electrophoretic bands (marked with * in Fig. 3) was confirmed by in-gel digestion with trypsin followed by MALDI-TOF/TOF analysis. These results are shown in Table 1. SDS-PAGE protein profiles of garden pea (Fig. 3A) show that lipoxigenase and convicilin, with molecular masses

higher than 75 kDa (bands 1 and 2, respectively), were partially hydrolyzed at the end of gastric digestion (120G), as well as legumin A (band 4, Table 1), and, therefore, new electrophoretic bands appeared under 15 kDa. Vicilin, a vicilin-derived fragment and legumin B (bands 3, 6 and 5, respectively, identified by MALDI-TOF/TOF, Table 1), were resistant to pepsin but they were hydrolyzed during the intestinal phase (120I). At the end of intestinal phase, the electrophoretic bands between 20 and 50 kDa corresponded to digestive enzymes, and electrophoretic bands under 15 kDa, corresponding to degradation products, became more intense. These results are in agreement with Le Gall et al. (2005) who showed a similar profile of pea flour digested for 3 h with pepsin. These authors identified vicilin fragments after 1.5 h of incubation of pea flour with trypsin. A reduced accessibility of proteins due to the protection of intact cell wall is the major factor that may decrease pea protein digestibility (Carre, Beaufils, & Melcion, 1991). As expected, protein degradation of grass pea was similar to that observed for garden pea (Fig. 3A). Electrophoretic bands 3, 8, 9 and 10 were identified by MALDI-TOF/TOF as vicilin, legumin A2, α-chain of legumin A2, and albumin-2 lathyrus, respectively (Table 1). These results are supported by Lioi et al. (2011), who identified vicilin and legumin fragments after a reducing step with β-mercaptoethanol. Convicilin and legumin fragments (electrophoretic bands 7 and 11) were also identified, as previously described (Rosa, Ferreira, & Teixeira, 2000).

Protein patterns of undigested soybean protein isolate, gastric and intestinal digests by SDS-PAGE are shown in Fig. 3B. Soybean protein isolate showed electrophoretic bands corresponding to different subunits of β-conglycinin and glycinin, (Fig. 3B, Table 1), as reported in previous studies (Ruan, Chen, Kong, & Hua, 2015; Picariello, Amigo-Benavent, del Castillo, & Ferranti, 2013). At the end of the gastric phase, bands 14 and 17 were identified as β-conglycinin (β-subunit) and glycinin (subunit B) showing resistance to the action of pepsin. These results are in accordance with Nguyen et al., 2015, who showed partial hydrolysis of β-conglycinin after 1 h with pepsin. Protein degradation of lentil protein isolate during gastrointestinal digestion is shown in Fig. 3B. Electrophoretic bands 18, 20, and 21 were identified by MALDI-

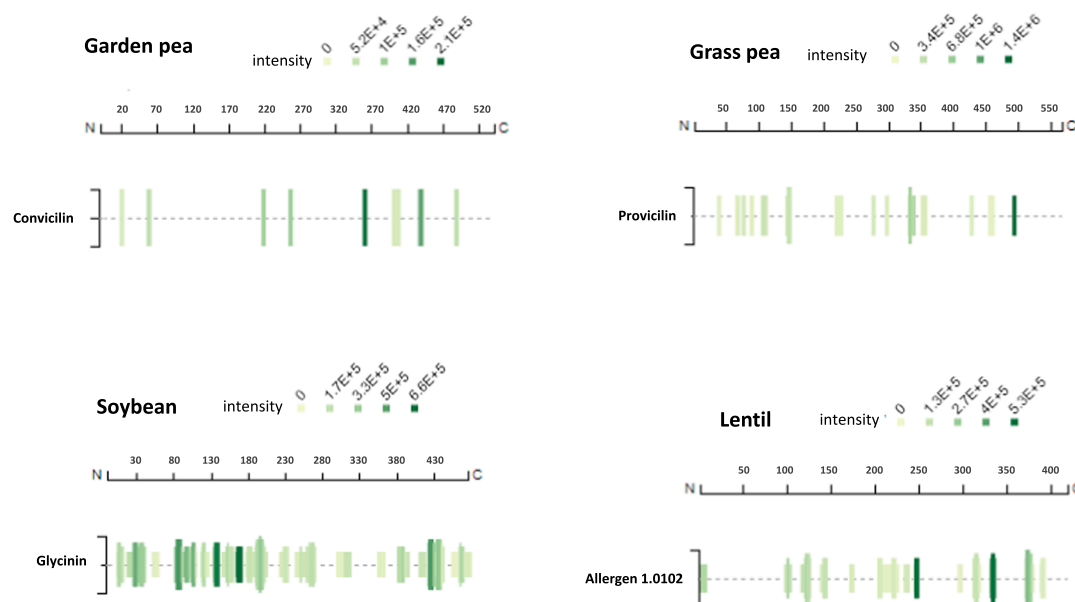


Fig. 5. Peptides from convicilin, provicilin, glycinin and allergen 1.0102 identified in garden pea, grass pea, soybean and lentil intestinal digests, respectively, were represented using the Peptigram Bioware tool. Each vertical bar corresponds to an amino acid identified as part of a peptide sequence. The height of the bar is proportional to the number of the peptides overlapping this position and the colour intensity is proportional to the sum of the intensities of the peptides overlapping a given position. Each line corresponds to an identified peptides at the end of intestinal digestion.

