Theses of Doctoral (PhD) Dissertation

INVESTIGATION OF EXTRACTION EFFICIENCY, PHYSICAL AND NUTRITIONAL PROPERTIES OF LEGUME PROTEIN

Chaima Neji

Dissertation supervisor: **Prof. Dr. Péter Sipos**

UNIVERSITY OF DEBRECEN

Doctoral School of Food Sciences

Debrecen, 2026

1. THE ANTECEDENTS AND THE OBJECTIVES OF THE DOCTORAL DISSERTATION

Industries are increasingly adopting fractionation to extract compounds with specific properties, valorize unused materials, or convert waste. The refined components can be used as food ingredients with particular nutritional or functional properties, such as emulsifiers or thickeners. These properties make the targeted products valuable to various users, including industries, retailers, and health-conscious consumers, as they have the potential to support personalized nutrition.

Protein, as a primary macronutrient, is subject to fractionation techniques to achieve a protein concentrate with high purity and specific functional properties. Nowadays, with the surge in people's awareness of the importance of a protein-rich daily diet (Patra et al., 2023), affordable, abundant, sustainable sources of protein are needed. Among plants, legumes offer several advantages from an environmental, agricultural, and nutritional perspective. They provide low greenhouse gas and water footprints and can enhance soil quality through nitrogen fixation (Semba et al., 2021). In addition to being affordable, they provide a sustainable source of protein (Semba et al., 2021), varying between 20% and 40% (Shevkani, 2023). Thanks to their properties, these proteins allow replacing existing animal-source proteins in various applications (Goldstein & Reifen, 2022).

From an industrial perspective, maintaining a balance between production cost-effectiveness, sustainability, and product quality is crucial. With the increasing emphasis on environmentally friendly approaches (ionic liquids, deep eutectic solvents), there is an urgent need to optimize extraction conditions. Therefore, establishing a modelling framework can improve process performance while reducing the number of trials, and the integration of computational tools and predictive modelling can provide a promising strategy to minimize resource use. However, to ensure the reliability of these results, validation is essential.

From a nutritional perspective, although plant-based proteins, especially those from legumes, are recognized for their nutritional value, their use is restricted due to allergens and antinutritional factors such as lectins, trypsin inhibitors, phytic acid, phenolics, oxalates, saponins, tannins, and cyanogenic glycosides. While fractionation and processing methods (such as enzymatic hydrolysis, fermentation, and roasting) can reduce or mitigate these effects, peptides with bioactive potential can be produced. Therefore, the characterisation of these peptides may reveal promising new functional properties. Additionally, in vivo model is also

available as a valuable tool for evaluating and validating the nutritional efficacy of diets that are either restricted to protein isolate or enriched with it.

This work includes a comprehensive study that integrates both industrial and nutritional aspects related to the plant protein extracts. Among legumes, soybean is the most widely produced legume crop (Semba et al., 2021) with a cultivated land area extended to nearly 136.9 million hectares in 2023, resulting in a harvest of 371.17 million tons (*FAOSTAT*, 2025). In this research, soybean was selected as a focus material to:

1. Establish a predictive model for estimating the purity of protein extracted using conventional and new methods (Deep Eutectic solvents).
2. Assess the physicochemical properties of the extracted protein using techniques such as SDS PAGE, fluorescence, and microscopy.

On the other hand, with the increase in single amino acid (Xiao et al., 2023) and protein extract supplement in food, the effect of the addition of lysine and soybean protein extract was studied with the aim to:

3. Elucidate the potential nutritional effect of a normal and an exclusive high-protein (soybean protein) or lysine-containing diet using the vivo model *Drosophila melanogaster*.
4. Among plant-based bioactive peptides, those originating from cereals and legumes exhibit notable health-promoting properties (Malaguti et al., 2014). Their wide range of physiological effects has been validated through in vitro studies and animal models (Malaguti et al., 2014). For instance, bioactive hydrolysates and peptides from pulses (such as peas, chickpeas, cowpeas, lupin, and particularly soybean and bean) have shown promising potential. However, while in vitro tests support their benefits, no clinical trials have yet been conducted, underscoring the need for further research (Garcés-RimónD et al., 2022). Therefore, as a start, this study focuses on the:
5. Analyse the primary sequences of proteins using bioinformatics tools for structural insights.

The aims of this study are outlined (Figure 1).

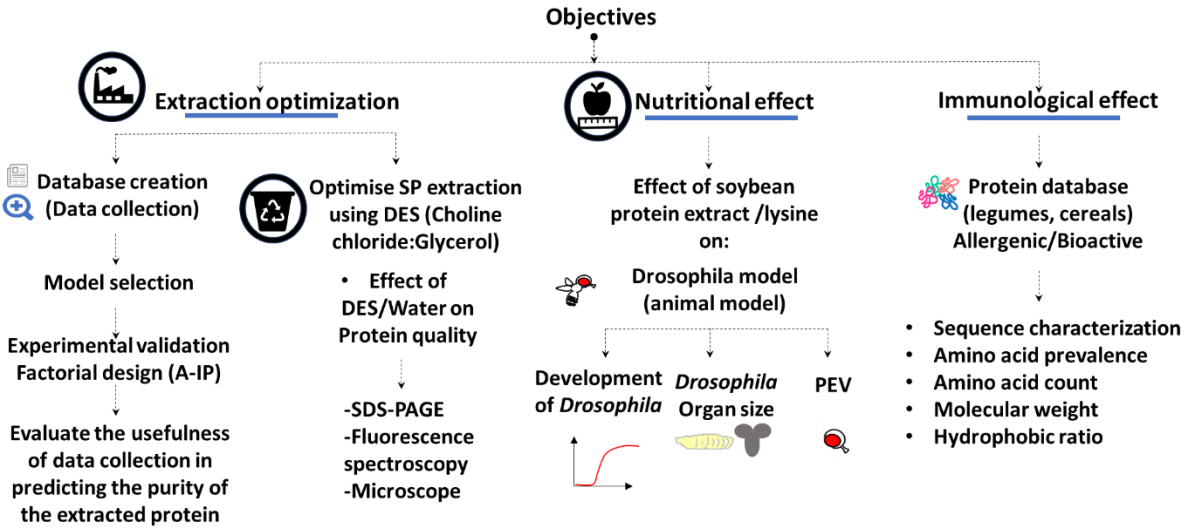

Figure 1. Study objectives and workflow

2. MATERIALS AND METHODS

2.1. Protein extraction

Es Mentor and Isidor are the used varieties of soybean that were cultivated at the University of Debrecen, Institutes for Agricultural Research and Educational Farm, Research Institute of Nyíregyháza (Nyíregyháza, Hungary) for protein extraction using the conventional extraction method (alkaline-isoelectric precipitation) and using deep eutectic solvent, respectively.

2.1.1. Extraction using data analysis and conventional method

To create the mini database, data including information about pH of extraction and precipitation, and the extraction temperature was collected and selected as factors. However, the purity of the protein, which reflects the crude protein content of the recovered protein, was considered as a response. The data was then divided into 70% for training and 30% for testing sets. For this research, a group of algorithms were identified: linear regression, a polynomial model, and machine learning algorithms.

This research employed a two-level full factorial design to validate the initial model. Extraction temperature was fixed at 25 °C and 45 °C, representing the low level (-1) and the high level (+1). A total of four randomized experiments based on two variables (2^2=4) were performed in triplicate according to the extraction and the adjusted procedure of Johansson et al, (2022).

The purity of the protein, solid yield (Eq. 1), and protein yield (Eq. 2) of all samples were calculated based on dry weight.

