

UNIVERSITY OF DEBRECEN

DOCTORAL SCHOOL OF FOOD SCIENCES

Head of Doctoral School:

Prof. Dr. Béla Róbert Kovács, DSc professor

Supervisor(s):

**Prof. Dr. Péter Sipos, PhD
professor**

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

Prepared by:

**Chaima Neji
PhD candidate**

Debrecen

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

Dissertation submitted in partial fulfilment of the requirements for the doctoral (PhD)
degree in Food Sciences

Written by Chaima Neji certified Food industry engineer

Prepared in the framework of the doctoral school of Food Sciences of the University of
Debrecen
(Food Sciences programme)

Dissertation supervisor: Prof. Dr. Péter Sipos

The official opponents of the dissertation:

.....
.....

The evaluation board:

Chairperson:

.....

members:

.....
.....

The date and venue of the dissertation defence:

TABLE OF CONTENTS

1. INTRODUCTION	1
2. LITERATURE REVIEW	3
2.1. Protein extraction: Efficiency optimization and sustainable techniques.....	4
2.1.1. Overview of extraction methodologies.....	4
2.1.2. Extraction optimization methods and quality prediction	9
2.1.2.1. Response surface methodology	9
2.1.2.2. Machine learning for extraction performance prediction	9
2.2. Characterization of the protein isolate	10
2.2.1. Structural morphology characterization of protein	10
2.2.2. Molecular weight of the protein/polypeptide.....	11
2.2.3. Fluorescence spectroscopy for protein characterization	11
2.3. Effect of different treatments on legume protein extract on nutritional efficiency	12
2.4. Evaluation and nutritional effect of soybean protein	14
2.4.1. <i>Drosophila melanogaster</i> model for nutritional evaluation.....	14
2.5. Importance of legumes and cereal sequences on the immunological properties .	18
3. MATERIALS AND METHODS	21
3.1. Optimization of protein extraction	21
3.1.1. Materials	21
3.1.2. Methods	21
3.1.2.1. Extraction using data analysis and conventional method	21
3.1.2.2. Extraction using Deep Eutectic Solvent (DES)	25
3.1.2.2.1. Preparation of Deep Eutectic Solvents	25
3.1.2.2.2. Protein extraction using DES/water mixture	25
3.1.2.2.3. Optimization of protein extraction.....	26
3.2. Evaluation of protein quality	28
3.2.1. Sample preparation	28
3.2.1.1. Protein extracted with DES solvent.....	28
3.2.1.2. Protein isolate modified with enzymatic hydrolysis.....	28
3.2.1.3. Protein isolate subjected to physical modification	29
3.2.2. Evaluation analysis	29

3.2.2.1. SDS-PAGE analysis	29
3.2.2.2. Fluorescence spectroscopy	29
3.2.2.3. Light microscope	30
3.3. Assessment of the nutritional effect of soybean protein	30
3.3.1. <i>Drosophila melanogaster</i> as a model for nutritional value assessment.....	30
3.3.1.1. Description of <i>Drosophila melanogaster in vivo</i> analytical system	30
3.3.1.2. Pupae length and width measurements	31
3.3.1.3. Determination of <i>Drosophila</i> Red Eye Pigment Concentration.....	31
3.3.1.4. Brain dissection and optic lobe length	32
3.4. Assessment of immunological properties of cereal and legume protein sequences	33
3.4.1. Dataset preparation and evaluation for epitope characterization.....	33
3.4.2. Dataset preparation	33
3.5. Statistical analysis.....	34
4. RESULTS	35
4.1. Optimization of protein extraction	35
4.1.1. Extraction using data analysis and conventional method	35
4.1.1.1. Understanding the data	35
4.1.1.2. Relation between predictors and dependent variables	35
4.1.1.3. Optimization of soybean protein isolate extraction conditions	36
4.1.1.3.1. Model fitting	36
4.1.1.3.2. Validation of protein extraction process	38
4.2. Extraction using DES.....	40
4.2.1. Selection of factors for the optimization of protein extraction.....	40
4.2.2. Model validation	45
4.2.3. Characterization of the extracted protein	47
4.2.3.1. Protein extracted using the DES-water solvent	47
4.2.3.2. Characterization of the treated protein isolate	53
4.3. Assessment of the nutritional effect of soybean protein and lysine.....	61
4.3.1. <i>Drosophila</i> model	61
4.3.1.1. Effect of the soybean protein and lysine on the development of <i>Drosophila</i>	61
4.3.1.2. Effect of diet on <i>Drosophila</i> 's eye color	70
4.3.1.3. Effect of the protein-restricted diet on the brain size.....	73
4.3.2. Allergenic and bioactive peptide characterization	77

4.3.2.1. Size of the peptide.....	77
4.3.2.2. Relationship between peptide characteristics and functional properties ...	85
4.3.2.3. Epitopes detection from bioactive sequences	86
5. CONCLUSIONS, RECOMMENDATIONS	91
6. NEW SCIENTIFIC RESULTS	92
7. PRACTICAL RESULTS	93
8. SUMMARY	94
9. BIBLIOGRAPHY	96
10. PUBLICATIONS IN THE FIELD OF RESEARCH.....	123
11. STATEMENTS	126
12. ANNEXES	127
13. ACKNOWLEDGEMENTS	140

LIST OF ABBREVIATIONS

11S: Glycinin

7S: β -conglycinin

0M: Zero media

A, Ala: Alanine

AA: Amino acids

AIC: Akaike's information criterion

A-IP: Alkaline isoelectric precipitation

BIC: Bayesian information criterion

C: Cysteine

CB: Central brain.