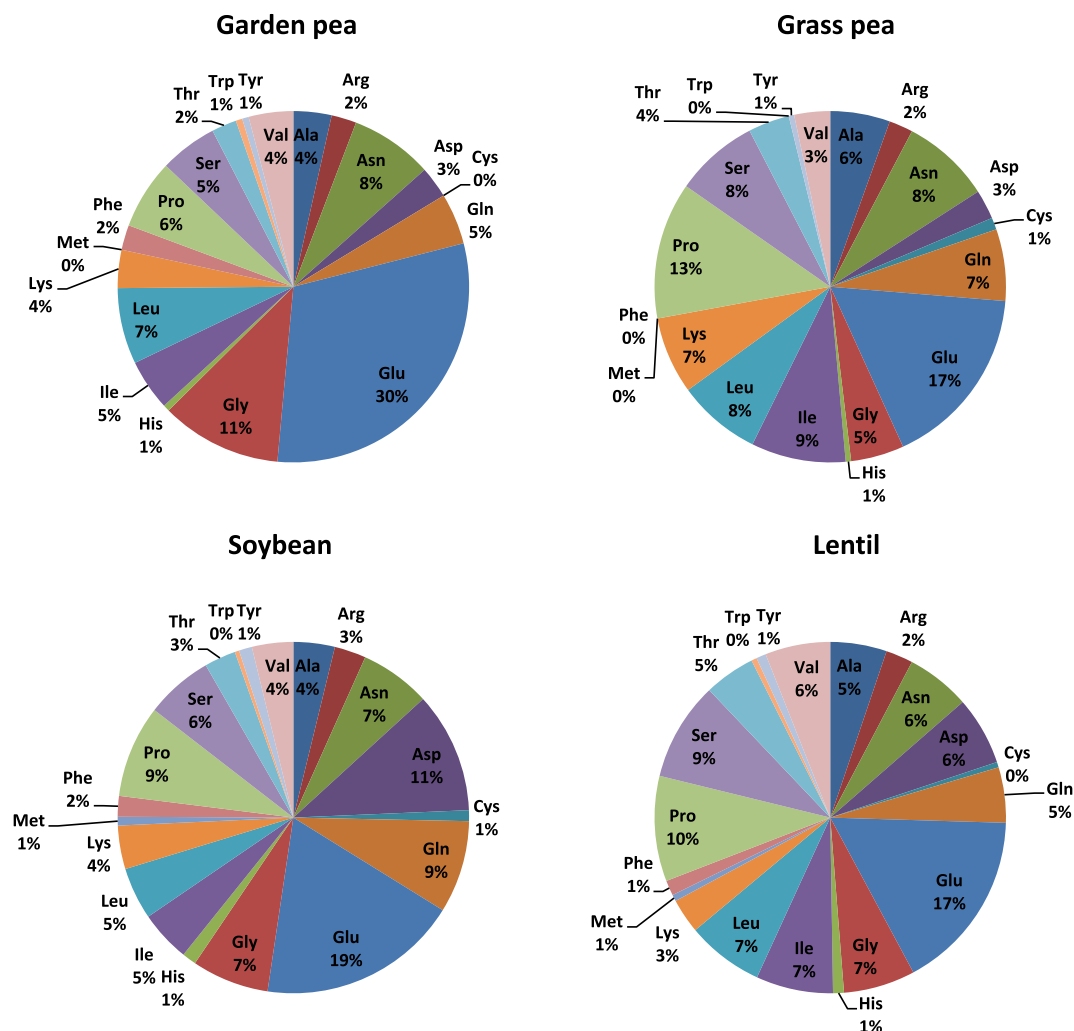


Fig. 6. Amino acid composition of the identified peptide sequences at the end of the *in vitro* gastrointestinal digestion of garden pea, grass pea, soybean and lentil. Repeated sequences are avoided.

Table 2

Identified peptides from soybean glycinin G1 longer than 8 amino acids and reported related epitopes. Common sequences are highlighted in bold. Homologous epitopes to the identified are searched by using the public resource Immune Epitope Database and Analysis Resource (IEDB) selecting linear epitopes with positive assays and 90% identity percentage.

Glycinin G1 - Uniprot: P04776 (GLYG1_SOYBN)						
Mass ^a	Position	Identified sequence	Related epitopes	Epitope ID ^b	Allergen	Organism ^c
1102.5	24–32	EQPQQNECQ	QPPQNECQIQ	913981	Gly m 6	Glycine max
1046.7	45–54	IESEGGLIET	IESEGGLIETWNPNNK RIESEGGLIETWNPNNK	187845 54145	Gly m 6 Ara h 3	Glycine max Arachis hypogaea
969.6	54–61	TWNPNNKP	IETWNPNNKQEFECAG GLIETWNPNNKPFQCA IESEGGLIETWNPNNK	25997 187829 187845	Ara h 3 Gly m 6 Gly m 6	Arachis hypogaea Glycine max Glycine max
1024.7	101–110	GMIYPGCPST	FGMIFPGCPSTYQEP	16009	Gly m 6	Glycine max
923.6	101–109	GMIYPGCPSP	PGCPSTFEPPQPPQQRGQS	809317	Gly m 6	Glycine max
735.4	103–109	IYPGCPSP	IYPGCPSTFEE	913490	Gly m 6	Glycine max
836.5	103–110	IYPGCPST				
873.6	121–128	GQSSRPQD	EPQKQGQSSRPQDRH GQSSRPQD	913164 913298	Gly m 6 Gly m 6	Glycine max Glycine max
913.5	138–146	EGDLIAVPT	GDLIAVPTGVAVWVWYNNE VHRFREGDLIAV RFDEGDIAVPTGVA GDLIAVPTGVAFWLY	538669 538776 53687 99266	Gly m 6 Gly m 6 Ara h 3 Ara h 3	Glycine max Glycine max Arachis hypogaea Arachis hypogaea
1173.8	170–179	SLENQLDQMP	NQLDQMPPRRFYLAGN	45574	Gly m 6	Glycine max
973.6	172–179	ENQLDQMP	ENQLDQ	913154	Gly m 6	Glycine max
1050.7	183–191	YLAGNQEQE	RFYLAGNQEQEFLKY RRFYLAGNQEQEFLR RRFYLAGNQEQEFL	53810 838383 914046	Gly m 6 Ara h 3 Gly m 6	Glycine max Arachis hypogaea Glycine max
997.6	196–204	QQEQGGHQS	QQEQGGHQ	913987	Gly m 6	Glycine max
1182.8	198–208	EQGGHQSQKGG	HQSQKG QQEQGGHQ	913364 913987	Gly m 6 Gly m 6	Glycine max Glycine max
1427.8	207–219	GKHQEEEEENEGGS	HSRRERAGQEEENEG	24844	Ara h 3	Arachis hypogaea
1242.7	209–219	HQEEEEENEGGS	QEEENEGGNIFSGFT	50572	Ara h 3	Arachis hypogaea
1355.8	209–220	HQEEEEENEGGSI	EENEGGNIFSGFTPE GQEEENEGGNIFSGF KHQEQE	99185 99293 913515	Ara h 3 Ara h 3 Gly m 6	Arachis hypogaea Arachis hypogaea Glycine max
1080.7	215–225	NEGGSILSGFT	GGSILSGFTLEFL ILSGFTLEFLHAFSV	20021 187850	Gly m 6 Gly m 6	Glycine max Glycine max
1002.7	281–288	EEEEDEKP	EEEEDEK	913087	Gly m 6	Glycine max
1105.7	282–290	EEEEDEKPPC				
1019.7	329–337	SSPDIYNPQ	KKNIGRNRSPDIYNP IGRNRSPDIYNPQAG NRSPDIYNPQAGSLK	31642 99331 99513	Ara h 3 Ara h 3 Ara h 3	Arachis hypogaea Arachis hypogaea Arachis hypogaea
1183.7	374–383	YNLNANSIHY	MFVPHYTLNANSIHY NANSIHYALNGRALV VPHYTLNANSIHYAL	41548 43262 70332	Gly m 6 Gly m 6 Gly m 6	Glycine max Glycine max Glycine max
1430.8	474–485	IKNNNPFFKFLVP	QARQLKNNNPFFKFFV NNNPFFKFFVPPSEQS PREQARQLKNNNPFFK QLKNNNPFFKFFVPPS	106026 105988 106008 106042	Ara h 3 Ara h 3 Ara h 3 Ara h 3	Arachis hypogaea Arachis hypogaea Arachis hypogaea Arachis hypogaea

^a Experimental mass.