Solids yield (%) = *Weight of protein isolate*100/ Weight of soybean flour* (Eq. 1) (Chiodza & Goosen, 2024)

Protein yield (%) = Mass of protein in protein isolate*100/Mass of protein in the soybean flour (Eq. 2) (Chamba et al., 2015)

Prediction error was calculated to evaluate the model's efficiency (Eq. 3 and 4) (Guang et al., 1995)

Percentage prediction error = *(Measured value-Predicted value)*100/ Measured value* (Eq. 3)

Percentage prediction error = *(Predicted value- Measured value)*100/ Measured value* (Eq. 4)

2.1.2. Extraction using Deep eutectic solvent

2.1.2.1. Preparation and characterisation of deep eutectic solvents

The preparation of deep eutectic solvents (DES) using Choline chloride and glycerol (molar ratio of 1:2) followed the method of Yue et al, (2021). Then, the mixture was stirred with a magnetic stirrer at 25 °C until a homogeneous and transparent liquid was formed, approximately after 20 minutes.

The pH, the electrical conductivity, and density of the different mixtures containing 30:70%, 50:50% and 60:40% of DES (ChCl: Gly): water was studied.

2.1.2.2. Protein extraction using DES/water mixture

Protein extraction was performed in accordance with the procedure by Q. Chen et al., (2021) with some modifications. 10 g of DES/water mixture was mixed with 1 g of soybean flour. After that, the samples were stirred at 160 rpm (Infors HT Ecotron, Switzerland) for a specific duration and temperature (Table 1), then centrifuged (Hettich, Germany) at 4020 RCF for 20 minutes to collect the supernatant. Each experimental condition was repeated three times.

2.1.2.3. Optimization of protein extraction

- **Identification of the factors affecting the protein extraction**

In this research, part of the experimental runs was conducted using Behnken design (Figure 2).

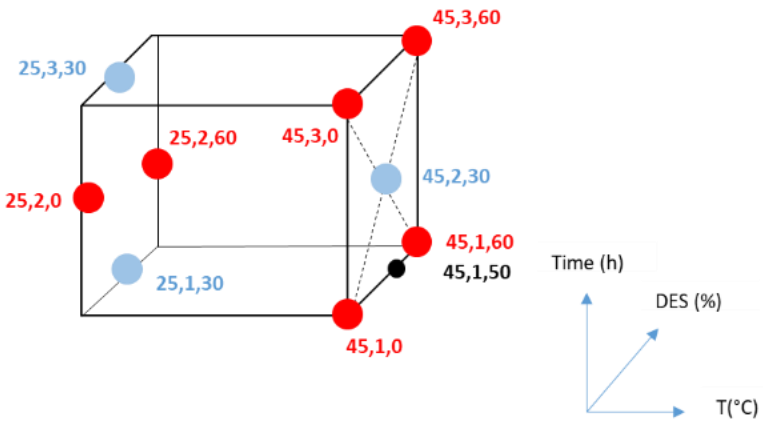

Figure 2. Applied combination of factors including extraction time (Time (h)), DES (%), and extraction temperature (T (°C))

The nitrogen content of each experiment was estimated using the Kjeldahl method was (Evers & Hughes, 2002) according to AOAC procedures AOAC 2001.11 (Latimer, 2023).

In this study 80% was designed for training while 20% was allocated for model testing. To enhance the model's reliability, K-fold cross-validation was employed with K set to 10.

2.2. Evaluation of protein quality

2.2.1. Sample preparation

- **Protein isolated with DES solvent**

2 g of soybean flour was incorporated into 20 ml of DES (ChCl:Glycerol): Water in the ratio of 60% of DES and 40% of water, to evaluate the impact of the binary mixture of DES: Water, since 65 °C is recognised as the denaturation temperature for the 7S fraction (Peng et al., 2016).

- **Protein isolate modified with enzymatic hydrolysis**

The sample was prepared according to the research of Z. Wang et al, (2014), with some modifications. The protein was dispersed in distilled water to a final ratio of 1:15 (w/v) and adjusted to pH 8.0 with 1 mol/L NaOH. The hydrolysis was carried out with 0.2% of trypsin at 37 °C for 90 minutes with continuous agitation. Throughout hydrolysis, the solution's pH was continuosly adjusted to pH 8.0 with 1 mol/L NaOH. To stop the reaction, the mixture was heated in boiling water for 10 minutes to inactivate the enzyme. After cooling to room temperature, the hydrolysate was centrifuged at 4020 RCF for 15 minutes. The resulting precipitate was freeze-dried.

- **Protein isolates subjected to physical modification**

The chemical modification of the soybean protein was prepared according to F. Liu & Tang, (2013) with some changes. Briefly, 6% (w/v) solution was prepared by dispersing the soybean protein isolate in distilled water with continuous stirring for two hours under a magnetic stirring and then kept overnight at 4 °C for full hydration of the proteins. The solution was then heated to 95 °C for 15 minutes in a bath. After cooling, the solution was sonicated at a frequency sweeping between 50-60 Hz for 30 minutes using an ultrasound processor model Olympus Endosonic supplied by the manufacturer (KeyMed, UK). After that, 300mM of NaCl was added and centrifuged for 15 minutes at 4020 RCF. Finally, the collected pellet was freeze-dried.

2.2.2. Evaluation of the characteristics of the extracted and treated protein

- Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was carried out following the study of Domokos-Szabolcsy et al., (2022) with some modifications.

10 mg of the homogenised sample was dissolved in 250 µL of (2× Laemmli). The sample solution was then mixed with the buffer. After incubation at 95 °C for 5 minutes, followed by centrifugation (13000 rpm, 20 minutes) at 4°C, 10 µL of the supernatant of each sample was loaded in a well alongside a standard marker (10-180 kDa). The resolving gel was 12.5%, and the stacking gel was 9%, prepared in a Mini-Protean Tetra Cell gel system (Bio-Rad Inc. Hercules, MI, USA). Both stacking and separating gels were run at 90 V for 15 minutes and 160 V for 59 minutes, respectively. After that, the gel was stained with Coomassie Brilliant Blue (G-250) solution. Scanning was performed using the Imaging System (BioRad ChemiDoc MP Imaging System 5.2.1, USA).

- The intrinsic fluorescence spectra of the isolated soybean protein were measured with an F-8500 Spectrofluorometer (Jasco, Oklahoma, United States).
- Light microscope: The images were captured at room temperature using a light microscope (Olympus Corporation SZX2-ILLK, Tokyo, Japan) and processed with OLYMPUS cellSens Entry 2.3 software.

2.3. Assessment of the nutritional effect of soybean protein

2.3.1. Description of *Drosophila melanogaster* in vivo analytical system

The evaluation of nutritional effects by *Drosophila melanogaster* in vivo digestion experiment was conducted using w^{m4h} (white mottled 4) strain sourced from Bloomington Stock Center, which was used for all the samples. The experiment was performed following the method described by (Aleya et al., 2023), with some modifications. The effect of soybean protein (SP), lysine and the mixture of soybean protein and lysine (2%, 1%, 0.5%) (Table 1) were evaluated.

Table 1. Experimental parameters

Total concentration of mixture/0N-diet (%)		
2%	1%	0.5%
Soybean protein+Lysine mixture composition (%)		
1	25-75	
2	50-50	
3	75-25	

Evaluated parameters

- Pupae length and width measurements

Pupal size (length and width) was determined according to the method of Enriquez et al, (2022). The length and width were measured using the freely accessible online software 'Imagej' (*ImageJ*, n.d.).

- Brain dissection and optic lobe length

The larval dissection to access the brain was done according to the protocol of (Bayala et al., 2024) used to access imaginal discs with some modifications. Pictures of the dissected brains were taken using the optical microscope Olympus Corporation (Model SZX2-ILLK, Japan), and the length of optic lobes was measured using OLYMPUS cellSens Entry 2.3 software.

- Determination of Drosophila Red Eye Pigment Concentration

To study the effect of diet on the flies, a close-up of the fly's head image was taken using the optical microscope Olympus Corporation (Model SZX2-ILLK, Japan). The analysis was carried out according to the study of Sun et al, (2002) with some modifications.