ChCl: Choline chloride

D or Asp: Aspartic acid

DES: Deep Eutectic Solvent

DOE: Design of Experiments

E or Glu: Glutamic acid

EAA: Essential amino acid

F or Phe: Phenylalanine

FAOSTAT: The Food and Agriculture Organization Corporate Statistical Database

FPFE: Polytetrafluoroethylene membrane filter

FTIR: Fourier transform infrared spectroscopy

G or Gly: Glycine

Gly: Glycerol

H or His: Histidine

HBA: Hydrogen bond acceptors

HBD: Hydrogen bond donors

HE: Hemisphere

HS: High Sugar diet

I or Ile: Isoleucine

IEDB-AR: The Immune Epitope Database

IEF: Isoelectric focusing

K or Lys: Lysine

L or Leu: Leucine

M or Met: Methionine

MS: Mass Spectrometry

N or Asn: Asparagine

N-EAA: Non-essential amino acid

NaCl: Sodium Chloride

NM: Normal diet

NMR: Nuclear Magnetic Resonance

P or Pro: Proline

PAGE: Polyacrylamide Gel Electrophoresis

PCA Principal Component Analysis

PEV: Position Effect Variegation

Q or Gln: Glutamine

R or Arg: Arginine

R^2 : Coefficient of determination

RCF: Relative centrifugal force

RMSE: Root Mean Squared Error

RPM: Revolutions per min

T: Temperature

T or Thr: Threonine

S or Ser: Serine

SEC-MALS: Size Exclusion Chromatography with Multiangle scattering

SDS-PAGE: Sodium dodecyl sulfate-polyacrylamide gel electrophoresis

SP: Soybean protein

V or Val: Valine

VNC: Ventral nerve cord

Vkg-GFP: Viking GFP

W or Trp: Tryptophan

w^{m4h}: white mottled 4 strain

Y or Tyr: Tyrosine

1. INTRODUCTION

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. Plant proteins are gaining continuous interest owing to their significance as sustainable food ingredients and their suitability for developing products such as meat analogues or supplements for bodybuilding and weight control.

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 (DES)), 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 characterization of these peptides may reveal promising new functional properties. Additionally, an *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, nutritional, and immunological aspects related to the plant origin protein extracts (Figure 1). The main objectives were to:

1. Apply a predictive model for estimating the purity of soybean protein extracted using the conventional method (Alkaline isoelectric precipitation (A-IP)), and the new method (DES).
2. Assess the physicochemical properties of the extracted soybean protein using techniques to elucidate its molecular weight through Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS PAGE), to give an overview of the protein's conformation and intermolecular contacts (fluorescence spectroscopy), and to visualise the microstructure (microscopy).
3. Elucidate the potential nutritional effect (viability, Position-effect variegation (PEV), organ size) of a normal and an exclusive high-protein diet (soybean protein) using the *in vivo* model *Drosophila melanogaster*. This model is characterized by its ease of maintenance in the laboratory, its fully mapped genome, and the similarity or equivalence of its genes to those in humans.

Analyse the primary sequences of plant proteins of different origins (cereals and legumes) possessing antagonist effect (allergenic and bioactive effect) using bioinformatics tools to derive insights and identify conserved motifs or characteristic patterns.