^b Epitope ID in the Immune Epitope Database and Analysis Resource (www.iedb.org).

^c Common names: *Glycine* max, soybean; *Arachis hypogaea*, peanut.

Boye, 2013).

MW distribution of peptides between 0.5 and 20 kDa was performed by MALDI-TOF. Peptide profiles of garden pea, grass pea, soybean and lentil are represented in Fig. 4, which shows the percentage of peptides within a given MW range, at the end of gastric and intestinal digestion. As foreseen, the percentage of peptides within the low MW range increased with the digestion time in all samples. For instance, after gastric digestion 23% of garden pea-derived peptides were comprised between 0.5 and 1 kDa, and this percentage increased up to 68% at the end of intestinal phase. Similarly, peptides between 6 and 20 kDa derived from garden pea represented 32% at the end of gastric phase, but peptides within this range could not be detected at the end of intestinal phase (Fig. 4). A similar profile was obtained with grass pea, soy bean and lentil intestinal digests. Remarkably, at the end of the gastric phase, soybean digest showed an important percentage of peptides (49%) within the MW range from 2 to 4 kDa, and garden pea was the substrate with a higher content of peptides, 32%, in the highest range (6–20 kDa). At the end of the intestinal phase, all substrates showed more than 47% of the peptides below 1 kDa. Pea and grass pea digests had 16–19% of peptides included in the range 4–6 kDa while lentil and

soybean showed similar percentages of peptides (24–35%) between 2 and 4 kDa.

3.3. Peptidomic characterization of intestinal *in vitro* digests

Garden pea, grass pea, soybean and lentil were analysed by HPLC-MS/MS after 120 min of gastrointestinal digestion and 25, 25, 169 and 100 different peptides were identified, respectively. Peptides from main proteins were represented using the Peptigram web application (Fig. 5). In this peptide profile tool, each vertical bar corresponds to an amino acid identified as a part of a peptide sequence; the height of this bar is proportional to the count of peptides overlapping this position while the colour intensity is proportional to the sum of the intensities of the peptides overlapping a given position. Fig. 5 shows the peptide profiles obtained for convicilin from garden pea, provicilin from grass pea, glycinin from soybean and allergen Len c 1.0102 from lentil at the end of intestinal phase. Under our mass spectrometric conditions, peptides with MW between 0.5 and 3 kDa are preferably detected, although a few peptides with longer sizes could be identified depending on their ionization capacity. The blank regions observed in the protein profile,

Table 3

Identified peptides from soybean β -conglycinin α -subunit 1 longer than 8 amino acids and reported related epitopes. Common sequences are highlighted in red. Homologous epitopes to the identified are searched by using the public resource Immune Epitope Database and Analysis Resource (IEDB) selecting linear epitopes with positive assays and 90% identity percentage.

β -conglycinin α subunit 1 - UniProt: PODO16 (GLCA1_SOYBN)						
Mass ^a	Position	Identified sequence	Related epitopes	Epitope ID ^b	Allergen ^c	Organism ^d
1800.3	63–77	VEKEECEGEIPRR	EECEGEIPRRPQHPEREPQQPGEKEE	181299	Gly m 5	Glycine max
1643.1	63–76	VEKEECEGEIPRP	EIPRRP	913122	I1NGH2	Glycine max
			EKEECEEG	913128	Gly m 5	Glycine max
1413.9	86–97	EPQQPGEKEEDE	EECEGEIPRRPQHPEREPQQPGEKEE	181299	Gly m 5	Glycine max
931.6	90–97	PGEKEEDE	HPEREPQQPG	913362	I1NGH2	Glycine max
			QQPGEKE	913991	I1NGH2	Glycine max
1100.7	119–126	QREEQEWPR	HEQREEQEWPRKEEKRGEKGSEEEDE	181334	Gly m 5	Glycine max
			RPIPFRRPQRQEEHEQREEQEWPR	181447	Gly m 5	Glycine max
			EHEQREEQE	913110	Gly m 5	Glycine max
1274.8	128–138	KEEKRGEKGSE	HEQREEQEWPRKEEKRGEKGSEEEDE	181334	Gly m 5	Glycine max
			KRGEKGSEEE	913565	CG-alpha-2	Glycine max
1281.7	136–146	GSEEEDEDEDE	SEEEDEDE	914069	Gly m 5	Glycine max
979.4	137–144	SEEEDEDE	AVRKEQQQRGRSEEEDEDE	190716	Ara h 1	Arachis hypogaea
1007.5	143–150	DEDEEQDE	QFPFRRPPHQKEERKQEEDEDEEQQR	181432	Gly m 5	Glycine max
			EDEEQDERQFP	913068	I1NGH2	Glycine max
			EEEDQDEDEEQK	913086	Gly m 5	Glycine max
1020.6	166–173	KQEEDEDE	QFPFRRPPHQKEERKQEEDEDEEQQR	181432	Gly m 5	Glycine max
			EDEDE	913084	I1NGH2	Glycine max
964.7	246–254	LLPNHADAD	LQNLDRYRILEFNSKPTLLPNHAD	181386	Gly m 5	Glycine max
877.6	244–251	TLLPNHA				
1133.7	267–276	SLVNDDRDS	TAILSLVNDDRDSYRLQSGDALRVPSGT	181474	Gly m 5	Glycine max
933.6	269–276	VNDDRDS	ILTLVNDDRDSYNLH	227783	Gly m 5	Glycine max
			LSGRAILTLVNDDRDR	227802	Gly m 5	Glycine max
1176.7	294–303	YVVPDNNEN	YVVPDNNENLRLITLAIPVNKPGRFES	181524	Gly m 5	Glycine max
1289.9	294–304	YVVPDNNENL				
1013.7	295–303	VVNPDDNNEN				
1562.0	303–316	NLRLITLAIPVNKP	YVVPDNNENLRLITLAIPVNKPGRFES	181524	Gly m 5	Glycine max
			AIPVNKPGRYDDFFLS	227744	Gly m 5	Glycine max
			KIKLAIPVNKPGRYD	227790	Gly m 5	Glycine max
1096.7	323–332	FLSSTEAQQS	FFLSSTEAQQSYLQGFSRNILE	181316	Gly m 5	Glycine max
1260.8	323–333	FLSSTEAQQSY	FFLSSTEAQ	913196	CG-alpha-2	Glycine max
			FLSSTEAQQSYLQGFSRNIL	913213	I1NGH2	Glycine max
			FLSSTQAQQSYLQGFS	913214	Gly m 5	Glycine max
1274.8	359–369	SREEGQQQGEQ	KVLFSREEGQQQ	913578	CG-alpha-2	Glycine max
1031.7	361–369	EEGQQQGEQ	NKVLFSREEGQQQG	913775		
			VLFSSREEGQQQGEQR	914238	I1NGH2	Glycine max
1150.8	397–406	KTISSEDKPF	EQIRALSKRAKSSSRKTISSEDKPFN	181307	Gly m 5	Glycine max
1002.7	397–405	KTISSEDKP	SEDKPFNLRSDPIYSNKLKGFKEITPEKN	181455	Gly m 5	Glycine max
875.6	398–405	TISSEDKP				
1022.8	398–406	TISSEDKPF				
1201.8	423–432	FEITPEKNPQ	DPIYSNKLKGFKEITPEKNPQLRDL	181292	Gly m 5	Glycine max
1054.8	424–432	EITPEKNPQ	SEDKPFNLRSDPIYSNKLKGFKEITPEKN	181455	Gly m 5	Glycine max
			GKFFEITPEKNPQLRD	227769	Gly m 5	Glycine max
			ITPEKNPQLRDLIFL	227787	Gly m 5	Glycine max
			YSNNFGKFFEITPEKN	227865	Gly m 5	Glycine max
1581.6	459–473	AIVILVINEGDANIE	FNSKAIVILVINEGDA	227765	Gly m 5	Glycine max
			VILVINE	914230	Gly m 5	Glycine max
1398.1	463–475	LVINEGDANIELV	FNSKAIVILVINEGDA	227765	Gly m 5	Glycine max
1298.9	463–474	LVINEGDANIEL	DANIELVGLKEQQQE	912980	CG-alpha-2	Glycine max
1185.9	464–474	VINEGDANIEL	LKEQQQE	913621	I1NGH2	Glycine max
1086.8	465–474	INEGDANIEL				
1072.8	466–475	NEGDANIELV				
973.6	466–474	NEGDANIEL				
1515.1	467–480	EGDANIELVGLKEQ				
1123.6	485–503	KYRAELSEQ	EVRYKRAELSE	913187	I1NGH2	Glycine max
1059.8	497–505	RAELSEQDI				
1182.8	575–584	ESYFVDAQPK	QRESYFVDAQPKKKEEGNKGKGPLSSILR	181434	Gly m 5	Glycine max
			DVERLLKKQRESYFVD	227756	Gly m 5	Glycine max
			LKKQRESYFVDAQPPQ	227797	Gly m 5	Glycine max
			NQRESYFVDAQPKK	913847	I1NGH2	Glycine max