2.3.2. Dataset preparation and evaluation for epitope characterisation

For peptides, a group of 263 selected sequences of allergens and bioactive peptides (immunomodulatory/anti-inflammatory/anti-proliferative) from cereals and legumes were collected from the online Structural Database of Allergenic Proteins (*SDAP 2.0 : Structural Database of Allergenic Proteins*, 2023) and research articles. During the data preparation, the sequences were characterized by their physicochemical properties and composition (amino acid count) after the removal of the duplicated sequences.

For epitopes, 1963 allergenic epitopes from different sources (legumes, fish, and cereals), and 32 immunosuppressive epitopes were extracted from the IEDB-AR database (IEDB.org: View Results and Refine Search). The characterisation of peptides includes the amino acid count (prevalence frequencies of 20 amino acids), molecular weight, the aliphatic index (relative volume of a protein occupied by aliphatic side chains (A, V, I, and L)(Ikai, 1980), aromaticity (relative frequency of aromatic amino acids (F, W, and Y)), hydrophobic ratio (GES scale, which is based on energetic considerations of residues in α-helices (Koehler et al., 2009), the frequency of amino acids in the termini of each sequence, the series of potential epitopes.

Concerning the epitopes, the amino acid count, molecular weight, charge, hydrophobicity, the prevalence of each amino acid in the epitope, and the relative position of W, F, and Y are the studied features.

2.4. Statistical analysis

All the results were expressed as means ± standard deviations of experimental triplicate from each independent experiment. Statistically significant values among groups were compared using the one-way analysis of variance (ANOVA). The value of $p < 0.05$ was regarded as statistically significant.

Minitab (version 21), Python (Jupyter notebook), R, GraphPad Prism version 9.5.0 (730) are the used softwares.

3. RESULTS

3.1. Protein extraction

3.1.1. Extraction using data analysis and conventional method

The correlation matrix (Figure 3) reveals a low positive correlation between the purity of the protein isolate (response) and the pH of the extraction. There is a moderately strong positive correlation between the pH of precipitation and the response, as well as between the pH of extraction and the pH of precipitation. Conversely, a weak correlation between extraction temperature and the response suggests this factor can be disregarded.

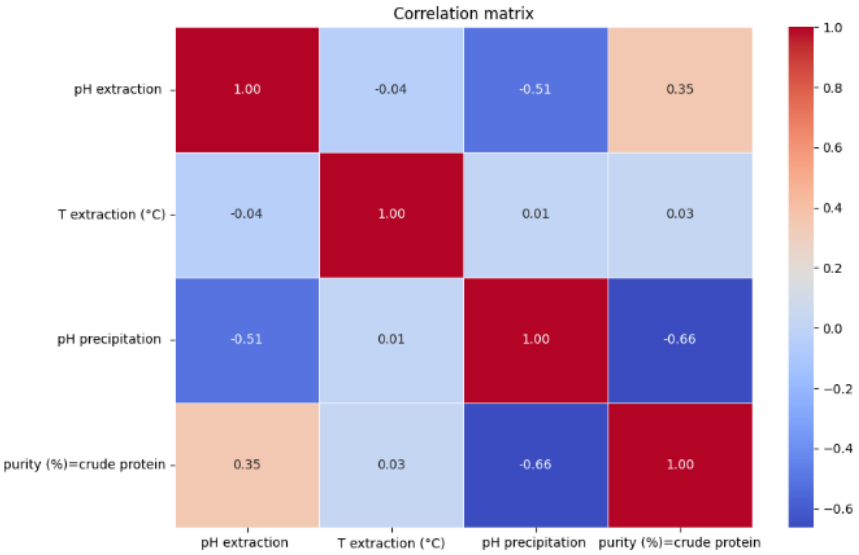

Figure 3. Correlation among extraction factors (temperature, pH of extraction, and pH of precipitation) and the response (purity of protein)

Optimization of soybean protein isolate extraction conditions

- **Model fitting**

Considering the results from various models, the voting regressor yields achieves the highest prediction accuracy, and the lowest mean squared error (RMSE). Using both methods—the train-test split (70:30) and 10-fold cross-validation, the Voting regressor achieved its highest R^2 value.

- **Validation of protein extraction process**

According to the results from the mini database, the alkaline extraction of soybean flour, presenting 37.76±0.32% of protein (dry basis), resulted in a relatively pure protein isolate

(Figure 4). This is likely because the alkaline solution can break down the plant cell matrix and solubilise the protein (Hadidi et al., 2023).

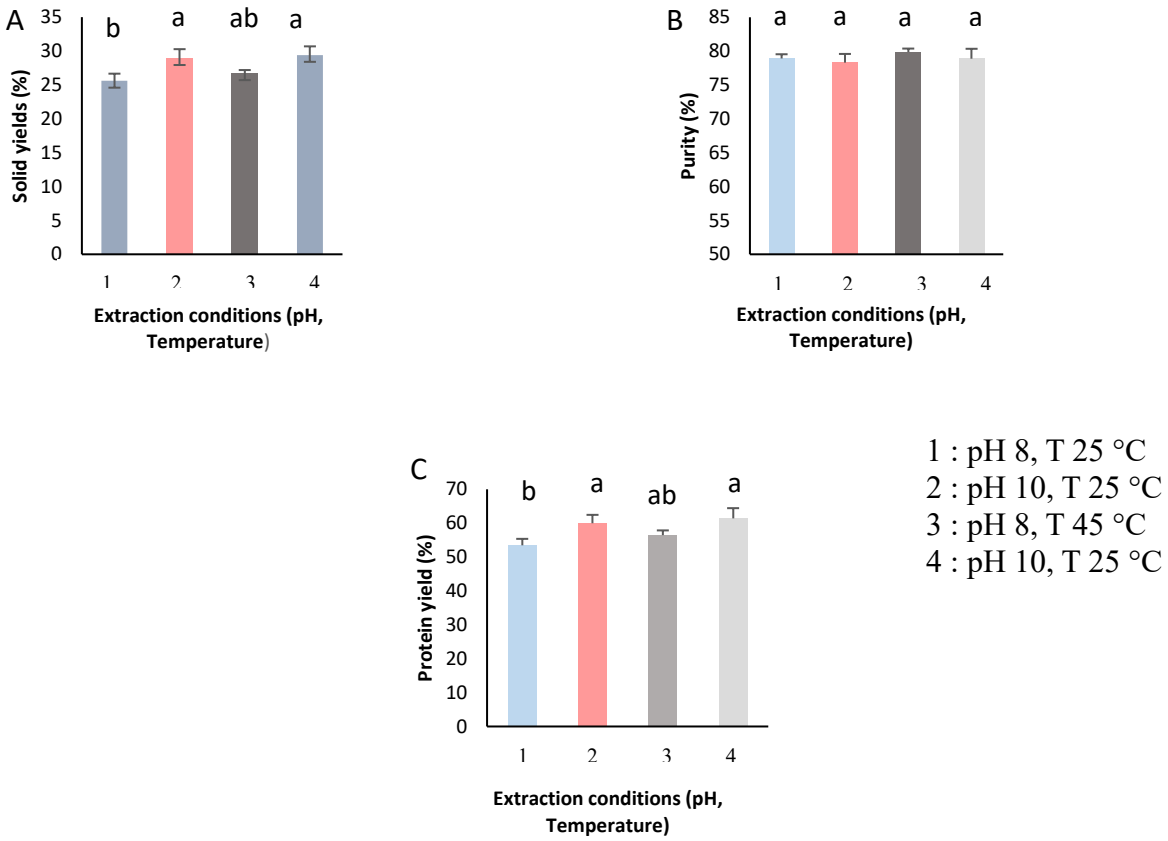

Figure 4. Soybean protein properties after extraction with different extraction pH (8 or 10) and temperature (25 °C or 45 °C). A: Solids yield of the extracted protein, B: Purity of the extracted protein, C: Yield of the extracted protein

The findings indicated that fluctuations in temperature and pH levels during the extraction process did exert a statistically significant effect on the purity of the protein isolate. The outcomes produced by the model were consistent with the experimental results. The selected model demonstrated that variations in the pH of extraction would yield comparable purity levels but with a high margin of error (Figure 4).