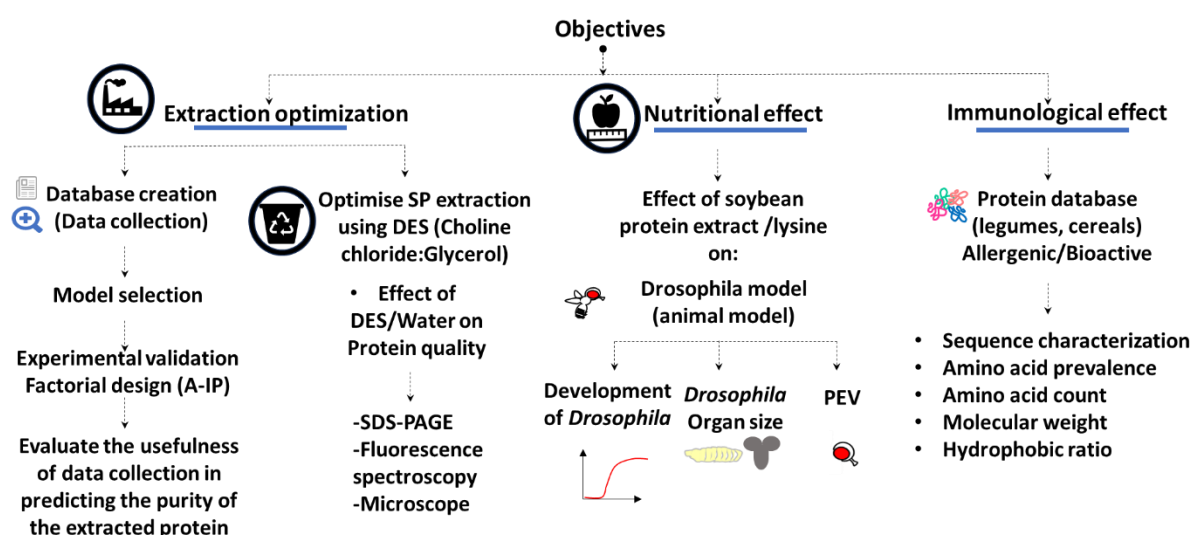


Figure 1. Study objectives

2. LITERATURE REVIEW

Nowadays, there is a surge in people's awareness of the importance of a protein-rich daily diet (Patra, Prasath, et al., 2023). Therefore, to align with the expanding nutrient demands of a growing global population, affordable, abundant, sustainable sources of protein are needed (Hadidi et al., 2024). Plant protein is recognized for being an eco-friendly, plentiful, and customizable plant-based food that can serve as a substitute for animal-based foods. Nevertheless, currently, available plant-based foods lack nutrients exclusive to animal food products (J. Huang et al., 2024).

From the environmental, agricultural, and nutritional perspectives, legumes offer several advantages. 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). While most legumes contain 20%–25% of protein, lupines and soybeans have exceptionally high protein content (Shevkani, 2023). In particular, soybean is the most widely produced legume crop (Semba et al., 2021). In 2023, the cultivated land area extended to nearly 136.9 million hectares, resulting in a harvest of 371.17 million tons (FAOSTAT, 2025). This production quantity is projected to increase shortly, alongside a growing demand driven by the ever-increasing population (Dilawari et al., 2022). Despite being one of the most cultivated genetically modified crops (Sieradzki et al., 2021) soybeans represent the most widely used source of protein, containing at least 30 allergens, such as Gly m 1 to Gly m 8, P28, and P34, which affect 0.5% of the general population (Pi et al., 2021). Being a worldwide cultivated crop with high-quality protein and edible oil (Giorgio et al., 2023), soybean (*Glycine max L.*) is considered a basic source of vegetable oil and a substitute for high-protein meat (Dilawari et al., 2022). However, their use as protein extracts extends their application in many other fields. In particular, these eco-friendly protein isolates can provide a high-quality food source with a well-balanced amino acid composition (Z. Huang et al., 2023). These proteins can be used as an ingredient in multiple food systems (bakery, breakfast cereals, beverages, infant formulas, enteral nutritionals, dairy) (Gilani & Lee, 2003) and non-food systems such as biodegradable films and gel systems (Meganaharshini et al., 2023).

2.1. Protein extraction: Efficiency optimization and sustainable techniques

2.1.1. Overview of extraction methodologies

During protein extraction, assessing extraction efficiency by evaluating the purity of the extract and the degree of protein separation (protein recovery or yield) is the primary indicator of process performance. Regarding protein purity, the denomination of protein concentrates or isolates depends on the protein content, with concentrates classified when the protein level reaches 65% and isolates when it reaches 90% (Boye et al., 2010; I. M. Rodrigues et al., 2012).

Whether derived from seeds or waste, the effectiveness of plant protein extraction methods depends mainly on the adopted methods (Hewage et al., 2022). In addition to the properties of protein (Meganaharshini et al., 2023), the quality and recovery yield of plant proteins depend on the extraction methods (Tang et al., 2024). For instance, wet extraction yields pure protein isolates, but it is less sustainable due to higher energy consumption and increased waste. Conversely, dry separation prioritizes the functioning of ingredients over molecular purity (Fatima et al., 2024). Together, dry and wet fractionation methods represent the conventional approaches for protein extraction (Eze et al., 2022).

Protein extraction using conventional methods can be applicable to a diverse category of plant matrices, encompassing oilseeds, cereals, pulses, and agro-residues (M. Kumar et al., 2021). However, they are known to be time-consuming, energy-intensive, and ecologically unsustainable (M. Kumar et al., 2021; Tang et al., 2024). These methods include wet extraction methods (alkaline, acid, and salt extraction) and dry fractionation, commonly utilised for plant protein extraction (Tang et al., 2024). Among these techniques, alkaline extraction is the most utilised (Momen et al., 2021)(Figure 2).

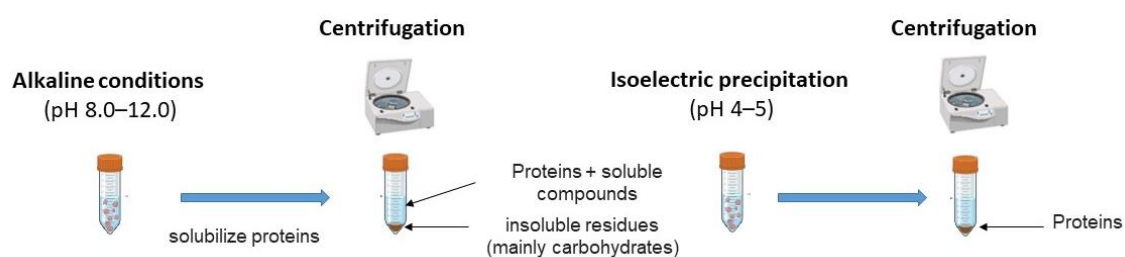


Figure 2. Alkaline conditions/Isoelectric precipitation

Since these conventional techniques for protein extraction are not sustainable, it is imperative to explore alternative, more environmentally friendly solutions (Hanafi et al., 2024). To comply with these criteria to be qualified as a green medium, criteria including accessibility, low toxicity, biodegradability, and cost effectiveness should be satisfied (Santos et al., 2022). In pursuit of improved extraction effectiveness, innovative extraction technologies have emerged (Figure 3) (Shawky et al., 2025).