^a Experimental mass.

^b Epitope ID in the Immune Epitope Database and Analysis Resource (www.iedb.org).

^c When the protein is not characterized the Uniprot number is given, e.g I1NGH2. CG-alpha-2, β -conglycinin α -subunit 2.

^d Common names: *Glycine* max, soybean; *Arachis hypogaea*, peanut.

where no peptides were identified, may correspond to peptide fragments too long to be solubilized or ionized under the analysis conditions, or may correspond to short peptides, free amino acids or peptides with low ionization capacity. Fig. 6 shows the amino acidic composition of the

resistant peptides at the end of intestinal digestion identified by HPLC-MS/MS. Garden pea resistant domains contained mainly glutamic acid and glycine (30% and 11%, respectively). However, grass pea peptides were rich in glutamic acid and proline (17% and 13%, respectively),

similar as lentil peptides (17% of glutamic acid and 10% of proline). Soybean resistant peptides were mainly composed by negatively charged residues, glutamic and aspartic acid (19% and 11%, respectively). These results are consistent with the lower percentage found for these two as free amino acids (Figure S1). Despite the amount of negatively charged amino acids, about 11% of the total identified peptides resistant to soybean gastrointestinal digestion had a theoretical pI > 10 and positive net charge. Interestingly, some identified peptides had been previously described because of their biological activities. For instance, lentil convicilin-derived peptide ⁵²³LTTPG⁵²⁷ had been identified as an inhibitor of angiotensin converting enzyme and renin activities (Aluko et al., 2015). Moreover, some of the identified sequences, ¹⁵⁴PFPRPP¹⁵⁹, ³⁰³NLRILITLAIPVNKP³¹⁶ and ³⁹⁴SSRKT³⁹⁸ from soy β-conglycinin α-subunit, and glycinin ⁴⁷⁴IKNNNPFKFLVP⁴⁸⁵ were predicted as antimicrobial peptides by CAMP_{R3} prediction software (Waghu, Barai, Gurung, & Idicula-Thomas, 2016). Other authors had described several soybean cationic peptides with antimicrobial activity (Taniguchi et al., 2020), although the peptides found in this work are different from those previously reported.

Identified peptides containing 8 or more amino acids were also matched with reported food epitopes by using Immune Epitope Database and Analysis Resource (www.iedb.org), searching for linear epitopes with positive assays and 90% of sequence homology. Most peptides showing sequence homology with described epitopes belonged to soybean glycinin G1 and β-conglycinin (Tables 2 and 3, respectively), which are considered two major soybean allergens. Matching results with other soybean proteins, glycinin G4, glycinin G5 and 2S albumin are shown in Supplementary Material (Table S2). Glycinin G1-derived peptides released during gastrointestinal digestion corresponded or were contained in previously reported Gly m 6 epitopes, but interestingly, some of them share motifs with Ara h 3 from *Arachis hypogaea* (peanut). For instance, glycinin G1 peptides ³²⁹SSPDYNPQ³³⁷ and ⁴⁷⁴IKNNNPFKFLVP⁴⁸⁵ matched exclusively with Ara h 3 epitopes (Table 2). This is agreement with the described IgE cross-reactivity between peanut and soybean that has been demonstrated clinically, as well as, by epitope mapping (Cabanillas, Jappe, & Novak, 2018). Structural molecular modelling had been previously found homology between peanut Ara h 1 and Ara h 3 with Gly m 5 and Gly m 6, respectively (Chruszcz et al., 2011). According to the location of the trypsin and chymotrypsin cleavage sites in the three-dimensional model of Ara h 3, the resistance of these peanut epitopes to gastrointestinal enzymes had been predicted (Rougé et al., 2009). Our results confirm the survival of these peptides and therefore their potential ability to elicit immune responses when exposed to the intestinal mucosa of allergic individuals. Similar epitopes were also found in a previous study with soybean flour digested with gastrointestinal enzymes, although the sequence of the gastrointestinal resistant peptides were not identical to those found in our work probably due to the differences in the digestion protocol or in the substrate used (soybean flour vs protein isolate) (De Angelis, Pilolli, Bavaro, & Monaci, 2017). This preliminary results should be confirmed with sera from allergic individuals. In addition, it would be interesting to compare the immunoreactivity of the digests obtained from the whole seed vs those from the protein isolates, to study if the better digestibility found for the protein isolates could decrease immunoreactivity.