3.1.2. Extraction using Deep eutectic solvent

- **Selection of factors for the optimization of protein extraction**

To assess N diffusion during protein extraction with different DES-water mixtures, different models were evaluated. The determination coefficient (R^2) indicates that both the RandomForest, and the polynomial model showed a good fit to the data.

The rise in the DES (%) from 0 to 60% correlates with a statistically significant increase in the total N content. This increase is partly attributable to the N derived from ChCl, which constitutes 10% of its composition (Figure 5), as Kjeldahl method measures the overall nitrogen levels (Prandi et al., 2019).

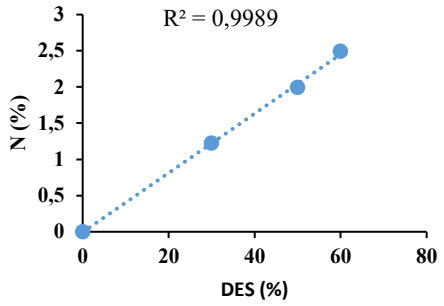

Figure 5. The N content of the DES:water solutions (0:100%, 30:70%,50:50%, and 60:40%). A strong linear relashionship between the percentage of DES and the N content as indicated by the coefficient of determination R^2

The density increased significantly by the addition of DES. It gives an idea about the solvation ability of DESs (Abdollahzadeh et al., 2022) (Figure 6).

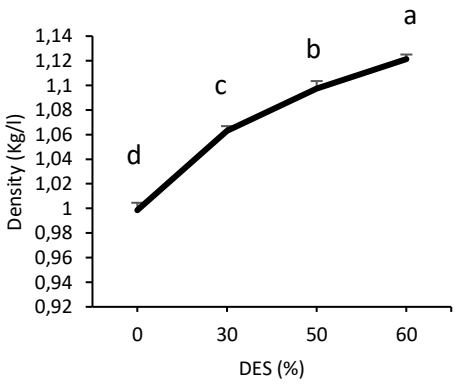

Figure 6. Density of the DES: water mixtures

With the addition of water, an adverse effect was obtained for the conductivity (Figure 7). As a result of the decrease in viscosity, it has been proved that a low amount of water may rise the conductivity by up to 100 times for some NADES.

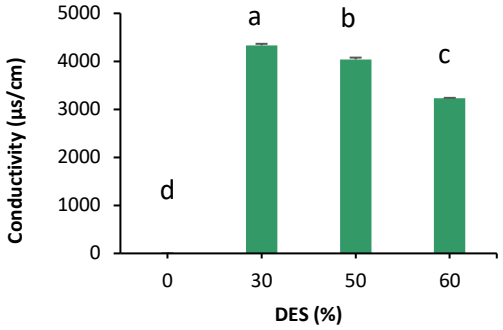

Figure 7. Conductivity of the DES:water mixtures

This result is further supported by the pH measurement of the used solvent (Figure 8), proving its neutrality.

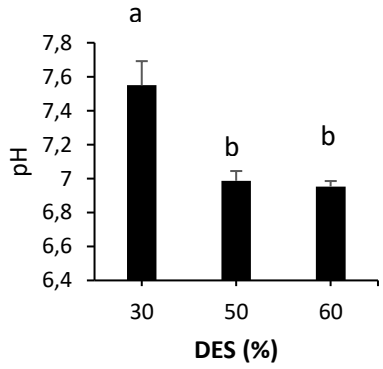

Figure 8. pH of the DES:water mixtures

3.2. Protein extract characterization

- **Protein extracted with DES-water solvent**

The SDS page analysis of both protein isolates showed that the protein extraction using a DES: water ratio (60%:40%) represents α′, α, and β subunits of 7S, and the acid subunits of the 11S glycinin fraction.

On the other hand, the DES extraction during heating resulted in an irreversible denaturation of soy protein after the use of pure DES mixture Q. Chen et al., (2021) (Figure 9).

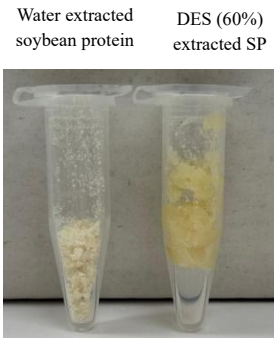

Figure 9. Deep eutectic solvent extracted soybean protein isolates and water extracted soybean protein

- **Treated protein isolate through physical and enzymatic process**

The electrophoretic analysis of both the modified and the native proteins revealed a broad range of molecular weights of ~ 17, 25, 34, 43, 55, 95, and 170 kDa. Following the enzymatic and physical treatment, no distinct red or blue shift occurred for the fluorescence spectrum of soybean protein isolate. However, the fluorescence intensity of the emitted light at 357 nm increased significantly (Figures 10). This effect can be related to the increase in the exposure of the amino acid residue (such as tryptophan and leucine) that were initially buried under the native state due to conformational changes (Li et al., 2020; Zhang et al., 2021), increasing the fluorescence intensity (Zhang et al., 2021).

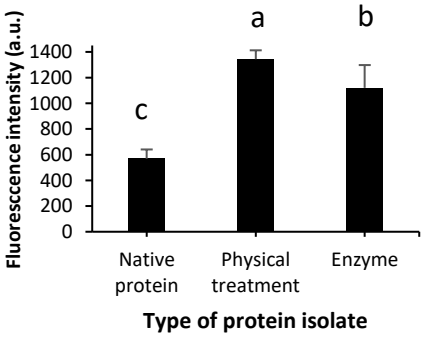

Figure 10. Fluorescence of native and modified soybean protein isolate (excitation:290 nm, emission: 357nm) Values with different superscripts differ significantly at $p < 0.05$.

The native soybean protein isolate has 80.61 ± 0.61% of protein, the rest which represent 20%. The oxidation of soybean oil generated a fluorescence with an excitation maximum at 365 nm and an emission maximum at 450 nm, and the intensity increased during storage (Liang, 1999). For these reasons, the reduction of the fluorescence intensity (Ex:360 nm /Em: 440 nm) of the treated samples (Figure 11) can be related to the removal of the oxidised oil by the addition of water and centrifugation.

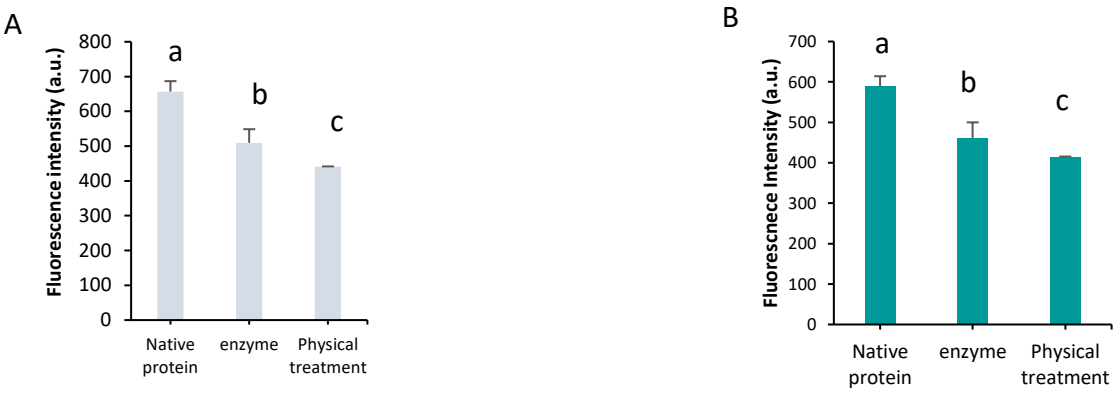

Figure 11. Fluorescence of native and modified soybean protein isolate A. Excitation:360 nm, emission: 450 nm, B. excitation:360 nm, emission: 440 nm
Value with different superscripts differ significantly at $p < 0.05$.