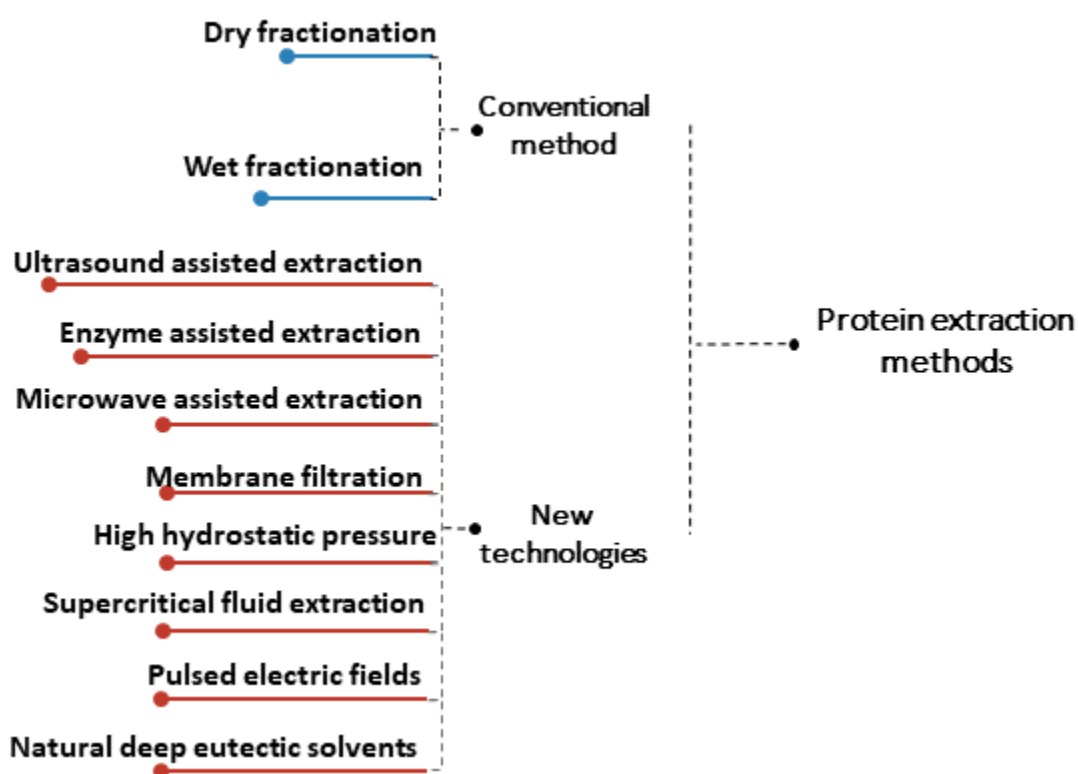


Figure 3. Protein extraction methods from legumes

The advantages and disadvantages of some of these new protein extraction methods are summarized in Table 1.

Table 1. Advantages and disadvantages of new technologies for protein extraction

Extraction method	Advantages	Disadvantages	References
Alkali-based extraction	-Low cost -Enhanced digestibility and bioavailability	-Reduced yield of extraction, -Reduced protein quality through denaturation, -Loss of amino acids, -Time-consuming, energy-intensive, -Non-eco-friendly	(Ampofo & Ngadi, 2022)

Microwave-assisted	-Environmentally friendly -Penetration power -Higher extraction rate -Rapid extraction -High-efficiency tool -High-power energy -Reducing process time -Low solvent volume	-Expensive equipment -Can cause denaturation -Higher extraction time -Formation of hot spots -Dielectric properties of the solid material to absorb microwaves -Clean-up step needed - Waiting time for the vessels to cool down	(Hadidi et al., 2023) (Anyiam et al., 2025)
Ultrasound assisted extraction	-Eco-innovative technique, green technology -Short processing -Reduce process time -Enhance extraction -Non-toxic -Low cost -Limited sensory and nutritional losses	-Modify the physical, structural, and functional properties, especially hydrophobicity, solubility, emulsion, and foaming. -High energy consumption	(Ampofo & Ngadi, 2022) (Rahman & Lamsal, 2021) (Anyiam et al., 2025)
Pulsed electric fields	-Low energy consumption -Short processing time, -Enhancing the nutritional quality of food products by altering the structure and function of anti-nutritional factors.	-High cost, -Protein digestibility of PEF-treated food products is yet to be thoroughly explored.	(S. Zhang et al., 2021) (Kheto et al., 2024)(Navaf et al., 2023)
Enzyme-aided extraction	-Fast, easy to be carried out, and eco-friendly	-High cost of enzymes, - increases the processing costs (need to inactivate enzymes by heat treatment at the end of the enzymatic reaction)	(Momen et al., 2021) (Tong et al., 2025)
DES	-Nontoxicity (based on non-toxic components), Sustainability, -Easy accessibility to compounds,	Cytotoxicity assessments reveal that they occasionally have a higher level of toxicity than their individual component.	(Patra, Arun Prasath, et al., 2023) (Boateng, 2023)