In addition to soybean, peptides from lentil and grass pea were also matched with described epitopes. Grass pea resistant peptides shared structural homology with reported epitopes from peanut (allergen Ara h1), soybean (Gly m 5) and English walnut (Jug r4), and in a similar way, peptides identified in lentil digests fit with fragments from lentil (allergen Len c1), soybean (Gly m 6) and English walnut (Jug r4) (Supplementary material Table S3).

4. Conclusion

The nitrogen content in the soluble fraction, as well as, the profile of free amino acids could be used to compare digestibility of protein

substrates and digestion performance. At the end of intestinal digestion of the studied protein isolates, 22–28% of the protein is in the form of free amino acids, particularly leucine, lysine, phenylalanine and arginine. The distribution of the nitrogen fraction into free amino acids and peptides of different legume protein isolates after gastrointestinal digestion was similar to that found for milk proteins, except for soybean, where the percentage of insoluble nitrogen was higher. Legume globulins were resistant to gastric digestion but they were hydrolysed into peptides and amino acids during the intestinal phase. In all cases, large-size peptides (6–20 kDa) were detected at the end of the gastric phase, but in the intestinal digests, fragments up to 6 kDa were found, except in the case of soybean, where peptides were below 4 kDa. These results demonstrate the digestibility of several legume protein isolates, being a good source of essential amino acids to be used as alternative ingredients in protein-rich food products. Several peptides resistant to gastrointestinal digestion matched with previously reported linear food epitopes in accordance with the potential to elicit immune responses when exposed to the intestinal mucosa of allergic individuals. Some peptides were matched with epitopes from different species, e.g. soybean and peanut, offering a structural basis founded on digestion-resistant fragments for the cross-reactivity between these species.

CRedit authorship contribution statement

Marta Santos-Hernández: Investigation, Formal analysis, Writing - original draft. **Fabio Alfieri:** Investigation, Formal analysis. **Veronica Gallo:** Investigation, Formal analysis. **Beatriz Miralles:** Writing - review & editing. **Paolo Masi:** Writing - review & editing. **Annalisa Romano:** Writing - review & editing. **Pasquale Ferranti:** Writing - review & editing. **Isidra Recio:** Conceptualization, Supervision.

Declaration of Competing Interest

None.

Acknowledgements

This work has received financial support from project AGL2015-66886-R from the Spanish Ministry of Economy and Competitiveness (MINECO) and project PID2019-107663RB-I00 from Ministry of Science and Innovation (MICINN). M. S-H. is the recipient of FPU grant (FPU15-03616) from the Ministry of Education, Culture and Sports. Dr. R. Olías is acknowledged for the determination of the enzyme inhibitory activity in the plant protein isolates.

Appendix A. Supplementary material

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.foodres.2020.109708>.

References

- Adibi, S. A., & Mercer, D. W. (1973). Protein digestion in human intestine as reflected in luminal, mucosal, and plasma amino acid concentrations after meals. *Journal of Clinical Investigation*, 52, 1586–1594.
- Aluko, R. E., Girgih, A. T., He, R., Malomo, S., Li, H., Offengenden, M., & Wu, J. (2015). Structural and functional characterization of yellow field pea seed (*Pisum sativum* L.) protein-derived antihypertensive peptides. *Food Research International*, 77, 10–16.
- Aryee, A. N. A., & Boye, J. I. (2016). Improving the digestibility of lentil flours and protein isolate and characterization of their enzymatically prepared hydrolysates. *International Journal of Food Properties*, 19, 2649–2665.
- Barac, M., Cabrilo, S., Pesic, M., Stanojevic, S., Zilic, S., Macej, O., & Ristic, N. (2010). Profile and functional properties of seed proteins from six pea (*Pisum sativum*) genotypes. *International Journal of Molecular Sciences*, 11, 4974–4991.
- Barbana, C., & Boye, J. I. (2013). In vitro protein digestibility and physico-chemical properties of flours and protein concentrates from two varieties of lentil (*Lens culinaris*). *Food & Function*, 4, 310–321.
- Bishnoi, S., & Khetarpaul, N. (1994). Protein digestibility of vegetables and field peas (*Pisum-sativum*) - varietal differences and effect of domestic processing and cooking methods. *Plant Foods for Human Nutrition*, 46, 71–76.