As shown in figure 11, the native soybean protein isolate has a higher fluorescence intensity than the treated ones at an excitation/emission wavelength of 360 nm/440 nm. This fluorescence can be generated by Maillard reaction products.

3.3. Assessment of nutritional effect of soybean protein and lysine

3.3.1. *Drosophila* model

- **Effect of the soybean protein and lysine on the development of *Drosophila***

The addition of 5-fold in sucrose to normal diet or 4% of soybean protein to zero nutrient diet (0M) showed 2-days of delay compared to normal media (NM) (Figure 12).

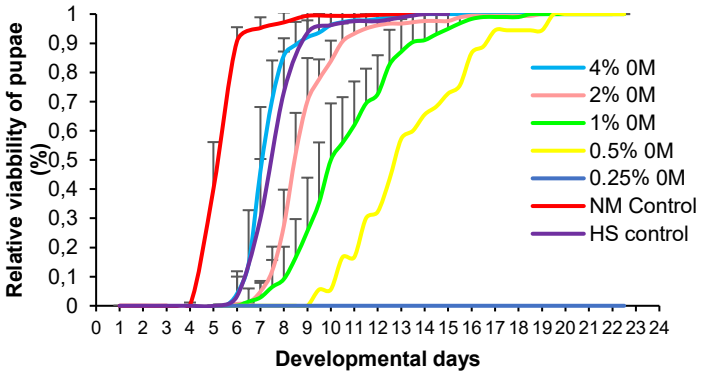

Figure 12. Progress of pupae's emergence after the addition of different percentages of SP in 0M over time

In 0M, 0.25 and 0.5% of Lysine, were able to make some embryos developed into larvae. However, these larvae could not pupate. In contrast, at high concentrations of lysine (1, 2, and 4%), no movement of larvae was detected. These findings indicate that lysine is necessary for the development of Drosophila. However, the low amount was sufficient, as higher quantities were toxic. These results emphasize the necessity of a balanced amino acid supply.

- **Effect of food restriction environment on pupae size**

In ideal conditions that promote rapid growth, internal checkpoints ensure that maturation does not occur until juvenile development is complete (Tennessen & Thummel, 2011). In particular, pupae size is one of the critical parameters for controlling growth and ensuring that the larvae attain the appropriate size at metamorphosis (Mirth & Riddiford, 2007).

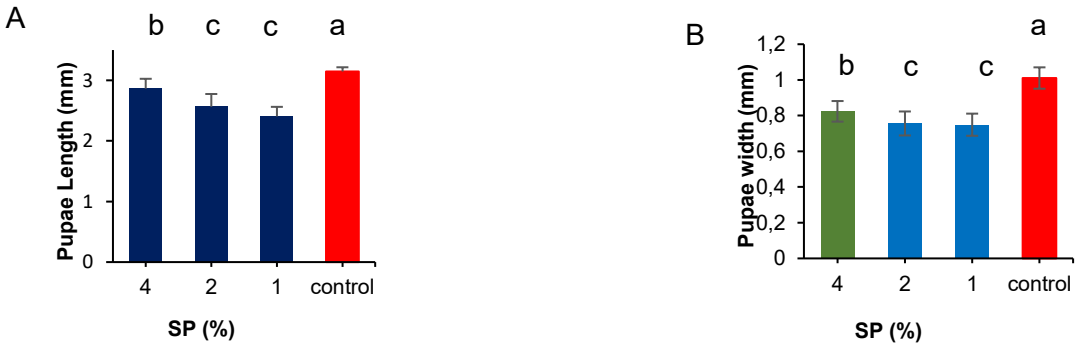

Figure 13. Pupal length (in mm) (A) and width (B) reared on different diets. Different letters denote significant differences based on estimated marginal mean comparisons (p value < 0.05). Control: NM

The analysis of the pupae size that emerged in OM with a variable SP concentration showed a significant difference in the average length and width (Figure 13). The largest pupae were found in NM (control), while the smallest pupae were observed in the 0M supplemented with (1 and 2% of SP). At low SP concentrations (0,5% and 0.25%), a limited number of pupae appeared, making it challenging to generalise the findings about their size.

- **Effect of diet on *Drosophila*'s eye color**

The visual assessment of the preliminary analysis of flies' eye color in the samples supplemented with SP isolate, assisted by the Drosopterin concentration, indicated a difference

in the eye color of the flies raised in NM and all those raised in 0M with different protein concentrations (1, 2, and 4 g/100 g 0M) was also identified (Figure 14).

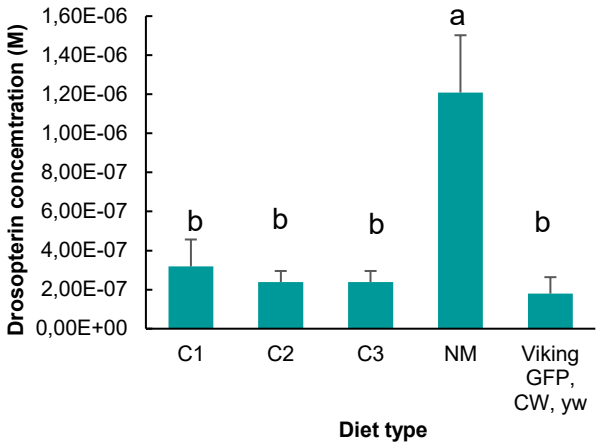

Figure 14. Drosopterin concentration of W^{m4h} *Drosophila* eyes reared in C1: 0M with 4% SP isolate, C2: 0M with 2% SP isolate, C3: 0M with 1% SP isolate, NM (Control), and of *Drosophila* mutation Viking GFP (Curly wings, yellow body) reared in NM

Our research revealed a variation in the shape of the brain (Figure 15). In particular, the length of the HS (hemispheres) of the larvae reared on a diet restricted to 2 and 4% of SP represents only 63 to 67% of the length of those grown under normal conditions.

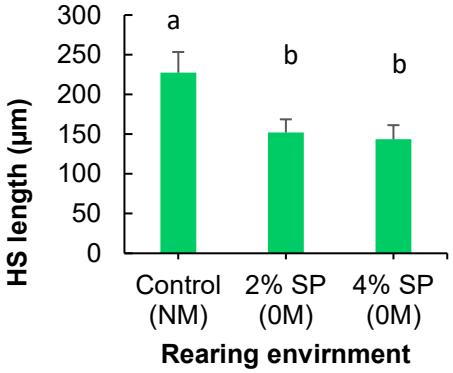

Figure 15. Length of HS of the brain of 3rd instar larvae cultivated in NM (control), 0M supplemented with 4% and 2% of SP. Different superscript denotes the significant difference at $p<0.05$

3.3.2. Allergenic and bioactive peptide characterisation

The comparison between the size of bioactive and allergic peptides reveals that there is a statistically significant difference between them. A smaller number of amino acids and a low molecular weight characterise the bioactive peptides compared to the allergic ones (241 amino acids).

The verification of their amino acid composition reveals that the peptides with allergic effect showed the dominance of Q (glutamine) with more than 11 % occurrence. However, bioactive peptides, showed a higher occurrence of E, P, G, K, S, and Q amino acids (>5%). Although they are characterised by their anti-inflammatory activities (Hasegawa et al., 2012; Nam et al., 2018), the ratio of cysteine in the allergic peptides is the highest, indicating their higher prevalence in allergic proteins (Figure 16).