	-Biodegradability/low impact on the environment, low cost, and ease of manufacture.		
High pressure-assisted	-Preserve protein functionality -Uniform treatment -Suitable for heat-sensitive proteins	-High equipment cost -Limited scalability -Protein denaturation -High energy consumption -Formation of unwanted by-products	(Anyiam et al., 2025) (Hadidi et al., 2023)

During protein extraction, the solvent is identified as a determining factor (Patra, Prasath, et al., 2023). Over the organic (such as alcohols and ketones, chloroform and toluene, formamide, pentane) and aqueous solvents (like alkaline-assisted aqueous extraction), DES emerged as a promising solvent to satisfy the environmental requirements while maintaining the protein quality (Patra, Prasath, et al., 2023). In particular, they present advantages compared to conventional alternatives in terms of extraction efficiency, environmental safety, and improved techno-functional properties (Hewage et al., 2022). The literal meaning of the term ‘eutectic’ originates from Greek (ευτηκτος), translated as ‘easy melting’. However, most of the literature coined the term ‘deep’ as those mixtures with a eutectic temperature far below that of an ideal liquid mixture (Ling & Hadinoto, 2022). These DESs have the specificity of being characterized by their significant depressions in melting points compared to those of the neat constituent components (Hansen et al., 2021). These solvents are not pure compounds (Martins et al., 2019), they are formed by combining hydrogen bond acceptors (HBA) and hydrogen bond donors (HBD) in specific molar ratios (Mannu et al., 2021). Basing on the type of mixed components, DESs are classified into five categories: type I (organic salt, metal salt), type II (organic salt, metal salt hydrate), type III (organic salt, molecular HBD), type IV (metal salt hydrate or metal salt, molecular HBD), and type V (molecular HBA, molecular HBD)(Souza et al., 2024). In particular, quaternary ammonium salts, especially choline chloride, are considered the HBA most widely used in DES formulations (Martins et al., 2019).

By extracting both soluble and insoluble proteins, DES provide a viable approach for valorising protein-rich material deriving from various sources (Hanafi et al., 2024).

Figure 4 shows the extraction steps. They were used to extract protein from diverse plant products, such as oat (Yue et al., 2021), soybean (Q. Chen et al., 2021), fava bean (Hewage et al., 2024), and byproducts such as brewer's spent grain (Wahlström et al., 2017), oilseed cakes (Grudniewska et al., 2018), sesame meal (Cao et al., 2023), tip of the bamboo shoot, and its processing byproducts (basal bamboo shoot and sheath) (Z. Lin et al., 2021).

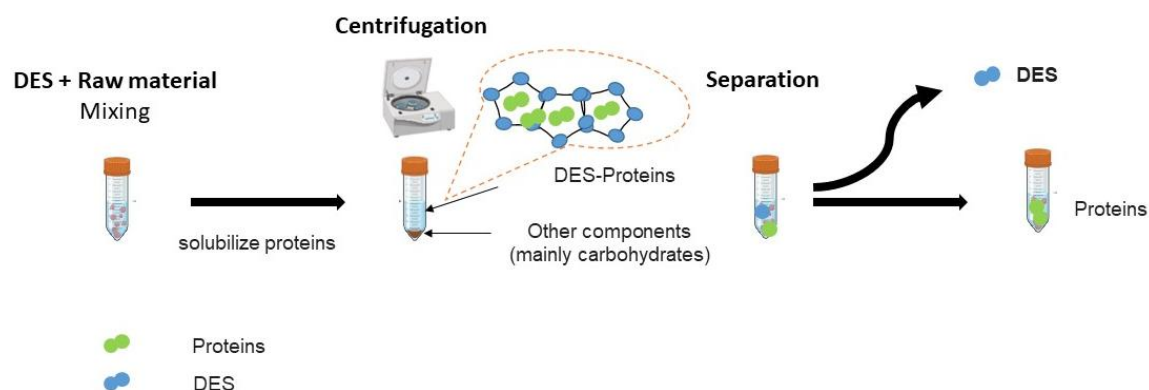


Figure 4. Solid-liquid extraction of proteins using DES solvents

To enhance the physicochemical properties of these mixtures, water can be added (Martins et al., 2019). In conventional solid-liquid extraction, forming hydrogen bonds among the DES molecules and between DES and water molecules aids in capturing protein complex molecules, facilitating the efficient protein separation (Patra, Arun Prasath, et al., 2023). The high viscosity of DES impedes its practical use in various industrial processes, such as extraction processes (Gabriele et al., 2019). Altering the chemical nature of the components, adjusting their molar ratio, and changing the temperature (Gabriele et al., 2019), or modifying the water content can reduce this effect (Gabriele et al., 2019; Hammond et al., 2017; C. Ma et al., 2018).

While adding a minimal water content could enhance the fluidity, conductivity, and polarity of the DESs (Gabriele et al., 2019) and decrease the viscosity (Dai et al., 2015), a higher amount of water can have a drastic effect on their physicochemical properties. Dilution of natural DES weakens the interactions gradually and even makes them disappear entirely at around 50% (v/v) water addition (Dai et al., 2015). However, another study indicated that such interactions can entirely disappear at around 75% (w/w) (Gabriele et al., 2019).