- Bohn, T., Carriere, F., Day, L., Deglaire, A., Egger, L., Freitas, D., ... Dupont, D. (2018). Correlation between in vitro and in vivo data on food digestion. What can we predict with static in vitro digestion models? *Critical Reviews in Food Science and Nutrition*, 58, 2239–2261.
- Brodtkorb, A., Egger, L., Alminger, M., Alvito, P., Assuncao, R., Ballance, S., ... Recio, I. (2019). INFOGEST static in vitro simulation of gastrointestinal food digestion. *Nature Protocols*, 14, 991–1014.
- Cabanillas, B., Jappe, U., & Novak, N. (2018). Allergy to peanut, soybean, and other legumes: Recent advances in allergen characterization, stability to processing and IgE cross-reactivity. *Molecular Nutrition & Food Research*, 62, 1700446.
- Carre, B., Beaufils, E., & Melcion, J. P. (1991). Evaluation of protein and starch digestibilities and energy value of pelleted or unpelleted pea-seeds from winter or spring cultivars in adult and young chickens. *Journal of Agricultural and Food Chemistry*, 39, 468–472.
- Chruszcz, M., Maleki, S. J., Majorek, K. A., Demas, M., Bublin, M., Solberg, R., ... Minor, W. (2011). Structural and immunologic characterization of Ara h 1, a major peanut allergen. *The Journal of Biological Chemistry*, 286, 39318–39327.
- Clemente, A., MacKenzie, D. A., Jeenes, D. J., & Domoney, C. (2004). The effect of variation within inhibitory domains on the activity of pea protease inhibitors from the Bowman-Birk class. *Protein Expression and Purification*, 36, 106–114.
- De Angelis, E., Pilolli, R., Bavaro, S. L., & Monaci, L. (2017). Insight into the gastro-duodenal digestion resistance of soybean proteins and potential implications for residual immunogenicity. *Food & Function*, 8, 1599–1610.
- Domoney, C., & Welham, T. (2008). Trypsin inhibitors in *Pisum*: Variation in amount and pattern of accumulation in developing seed. *Seed Science Research*, 2, 147–154.
- Egger, L., Schlegel, P., Baumann, C., Stoffers, H., Guggisberg, D., Brugger, C., ... Portmann, R. (2017). Physiological comparability of the harmonized INFOGEST in vitro digestion method to in vivo pig digestion. *Food Research International*, 102, 567–574.
- El, S. N., Karakaya, S., Simsek, S., Dupont, D., Menfaatli, E., & Eker, A. T. (2015). In vitro digestibility of goat milk and kefir with a new standardised static digestion method (INFOGEST cost action) and bioactivities of the resultant peptides. *Food & Function*, 6, 2322–2330.
- Faris, R. J., Wang, H., & Wang, T. (2008). Improving digestibility of soy flour by reducing disulfide bonds with thioredoxin. *Journal of Agricultural and Food Chemistry*, 56, 7146–7150.
- Han, I. H., Swanson, B. G., & Baik, B. K. (2007). Protein digestibility of selected legumes treated with ultrasound and high hydrostatic pressure during soaking. *Cereal Chemistry*, 84, 518–521.
- Karamac, M., & Rybarczyk, A. (2008). Chymotryptic hydrolysis of lentil meal proteins and characteristics of the resulting hydrolysates. *Polish Journal of Food and Nutrition Sciences*, 58, 351–357.
- Kuipers, B. J. H., van Koningsveld, G. A., Alting, A. C., Driehuis, F., Voragen, A. G. J., & Gruppen, H. (2006). Opposite contributions of glycinin- and beta-conglycinin-derived peptides to the aggregation behavior of soy protein isolate hydrolysates. *Food Biophysics*, 1, 178–188.
- Le Gall, M., Gueguen, J., Seve, B., & Quillien, L. (2005). Effects of grinding and thermal treatments on hydrolysis susceptibility of pea proteins (*Pisum sativum* L.). *Journal of Agricultural and Food Chemistry*, 53, 3057–3064.
- Lioi, L., Sparvoli, F., Sonnante, G., Laghetti, G., Lupo, F., & Zaccardelli, M. (2011). Characterization of Italian grasspea (*Lathyrus sativus* L.) germplasm using agronomic traits, biochemical and molecular markers. *Genetic Resources and Crop Evolution*, 58, 425–437.
- Longobardi, F., Innamorato, V., Di Gioia, A., Ventrella, A., Lippolis, V., Logrieco, A. F., ... Agostiano, A. (2017). Geographical origin discrimination of lentils (*Lens culinaris* Medik.) using H-1 NMR fingerprinting and multivariate statistical analyses. *Food Chemistry*, 237, 743–748.
- Manguy, J., Jehl, P., Dillon, E. T., Davey, N. E., Shields, D. C., & Holton, T. A. (2017). Peptigram: A web-based application for peptidomics data visualization. *Journal of Proteome Research*, 16, 712–719.
- Mariotti, F., Tomé, D., & Mirand, P. P. (2008). Converting nitrogen into protein—beyond 6.25 and Jones' factors. *Critical Reviews in Food Science and Nutrition*, 48, 177–184.
- Nguyen, T. T. P., Bhandari, B., Cichero, J., & Prakash, S. (2015). Gastrointestinal digestion of dairy and soy proteins in infant formulas: An in vitro study. *Food Research International*, 76, 348–358.
- Picariello, G., Amigo-Benavent, M., del Castillo, M. D., & Ferranti, P. (2013). Structural characterization of the N-glycosylation of individual soybean beta-conglycinin subunits. *Journal of Chromatography A*, 1313, 96–102.
- Porzucek, H., Larssonraznikiewicz, M., & Klepacka, M. (1991). In vitro protein digestibility of flours and protein isolates from seeds of some leguminous plants. *Swedish Journal of Agricultural Research*, 21, 49–53.
- Reynaud, Y., Lopez, M., Riaublanc, A., Souchon, I., & Dupont, D. (2020). Hydrolysis of plant proteins at the molecular and supra-molecular scales during in vitro digestion. *Food Research International*, 134, Article 109204.
- Romano, A., Giosafatto, C. V. L., Al-Asmar, A., Masi, P., Aponte, M., & Mariniello, L. (2019b). Grass pea (*Lathyrus sativus*) flour: Microstructure, physico-chemical properties and in vitro digestion. *European Food Research and Technology*, 245, 191–198.
- Romano, A., Giosafatto, C. V. L., Al-Asmar, A., Masi, P., Romano, R., & Mariniello, L. (2019a). Structure and in vitro digestibility of grass pea (*Lathyrus sativus* L.) flour following transglutaminase treatment. *European Food Research and Technology*, 245, 1899–1905.
- Rosa, M. J. S., Ferreira, R. B., & Teixeira, A. R. (2000). Storage proteins from *Lathyrus sativus* seeds. *Journal of Agricultural and Food Chemistry*, 48, 5432–5439.
- Rougé, P., Culierrier, R., Sabatier, V., Granier, C., Rancé, F., & Barre, A. (2009). Mapping and conformational analysis of IgE-binding epitopic regions on the molecular surface of the major Ara h 3 legumin allergen of peanut (*Arachis hypogaea*). *Molecular Immunology*, 46, 1067–1075.
- Ruan, Q. J., Chen, Y. M., Kong, X. Z., & Hua, Y. F. (2015). Analysis using fluorescence labeling and mass spectrometry of disulfide-mediated interactions of soy protein when heated. *Journal of Agricultural and Food Chemistry*, 63, 3524–3533.
- Rubio, L. A., & Clemente, A. (2009). In vivo (rat) and in vitro (Caco-2 cells) absorption of amino acids from legume protein isolates as compared to lactalbumin or casein. *Archives of Animal Nutrition*, 63, 413–426.
- Sanchón, J., Fernández-Tomé, S., Miralles, B., Hernández-Ledesma, B., Tomé, D., Gaudichon, C., & Recio, I. (2018). Protein degradation and peptide release from milk proteins in human jejunum. Comparison with in vitro gastrointestinal simulation. *Food Chemistry*, 239, 486–494.
- Santos-Hernández, M., Amigo, L., & Recio, I. (2020). Induction of CCK and GLP-1 release in enteroendocrine cells by egg white peptides generated during gastrointestinal digestion. *Food Chemistry*, 329, Article 127188.
- Santos-Hernández, M., Tomé, D., Gaudichon, C., & Recio, I. (2018). Stimulation of CCK and GLP-1 secretion and expression in STC-1 cells by human jejunal contents and in vitro gastrointestinal digests from casein and whey proteins. *Food & Function*, 9, 4702–4713.
- Schreuders, F. K. G., Dekkers, B. L., Bodnár, I., Erni, P., Boom, R. M., & van der Goot, A. J. (2019). Comparing structuring potential of pea and soy protein with gluten for meat analogue preparation. *Journal of Food Engineering*, 261, 32–39.
- Tagliazucchi, D., Martini, S., Shamsia, S., Helal, A., & Conte, A. (2018). Biological activities and peptidomic profile of in vitro-digested cow, camel, goat and sheep milk. *International Dairy Journal*, 81, 19–27.
- Taniguchi, M., Aida, R., Saito, K., Kikura, T., Ochiai, A., Saitoh, E., & Tanaka, T. (2020). Identification and characterization of multifunctional cationic peptides from enzymatic hydrolysates of soybean proteins. *Journal of Bioscience and Bioengineering*, 129, 59–66.
- Tomé, D. (2013). Digestibility issues of vegetable versus animal proteins: Protein and amino acid requirements-functional aspects. *Food and Nutrition Bulletin*, 34, 272–274.
- Waghu, F. H., Barai, R. S., Gurung, P., & Idicula-Thomas, S. (2016). CAMP(R3): A database on sequences, structures and signatures of antimicrobial peptides. *Nucleic Acids Research*, 44, D1094–D1097.
- Yan, Z. Y., Spencer, P. S., Li, Z. X., Liang, Y. M., Wang, Y. F., Wang, C. Y., & Li, F. M. (2006). *Lathyrus sativus* (grass pea) and its neurotoxin ODAP. *Phytochemistry*, 67, 107–121.