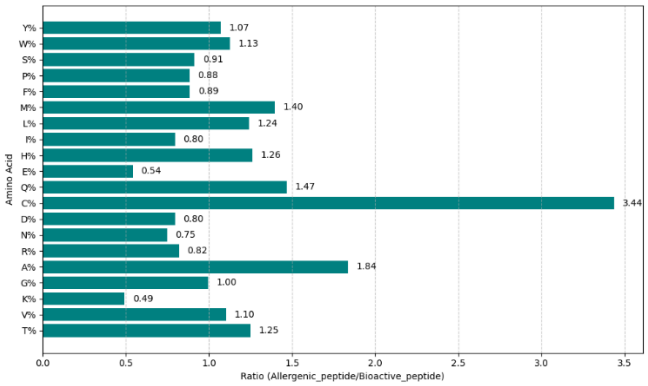

Figure 16. Average ratio of each amino acid in the allergic peptide / amino acid in the immunomodulatory peptide

The study of the correlation between all the features (amino acid count, aliphatic index, molecular weight, total hydrophobicity, the occurrence of thiol, E, and the amine group (K) and aromatic amino acids), and the nature of the peptide (allergenic or bioactive) referred to as its functional properties (Function) revealed a relatively high negative correlation between the molecular weight and the number of amino acids in the sequence.

In contrast there is a weak positive correlation between the prevalence of amine amino acid (K) and the response. However, the aromatic amino acids show almost no correlation with the type of the peptide, indicating their independence.

- **Epitopes detection from bioactive sequences**

In this research, the study of the physicochemical properties of the epitope that induces allergy and immunosuppressive effect (molecular weight, hydrophobicity, charge, amino acid count) and their composition of amino acids revealed a weak correlation between these properties and the functional property of the epitope (Function).

By detecting the pattern in the amino acid sequences of the bioactive peptides, a group of potential epitopes were identified. A similarity test (cosine similarity) using the main properties (the number of amino acids, the molecular weight, hydrophobicity, and the prevalence of the amino acid W%) was conducted to compare the identified epitopes with the immunosuppressive ones.

Using only W, as a parameter, the similarity revealed a notable difference between the sequences. However, featuring a group of parameters a high similarity between the potential epitope and the immunosuppressive epitope is unveiled. These results provide an overview on the potential use of these epitopes in treating autoimmune disorders, highlighting the potential role of structural similarity in predicting protein function. However, further experimental analysis is necessary to validate these findings.

4. NEW CONTRIBUTIONS TO ACADEMIC KNOWLEDGE

1. Data collection from previous scientific studies provided valuable data for determining optimal conditions for soy protein extraction, and the predictive model for protein purity showed a good fit with a high coefficient of determination (0.83). The experiment confirmed the adequacy of the model held under mild extraction conditions using a full factorial design, with temperature (25 and 45 °C) and pH of extraction (8 and 10). The conducted experiments confirmed that the temperature and pH of extraction did not influence the purity; however, they revealed that the pH of extraction significantly impacts the protein yield and solid yield. The highest yield was achieved at a temperature of 45 °C and a pH of 10.

   2. Adding different percentages of water to DES (Choline chloride: Glycerol (1:2)) and carrying out the extraction at different durations and mild temperatures (25-45 °C) may affect the proteinous nitrogen diffusion. Incorporating a high water content (~50-70%), which shifts toward a "DES-in-water" system, resulted in a nitrogen transfer similar to that of pure water. This might indicate a disruption of the specific H-bonded DES complex. The quality assessment of the protein extracted with 60% of DES, at 45 °C, revealed the potential exemption from impurities such as oxidized lipids or Maillard reaction conjugates, but it indicated the denaturation of the protein.

   3. Using the *Drosophila melanogaster* model through the indirect method (determination of the red eye pigment) reveals that a diet relying on just one single source of protein (soybean protein (1, 2, and 4%)) can induce epigenetic effects. On the other hand, the addition of lysine as a single amino acid to the normal diet of *Drosophila* to enhance the nutritional value of the diet can create a toxic effect at high concentrations (2% and 4%).

5. PRACTICAL USE OF THE RESULTS

1. The presented work highlights the importance of utilizing data from existing scientific research to advance process optimization. By compiling experimental results from diverse studies conducted under different conditions, this approach enables the development of data-driven models applicable to different aspects of the industry. This approach can minimize the necessity for extensive trials, offering a cost-effective and scientifically rigorous strategy.

2. Investigating the effect of water addition to DES is part of identifying strategies for enhancing the performance of the sustainable extraction process. From an industrial perspective, this method can generate higher efficiency corresponding to lower operational cost and higher profit.

3. This study highlights the importance of accurately determining protein intake to prevent nutrient deficiency and excessive consumption.

5. The toxicity of lysine for *Drosophila* when supplemented at a concentration of 2, and 4% in the normal diet can provide products for industrial applications, such as pesticides.

6. BIBLIOGRAPHY

Abdollahzadeh, M., Khosravi, M., Hajipour Khire Masjidi, B., Samimi Behbahan, A., Bagherzadeh, A., Shahkar, A., & Tat Shahdost, F. (2022). Estimating the density of deep eutectic solvents applying supervised machine learning techniques. *Scientific Reports*, *12*(1), 4954. https://doi.org/10.1038/s41598-022-08842-5

Aleya, A., Mihok, E., Pecsenye, B., Jolji, M., Kertész, A., Bársony, P., Vígh, S., Cziaky, Z., Máthé, A.-B., Burtescu, R. F., Oláh, N.-K., Neamțu, A.-A., Turcuș, V., & Máthé, E. (2023). Phytoconstituent Profiles Associated with Relevant Antioxidant Potential and Variable Nutritive Effects of the Olive, Sweet Almond, and Black Mulberry Gemmotherapy Extracts. *Antioxidants*, *12*(9), Article 9. https://doi.org/10.3390/antiox12091717

Bayala, E. X., Sinha, P., & Wittkopp, P. J. (2024). Protocol for dissecting *Drosophila* pupae and visualizing RNA expression using hybridization chain reaction. *STAR Protocols*, *5*(4), 103456. https://doi.org/10.1016/j.xpro.2024.103456

Chamba, M. V. M., Hua, Y., Murekatete, N., & Chen, Y. (2015). Effects of synthetic and natural extraction chemicals on yield, composition and protein quality of soy protein isolates extracted from full-fat and defatted flours. *Journal of Food Science and Technology*, *52*(2), 1016–1023. https://doi.org/10.1007/s13197-013-1084-x

Chen, Q., Chaihu, L., Yao, X., Cao, X., Bi, W., Lin, J., & Chen, D. D. Y. (2021a). Molecular Property-Tailored Soy Protein Extraction Process Using a Deep Eutectic Solvent. *ACS Sustainable Chemistry & Engineering*, *9*(30), 10083–10092. https://doi.org/10.1021/acssuschemeng.1c01848

Chiodza, K., & Goosen, N. J. (2024). Influence of mixing speed, solids concentration and enzyme dosage on dry solids yield and protein recovery during enzymatic hydrolysis of sardine (Sardina pilchardus) processing by-products using Alcalase 2.4L: A multivariable optimisation approach. *Biomass Conversion and Biorefinery*, *14*(22), 29045–29067. https://doi.org/10.1007/s13399-023-03829-2

Domokos-Szabolcsy, É., Elhawat, N., Domingos, G. J., Kovács, Z., Koroknai, J., Bodó, E., Fári, M. G., Alshaal, T., & Bákonyi, N. (2022). Comparison of Wet Fractionation Methods for Processing Broccoli Agricultural Wastes and Evaluation of the Nutri-Chemical Values of Obtained Products. *Foods*, *11*(16), Article 16. https://doi.org/10.3390/foods11162418

Enriquez, T., Lievens, V., Nieberding, C. M., & Visser, B. (2022). Pupal size as a proxy for fat content in laboratory-reared and field-collected Drosophila species. *Scientific Reports*, *12*(1), 12855. https://doi.org/10.1038/s41598-022-15325-0