2.1.2. Extraction optimization methods and quality prediction

2.1.2.1. Response surface methodology

In general, to achieve maximum economic yield and purity of phytonutrients, extraction and purification methods need to be optimised (Belwal et al., 2020). Therefore, it is necessary to plan and design experiments, and analyse the results (Aydar, 2018) using different statistical models or algorithms. By using the design of experiment, it is ensured that the samples are spread uniformly within the constrained domain, generating pre-processed data to remove eventual outliers and especially non-converged simulations (Galeazzi et al., 2022). Response Surface Methodology (Aydar, 2018), Taguchi design, and artificial neural networks are commonly used optimization models (Belwal et al., 2020).

Based on the principles of the design of experiments, (Veza et al., 2023) Response Surface Methodology is a widely used mathematical tool (Cano-Lamadrid et al., 2023; Fini & Gharehghani, 2024) for process optimization and enhancement. Compared to conventional experimental methods, it can generate considerable information from fewer experimental runs with a minimum experimentation cost (Veza et al., 2023). Full factorial design, central composite design, and Box-Behnken design are identified as the major classes (Veza et al., 2023). Among these designs, the Box–Behnken design is considered a popular one (T.-F. Zhang et al., 2011). Having three different levels for each factor, the Box-Behnken design has the specificity to avoid experiments at the extreme combinations of all factors. Rather than that, it offers improved prediction precision in the center of the factor space (Fini & Gharehghani, 2024).

2.1.2.2. Machine learning for extraction performance prediction

Process optimization is a crucial step for creating a feasible and cost-effective process design (Min et al., 2022). It is used to systematically probe and resolve a range of issues and challenges related to research, development, and production (Lundstedt et al., 1998). To make decisions, data is indispensable (Castro et al., 2022). In today's era of the fourth industrial revolution, the digital landscape is saturated with immense amounts of data (Sarker, 2021). In industry, three primary types of data are generated: The production data, which is produced from materials, equipment, and products (Kong et al., 2022), and is increasingly collected (Arantes et al., 2021), management data (enterprise resource

planning, product lifecycle management, and supply chain management), and external data (economy, market, customers, government) (Kong et al., 2022).

Big data analytics and data science technologies are gaining prominence in industry and engineering (Castro et al., 2022), and businesses in general, since it is widely relied upon by companies to enhance the reliability, speed, quality, and efficiency of their operations (Madugula et al., 2023). In particular, it can be used to identify anomalies in industrial machines and halt their operation before reaching a state of failure (Arantes et al., 2021) and to predict energy consumption (Kapp et al., 2023).

The key to analysing these data and developing applications can be the knowledge and use of artificial intelligence, particularly machine learning (supervised, unsupervised, semi-supervised, reinforcement learning, deep learning) (Sarker, 2021). In particular, machine learning has transformed many fields and has recently found applications in many fields such as chemistry and materials science (Jablonka et al., 2024), food compound analysis (Tseng et al., 2023), and agriculture. For instance, in agriculture, various machine learning models were used to identify the most influential variables for crop yield predictions (Dhaliwal & Williams, 2024).

2.2. Characterization of the protein isolate

Characterised by their complexity, food proteins can be categorised by their composition, structure, biological function, or solubility properties (Jain et al., 2020) or other functional properties. Despite significant advances in understanding the functional and physicochemical properties of proteins, further research is necessary to fully characterize and modulate their functionality (M. Kumar et al., 2022).

2.2.1. Structural morphology characterization of protein

In general, food microstructure is related to its chemical composition, rheological behaviour, and sensory attributes (Kaláb, 1995). Various methods exist for determining protein structure, each with distinct features and, in most cases, complementary (Lan et al., 2024) (Table 2). Among the used techniques, conventional light microscopy is the oldest microscopic technique available, providing basic information (Panato et al., 2017). Additionally, many new atomic force microscopy techniques, such as quantitative nanomechanical mapping, near-field scanning optical microscopy, Raman, and infrared nanospectroscopy, have been developed (Ruggeri et al., 2019).

Table 2. Methods for measuring protein structure

Detection method	Measured structure	References
MS, Edman degradation	Primary structure	(Lan et al., 2024)
FTIR fluorescence spectroscopy	Secondary and tertiary structure	
Raman spectroscopy, Circular Dichroism spectroscopy, Fourier infrared spectroscopy	Secondary structure	
NMR	Secondary, tertiary, and quaternary structure	

2.2.2. Molecular weight of the protein/polypeptide

In order to characterize the protein based on its molecular weight, PAGE (SDS-PAGE, native PAGE, 2-dimensional gel electrophoresis (2D-gel), IEF, and SEC-MALS) is a commonly used technique (M. Kumar et al., 2022). In particular, SDS-PAGE enables the separation, identification, and characterization of proteins in various food products (Joshi et al., 2025). It can assess key storage proteins (gliadins, glutenins, globulins, and caseins), providing insight into protein quality and functionality, and nutritional value (Joshi et al., 2025).

2.2.3. Fluorescence spectroscopy for protein characterization

Considered one of the most potent techniques in biophysics, emission fluorescence provides insight into protein conformation and intermolecular contacts, such as in protein-ligand and protein-protein interactions (F. H. dos Santos Rodrigues et al., 2023). They are characterized by an electronic absorption spectra over the ultraviolet region (185–320 nm) of the electromagnetic spectrum (Prasad et al., 2017). The intrinsic or natural fluorescence of proteins is related to the fluorescent amino acids, phenylalanine, tyrosine, and tryptophan, which are relatively rare in proteins (Lakowicz, 2000).