Compared digestibility of plant protein isolates by using the INFOGEST digestion protocol

Marta Santos-Hernández¹, Fabio Alfieri², Veronica Gallo², Beatriz Miralles¹, Paolo Masi², Annalisa Romano², Pasquale Ferranti², Isidra Recio^{1*}

¹Instituto de Investigación en Ciencias de la Alimentación, CIAL (CSIC-UAM, CEI UAM+CSIC), Nicolás Cabrera, 9, 28049 Madrid, Spain.

²Department of Agricultural Sciences, Division of Food Science and Technology, University of Naples Federico II, Via Università 100, 80055 Portici, Naples, Italy.

* Corresponding author: I. Recio

Nicolás Cabrera, 9. 28049 Madrid, Spain

Phone: +34 910017940

Fax: +34 910017905

e-mail: i.recio@csic.es

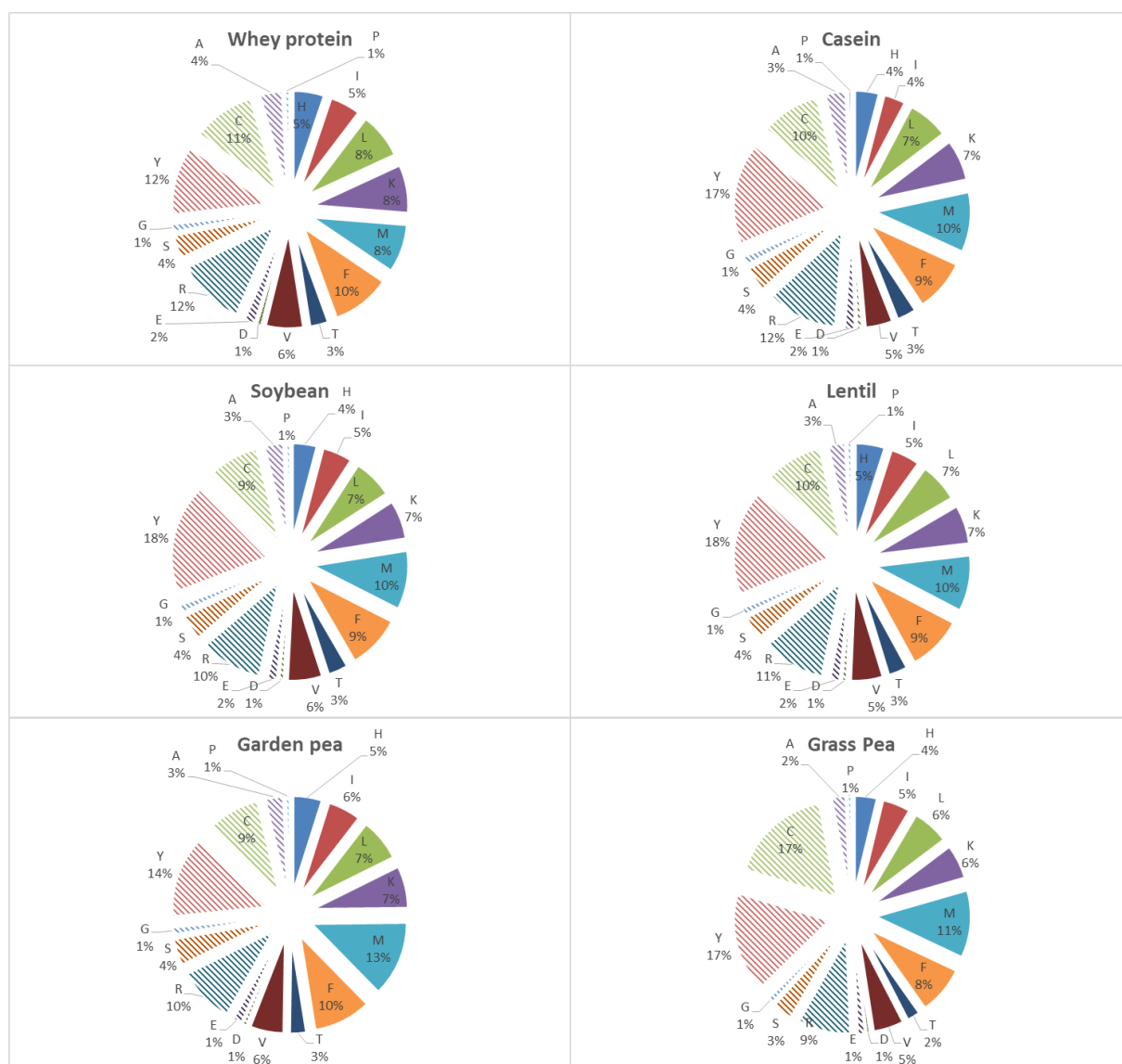


Figure S1: Free amino acids at the end of the gastrointestinal digestion expressed as percentage of the total amount of each amino acid in the soluble part of the gastrointestinal digests from the following protein isolates: garden pea, grass pea, soybean and lentil, casein and whey protein. Solid colors correspond to essential amino acids and frame-colored sectors correspond to non-essential amino acids.