Evers, J. M., & Hughes, C. G. (2002). ANALYSIS | Chemical Analysis. In H. Roginski (Ed.), *Encyclopedia of Dairy Sciences* (pp. 34–40). Elsevier. https://doi.org/10.1016/B0-12-227235-8/00015-8

FAOSTAT. (2025). https://www.fao.org/faostat/en/#data/QCL

Garcés-RimónD, M., Morales, D., & Miguel-Castro, M. (2022). Potential Role of Bioactive Proteins and Peptides Derived from Legumes towards Metabolic Syndrome. *Nutrients*, *14*(24). https://doi.org/10.3390/nu14245271

Goldstein, N., & Reifen, R. (2022). The potential of legume-derived proteins in the food industry. *Grain & Oil Science and Technology*, *5*(4), 167–178. https://doi.org/10.1016/j.gaost.2022.06.002

Guang, W., Baraldo, M., & Furlanut, M. (1995). Calculating percentage prediction error: A user's note. *Pharmacological Research*, *32*(4), 241–248. https://doi.org/10.1016/S1043-6618(05)80029-5

Hadidi, M., Aghababaei, F., & McClements, D. J. (2023). Enhanced alkaline extraction techniques for isolating and modifying plant-based proteins. *Food Hydrocolloids*, *145*(Journal Article), 109132. https://doi.org/10.1016/j.foodhyd.2023.109132

Hasegawa, S., Ichiyama, T., Sonaka, I., Ohsaki, A., Okada, S., Wakiguchi, H., Kudo, K., Kittaka, S., Hara, M., & Furukawa, S. (2012). Cysteine, histidine and glycine exhibit anti-inflammatory effects in human coronary arterial endothelial cells. *Clinical and Experimental Immunology*, *167*(2), 269–274. https://doi.org/10.1111/j.1365-2249.2011.04519.x

Ikai, A. (1980). Thermostability and Aliphatic Index of Globular Proteins. *Journal of Biochemistry (Tokyo)*, *88*(6), 1895–1898. https://doi.org/10.1093/oxfordjournals.jbchem.a133168

ImageJ. (n.d.). Retrieved December 8, 2025, from https://imagej.net/ij/

Johansson, M., Johansson, D., Ström, A., Rydén, J., Nilsson, K., Karlsson, J., Moriana, R., & Langton, M. (2022). Effect of starch and fibre on faba bean protein gel characteristics. *Food Hydrocolloids*, *131*(Journal Article), 107741. https://doi.org/10.1016/j.foodhyd.2022.107741

Koehler, J., Woetzel, N., Staritzbichler, R., Sanders, C. R., & Meiler, J. (2009). A Unified Hydrophobicity Scale for Multi-Span Membrane Proteins. *Proteins, Structure, Function, and Bioinformatics*, *76*(1), 13–29. https://doi.org/10.1002/prot.22315

Latimer, G. W., Jr. (Ed.). (2023). AOAC Official Method 2001.11Protein (Crude) in Animal Feed, Forage (Plant Tissue), Grain, and Oilseeds: Block Digestion Method Using Copper Catalyst and Steam Distillation into Boric Acid. In *Official Methods of Analysis of AOAC INTERNATIONAL* (p. 0). Oxford University Press. https://doi.org/10.1093/9780197610145.003.1403

Li, X., Chen, L., Hua, Y., Chen, Y., Kong, X., & Zhang, C. (2020). Effect of preheating-induced denaturation during protein production on the structure and gelling properties of soybean proteins. *Food Hydrocolloids*, *105*(Journal Article), 105846. https://doi.org/10.1016/j.foodhyd.2020.105846

Liang, J.-H. (1999). Fluorescence due to interactions of oxidizing soybean oil and soy proteins. *Food Chemistry*, *66*(1), 103–108. https://doi.org/10.1016/S0308-8146(98)00250-7

Liu, F., & Tang, C.-H. (2013). Soy Protein Nanoparticle Aggregates as Pickering Stabilizers for Oil-in-Water Emulsions. *Journal of Agricultural and Food Chemistry*, *61*(37), 8888–8898. https://doi.org/10.1021/jf401859y

Malaguti, M., Dinelli, G., Leoncini, E., Bregola, V., Bosi, S., Cicero, A. F. G., & Hrelia, S. (2014). Bioactive Peptides in Cereals and Legumes: Agronomical, Biochemical and Clinical Aspects. *International Journal of Molecular Sciences*, *15*(11), Article 11. https://doi.org/10.3390/ijms151121120

Mirth, C. K., & Riddiford, L. M. (2007). Size assessment and growth control: How adult size is determined in insects. *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology*, *29*(4), 344–355. https://doi.org/10.1002/bies.20552

Nam, S.-Y., Kim, H.-M., & Jeong, H.-J. (2018). Cysteine ameliorates allergic inflammatory reactions by suppressing thymic stromal lymphopoietin production in activated human mast cells. *Nutrition Research (New York, N.Y.)*, *49*(49), 79–87. https://doi.org/10.1016/j.nutres.2017.11.005

Patra, A., Prasath, V. A., & Pandiselvam, R. (2023). Deep eutectic solvent: An emerging trend for extraction of plant proteins. *Journal of Molecular Liquids*, *389*, 122887. https://doi.org/10.1016/j.molliq.2023.122887

Peng, X., Ren, C., & Guo, S. (2016). Particle formation and gelation of soymilk: Effect of heat. *Trends in Food Science & Technology*, *54*, 138–147. https://doi.org/10.1016/j.tifs.2016.06.005

Prandi, B., Faccini, A., Lambertini, F., Bencivenni, M., Jorba, M., Van Droogenbroek, B., Bruggeman, G., Schöber, J., Petrusan, J., Elst, K., & Sforza, S. (2019). Food wastes from agrifood industry as possible sources of proteins: A detailed molecular view on the composition of the nitrogen fraction, amino acid profile and racemisation degree of 39 food waste streams. *Food Chemistry*, *286*, 567–575. https://doi.org/10.1016/j.foodchem.2019.01.166

SDAP 2.0: Structural Database of Allergenic Proteins. (2023). https://fermi.utmb.edu/

Semba, R. D., Ramsing, R., Rahman, N., Kraemer, K., & Bloem, M. W. (2021). Legumes as a sustainable source of protein in human diets. *Global Food Security*, *28*, 100520. https://doi.org/10.1016/j.gfs.2021.100520

Shevkani, K. (2023). Chapter 3—Protein from land—Legumes and pulses. In B. K. Tiwari & L. E. Healy (Eds.), *Future Proteins* (pp. 35–68). Academic Press. https://doi.org/10.1016/B978-0-323-91739-1.00003-9

Sun, H., Merugu, S., Gu, X., Kang, Y. Y., Dickinson, D. P., Callaerts, P., & Li, W.-H. (2002). Identification of Essential Amino Acid Changes in Paired Domain Evolution Using a Novel Combination of Evolutionary Analysis and In Vitro and In Vivo Studies. *Molecular Biology and Evolution*, *19*(9), 1490–1500. https://doi.org/10.1093/oxfordjournals.molbev.a004212

Tennessen, J. M., & Thummel, C. S. (2011). Coordinating Growth and Maturation—Insights from *Drosophila*. *Current Biology*, *21*(18), R750–R757. https://doi.org/10.1016/j.cub.2011.06.033

Wang, Z., Li, Y., Jiang, L., Qi, B., & Zhou, L. (2014). Relationship between Secondary Structure and Surface Hydrophobicity of Soybean Protein Isolate Subjected to Heat Treatment. *Journal of Chemistry*, *2014*(Journal Article), 1–10. https://doi.org/10.1155/2014/475389

Xiao, C.-W., Hendry, A., Kenney, L., & Bertinato, J. (2023). L-Lysine supplementation affects dietary protein quality and growth and serum amino acid concentrations in rats. *Scientific Reports*, *13*(1), 19943. https://doi.org/10.1038/s41598-023-47321-3