In addition to these amino acids, the common fluorescent markers in food include the structural proteins such as collagen and elastin, the enzymes and coenzymes NADH and FAD, various vitamins, lipids, porphyrins, and certain mycotoxins (Radotić et al., 2023). The fluorescence spectroscopy method is a rapid method used for many applications, such

as the characterization of oil composition (J. Xu et al., 2016), and to assess beef quality and detect signs of deterioration (H. Liu et al., 2019). Concerning protein, the intrinsic fluorescence spectrum reflects changes in polarity of the microenvironment surrounding aromatic amino acids (phenylalanine, tyrosine, tryptophan), indicating the tertiary conformational alterations (H. Wang et al., 2022). For instance, tryptophan, which is characterized by an excitation/emission wavelength of 270/280-430 nm (Militello et al., 2003), is commonly used to monitor protein conformational changes and infer details about local structural environments and dynamics (Vivian & Callis, 2001).

2.3. Effect of different treatments on legume protein extract on nutritional efficiency

Protein and its constituent amino acids are essential components of any healthy diet (Fukagawa, 2013). Despite the lack of scientific validation regarding the long-term efficacy and safety of high-protein diets, they remain among the most popular diets for rapid weight loss (Uchegbu & Kopelman, 2013). Preclinical research (animal experiment) demonstrated that prolonged consumption of this diet may be an effective strategy to reduce energy intake and expenditure while preserving body lean mass (Vu et al., 2017). While it restricts food choices, especially carbohydrates, it provides a diverse selection of protein-rich foods (Uchegbu & Kopelman, 2013). Interestingly, the emphasis is on animal protein over plant protein, inducing a higher intake of total fat, mainly saturated fat, and cholesterol (Uchegbu & Kopelman, 2013). Over the past few years, interest in plant protein and its utilisation has grown (Navaf et al., 2023), driven by emerging environmental and health awareness concerns (growing obesity epidemic) (Ewy et al., 2022). These protein proved their health benefits (D. Lin & Miao, 2021) and their ability to optimize the functional, textural, and nutritional quality of the food product (Meganaharshini et al., 2023). However, the transition from raw material to undergoing one (such as extracted protein) or several processing steps, especially after the incorporation into the final food, may affect the nutritional quality differently (Duque-Estrada & Petersen, 2023). While food processing can improve protein digestibility, the application of specific process conditions may induce an opposite effect (Duque-Estrada & Petersen, 2023). Figure 5 presents the potential impact of various processing techniques (such as thermal, high-pressure, and extrusion) on protein digestibility. This underscores the significance of the study of the nutritional properties of the extracted protein.