Table S1: Peptides identified by HPLC-MS/MS at the end of the gastrointestinal digestion of the following plant protein isolates: garden pea, grass pea, soybean and lentil. For each peptide, the precursor ion used for fragmentation, the experimental and the theoretical mass together with the position in the protein and sequence are given. The precursor protein with the Uniprot code of each protein is given (See Supplementary Excel file).

Precursor Ion	z	Experimental Mass	Theoretical Mass	Range		Sequence
Glycinin G1 - Uniprot code P04776 (GLYG1_SOYBN) - Glycine max						
657.5	1+	656.5	656.324	23	27	REQPQ
1103.5	1+	1102.5	1102.435	24	32	EQPQQNECQ
975.6	1+	974.6	974.376	24	31	EQPQQNEC
771.5	1+	770.5	770.392	37	43	NALKPDN
747.5	1+	746.5	746.356	44	50	RIESEGG
1047.7	1+	1046.7	1046.513	45	54	IESEGGLIET
704.5	1+	703.5	703.339	45	51	IESEGGL
873.6	1+	872.6	872.403	52	58	IETWNPN
970.6	1+	969.6	969.467	54	61	TWNPNNKP
683.5	1+	682.5	682.34	56	61	NPNNKP
1025.7	1+	1024.7	1024.436	101	110	GMIYPGCPST
924.6	1+	923.6	923.388	101	109	GMIYPGCPS
736.4	1+	735.4	735.326	103	109	IYPGCPS
837.5	1+	836.5	836.374	103	110	IYPGCPST
597.3	1+	596.3	596.292	114	118	PQQPQ
874.6	1+	873.6	873.394	121	128	GQSSRPQD
759.5	1+	758.5	758.367	121	127	GQSSRPQ
758.6	1+	757.6	757.383	124	129	SRPQDR
736.4	1+	735.4	735.355	136	141	FREGDL
589.4	1+	588.4	588.287	137	141	REGDL
914.7	1+	913.7	913.476	138	146	EGDLIAVPT
951.6	1+	950.6	950.398	153	160	YNNEDTPV

.... (see Supplementary Excel file)

Table S2: Identified peptides derived from soybean albumin, glycinin G5, and glycinin G4 longer than 8 amino acids and reported related epitopes. Common sequences are highlighted in red. Homologous epitopes to the identified are searched by using the public resource Immune Epitope Database and Analysis Resource (IEDB) selecting linear epitopes with positive assays and 90% identity percentage.

2S albumin - UniProt: P19594 (2S5_SOYBN)						
Mass ^a	Position	Identified sequence	Related epitopes	Epitope ID ^b	Allergen	Organism ^c
986.6	89-96	EEEGHMQK	EE EEGHMQK C	561645	Gly m 8	<i>Glycine max</i>
			EEGHMQK CCSEMSEL	561647	Gly m 8	<i>Glycine max</i>
862.8	106-113	SPKCQCKA	SELR SPKCQCKA	561749	Gly m 8	<i>Glycine max</i>
Glycinin G5 – UniProt: P04347 (GLYG5_SOYBN)						
1020.6	145-154	DVLVIPLGVP	LV IPLG	913666	Glycinin G5	<i>Glycine max</i>
			RARLSPG DVLVIPAG	141735	Len c 1	<i>Lens culinaris</i>
1291.8	189-200	LAGNPDIEHPET	DIEHPET	913002	Glycinin G5	<i>Glycine max</i>
1178.8	190-200	AGNPDIEHPET				
1206.7	275-283	WQEQEDEDE	WQEQEDED	914280	Glycinin G5	<i>Glycine max</i>
1020.6	276-283	QEQEDEDE				
1123.6	280-288	DEDEDEDEE	EDEDEE EEGSNREVR	99173	Ara h 1	<i>Arachis hypogaea</i>
2234.3	300-318	PSHGKHEDDEDEDEEDQP	EDEEDQ RP	913066	Glycinin G5	<i>Glycine max</i>
Glycinin G4 – UniProt: P02858 (GLYG4_SOYBN)						
893.5	158-165	YNTGDEPV	GV ^c PYWT YNTGDEPV V	187839	Glycinin G4	<i>Glycine max</i>

^a Experimental mass

^b Epitope ID in the Immune Epitope Database and Analysis Resource (www.iedb.org)

^c Common names: *Glycine max*, soybean; *Arachis hypogaea*, peanut; *Lens culinaris*, lentil.

Table S3: Identified peptides from lentil protein isolate longer than 8 amino acids and reported related epitopes. Common sequences are highlighted in red. Homologous epitopes to the identified are searched by using the public resource Immune Epitope Database and Analysis Resource (IEDB) selecting linear epitopes with positive assays and 90% identity percentage.

Allergen Len c. 1.0101 - UniProt: Q84UI1 (Q84UI1_LENCU)						
Mass ^a	Position	Identified sequence	Related epitopes	Epitope ID ^b	Allergen	Organism ^c
932.4	371-379	AGEEDNVIS	EEDNVISQIQRPVKELAFPGSSREVDR	141628	Len c 1	Lens culinaris
1045.5	370-379	LAGEEDNVIS	LIGFGINAKNNQRNFLAGEEDNVIS	141700	Len c 1	Lens culinaris
			NNQRNFLAGEEDNVI	141716	Len c 1	Lens culinaris
			RNFLAGEEDNVISQI	141741	Len c 1	Lens culinaris
Legumin J - UniProt: P05692 (LEGJ_PEA)						
1223.5	54-63	TETWNPNHPE	RIESEGGYIETWNP	54145	Ara h 3	Arachis hypogaea
			GLIETWNPNNKPFQCA	187829	Gly m 6	Glycine max
			IESEGGLIETWNPNNK	187845	Gly m 6	Glycine max
949.4	82-90	LSFPGCPET	TGVLFPGCPETFEES	114627	Jug r 4	Juglans regia
			AFPGCP	912924	Glycinin G5	Glycine max
Legumin A2 primary translation product - UniProt: Q99304_VICFA						
1064.6	481-498	SNNPFKFLV	PREQARQLKNNNPFK	106008	Ara h 3	Arachis hypogaea

^a Experimental mass

^b Epitope ID in the Immune Epitope Database and Analysis Resource (www.iedb.org)

^c Common names: *Glycine max*, soybean; *Arachis hypogaea*, peanut; *Juglans regia*, English walnut.