Yue, J., Zhu, Z., Yi, J., Lan, Y., Chen, B., & Rao, J. (2021). Structure and functionality of oat protein extracted by choline chloride–dihydric alcohol deep eutectic solvent and its

water binary mixtures. *Food Hydrocolloids*, *112*, 106330. https://doi.org/10.1016/j.foodhyd.2020.106330

Zhang, A., Cui, Q., Wang, X., & Zhao, X. (2021). Effect of temperature of preheated soy protein isolate on the structure and properties of soy protein isolate heated–vitamin D3 complex. *Journal of Food Biochemistry*, *45*(6), e13733-n/a. https://doi.org/10.1111/jfbc.13733

7. PUBLICATIONS IN THE FIELD OF RESEARCH

UNIVERSITY of DEBRECEN

UNIVERSITY AND NATIONAL LIBRARY
UNIVERSITY OF DEBRECEN
H-4002 Egyetem tér 1, Debrecen
Phone: +3652/410-443, email: publikaciok@lib.unideb.hu

Registry number: DEENK/2/2026.PL
Subject: PhD Publication List

Candidate: Chaima Neji
Doctoral School: Doctoral School of Nutrition and Food Sciences, Doctoral Programme of Food Sciences
MTMT ID: 10087077

List of publications related to the dissertation

Foreign language scientific articles in international journals (5)

1. **Neji, C.**, Muthu, A., Huzsvai, L., Ungai, D., Seres, E., Prokisch, J., Máthé, E., Sipos, P.: Assessing data to optimize soybean protein extraction.
   LWT-Food Sci. Technol. 217, 1-12, 2025. ISSN: 0023-6438.
   DOI: http://dx.doi.org/10.1016/j.lwt.2025.117401
   IF: 6.6 (2024)

2. **Neji, C.**, Muthu, A., Ungai, D., Seres, E., Domokos-Szabolcsy, É., Prokisch, J., Máthé, E., Sipos, P.: Assessing Deep Eutectic Solvent Efficiency by Optimizing Soybean Protein Extraction.
   Food Bioprocess Technol. 18 (11), 9423-9437, 2025. ISSN: 1935-5130.
   DOI: http://dx.doi.org/10.1007/s11947-025-03973-9
   IF: 5.8 (2024)

3. Kamani, M. H., **Neji, C.**, Fitzsimons, S. M., Fenelon, M. A., Murphy, E. G.: Unlocking the nutritional and functional potential of legume waste to produce protein ingredients.
   Crit. Rev. Food Sci. Nutr. 64 (21), 7311-7329, 2024. ISSN: 1040-8398.
   DOI: http://dx.doi.org/10.1080/10408398.2023.2184322
   IF: 8.8

4. **Neji, C.**, Semwal, J., Máthé, E., Sipos, P.: Dough Rheological Properties and Macronutrient Bioavailability of Cereal Products Fortified through Legume Proteins.
   Processes. 11 (2), 1-16, 2023. EISSN: 2227-9717.
   DOI: http://dx.doi.org/10.3390/pr11020417
   IF: 2.8

5. **Neji, C.**, Semwal, J., Kamani, M. H., Máthé, E., Sipos, P.: Legume Protein Extracts: The Relevance of Physical Processing in the Context of Structural, Techno-Functional and Nutritional Aspects of Food Development.
   Processes. 10 (12), 1-27, 2022. EISSN: 2227-9717.
   DOI: http://dx.doi.org/10.3390/pr10122586
   IF: 3.5

UNIVERSITY AND NATIONAL LIBRARY
UNIVERSITY OF DEBRECEN
H-4002 Egyetem tér 1, Debrecen
Phone: +3652/410-443, email: publikaciok@lib.unideb.hu

List of other publications

Foreign language scientific articles in international journals (5)

6. Muthu, A., Nguyen, H. H. D., Ferroudj, A., Prokisch, J., El-Ramady, H., **Neji, C.**, Béni, Á.: Green Synthesised Carbon Nanodots Using the Maillard Reaction for the Rapid Detection of Elemental Selenium in Water and Carbonated Beverages.
Nanomaterials. 15 (15), 1-16, 2025. ISSN: 2079-4991.
DOI: http://dx.doi.org/10.3390/nano15151161
IF: 4.3 (2024)

7. Muthu, A., Nguyen, H. H. D., **Neji, C.**, El-Ramady, H., Prokisch, J., Béni, Á.: Nanohybrid-based fluorescent "turn-on" system for selective detection of mercury (II) in aqueous solution.
J. Food Compos. Anal. 148, 1-10, 2025. ISSN: 0889-1575.
DOI: http://dx.doi.org/10.1016/j.jfca.2025.108653
IF: 4.6 (2024)

8. Muthu, A., Nguyen, H. H. D., **Neji, C.**, Seresné Törős, G., Ferroudj, A., Atieh, R., Prokisch, J., El-Ramady, H., Béni, Á.: Nanomaterials for Smart and Sustainable Food Packaging: Nano-Sensing Mechanisms, and Regulatory Perspectives.
Foods. 14 (15), 2-29, 2025. EISSN: 2304-8158.
DOI: https://doi.org/10.3390/foods14152657
IF: 5.1 (2024)

9. Prokisch, J., Seresné Törős, G., Nguyen, H. H. D., **Neji, C.**, Ferroudj, A., Sári, D., Muthu, A., Brevik, E. C., El-Ramady, H.: Nano-Food Farming: Toward Sustainable Applications of Proteins, Mushrooms, Nano-Nutrients, and Nanofibers.
Agron. J. 14, 1-30, 2024. ISSN: 0002-1962.
DOI: http://dx.doi.org/https://doi.org/10.3390/agronomy14030606
IF: 2

10. Nagy, R., **Neji, C.**, Gálné Remenyik, J., Sipos, P.: Evaluation of functional properties of physically treated Sorghum flours and development of gluten-free Sorghum breads.
J. Microbiol. Biotech. Food Sci. 13 (3), 1-6, 2023. ISSN: 1338-5178.
DOI: http://dx.doi.org/10.55251/jmbfs.9959
IF: 0.6

Foreign language abstracts (3)

11. Muthu, A., **Neji, C.**, Nguyen, H. H. D., Prokisch, J., Béni, Á.: Application of Machine learning for predicting survival outcomes of nano selenium dosage in different in-vivo and in-vitro models.
In: Reshaping Global Food Systems / Dongxiao Sun-Waterhouse, Paul Finglas, Sian Astley, Nandika Bandara, Elsevier, Glasgow, 21, 2025.

UNIVERSITY of DEBRECEN

UNIVERSITY AND NATIONAL LIBRARY
UNIVERSITY OF DEBRECEN
H-4002 Egyetem tér 1, Debrecen
Phone: +3652/410-443, email: publikaciok@lib.unideb.hu

12. **Neji, C.**, Sipos, P.: Optimisation of soybean protein extracted from binary mixtures of choline chloride and glycerol.
In: Abstract book. Ed.: Cristina L.M. Silva, Teresa R.S. Brand, The European Federation of Food Science and Technology, Portugal, 602, 2025.
13. **Neji, C.**, Ungai, D., Seres, E., Sipos, P.: Effect of nitrogen fertilization on the quality of soybean flour and protein isolate.
In: Book of abstracts : VIII International Conference Sustainable Postharvest and Food Technologies - Inoptep 2023 and XXXV. Scientific-Professional Conference Processing and Energy in Agriculture - PTEP 2023. Eds.: Filip Kulić; Ivan Pavkov, National Society of Processing and Energy in Agriculture, Novi Sad, 88, 2023. ISBN: 9788675205814

Total IF of journals (all publications): 44,1
Total IF of journals (publications related to the dissertation): 27,5

The Candidate's publication data submitted to the Tudóstér have been validated by DEENK on the basis of the Journal Citation Report (Impact Factor) database.

07 January, 2026