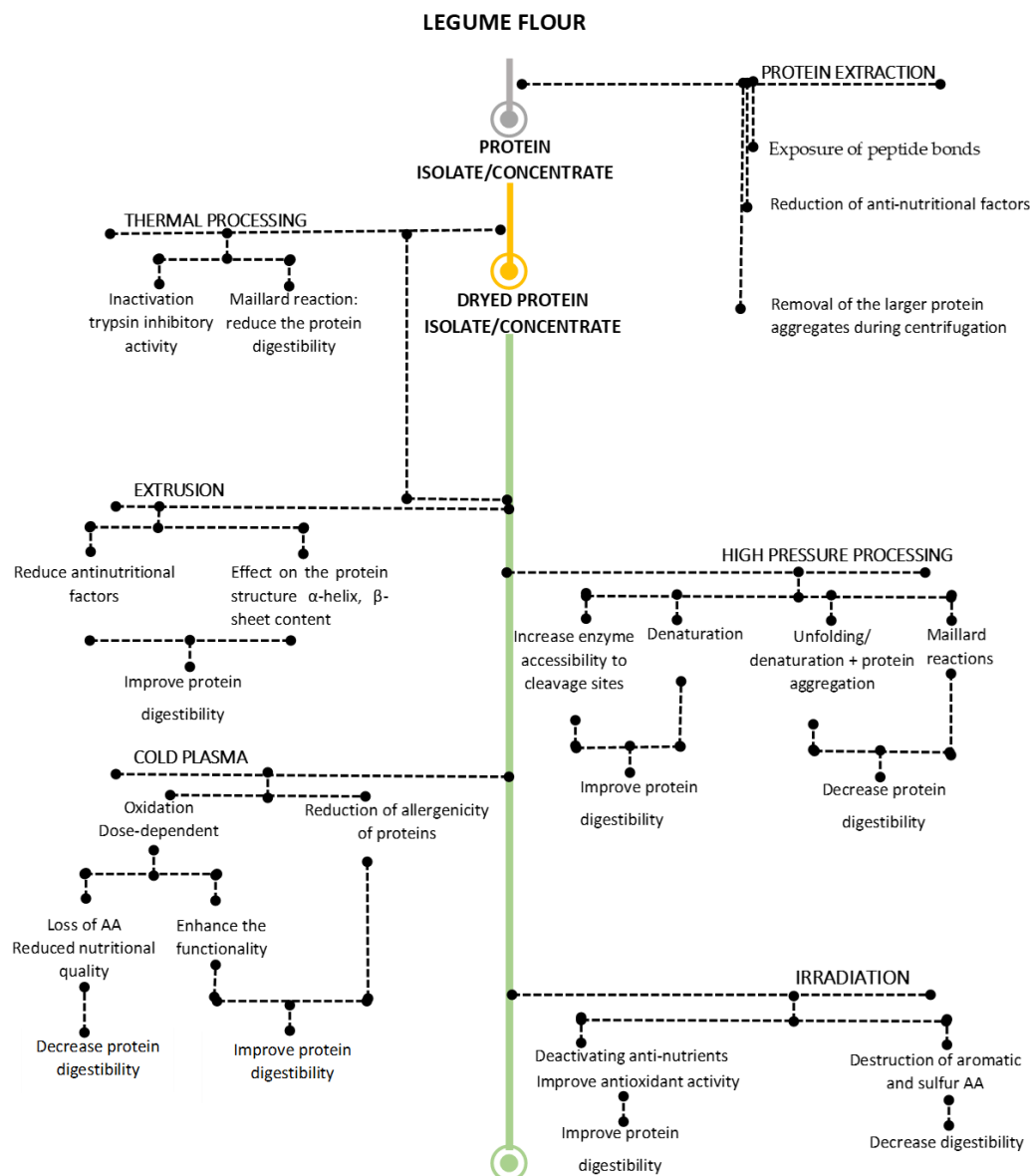


Figure 5. Impact of extraction and processing methods on the digestibility of protein isolate/concentrate (Neji et al., 2022)

On the other hand, single amino acid supplements in foods are increasing (Xiao et al., 2023). For example, being one of the essential amino acids and a necessary element in protein synthesis and various physiological activities, lysine was extensively supplemented to support the growth and development of children, especially in developing countries where diets are mainly based on cereal-based products (Gunaratne et al., 2025). Despite the uncertainty surrounding their potential nutritional and physiological impacts (Xiao et al., 2023), a recent study has demonstrated the effectiveness of the supplementation of a low-protein diet with essential amino acids to promote growth outcomes and repudiate the validity of the single-limiting amino acid

theory used in growth modelling (W. Liu et al., 2024). These results prompt the need for further experimentation to evaluate the effect of diet.

2.4. Evaluation and nutritional effect of soybean protein

2.4.1. *Drosophila melanogaster* model for nutritional evaluation

- **Stages of *Drosophila* development**

Different models can be used, such as the non-mammalian models (*Drosophila*, zebrafish, and nematodes) (Cai et al., 2025). *Drosophila melanogaster*, the fruit fly, serves as a versatile model organism widely used in biomedical research and biology studies, including complex human diseases (Giansanti et al., 2025; Jennings, 2011), ranging from heart disease to cancer to neurological disorders (Van Alphen & Van Swinderen, 2013). These models have the advantages of having low cost and short life cycles (Cai et al., 2025). In the case of *Drosophila*, the development starts with an embryonic development followed by the embryo hatching out of the egg shell to become a larva (Figure 6).

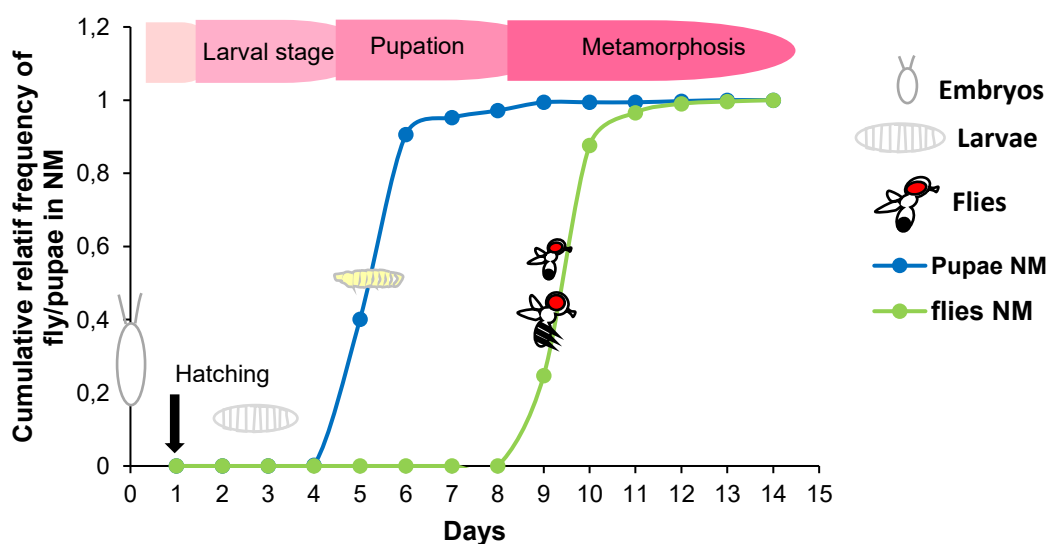


Figure 6. Synopsis of *Drosophila* development from egg to adult fly (Alberts et al., 2002)

Then the larva undergoes three distinct developmental stages, or instars. These phases are delineated by molting events during which it sheds its old coat of cuticle and lays down a larger one. The pupation occurs at the end of the third instar. The pupation starts

4 to 5 days after putting the embryos. The metamorphosis of pupae generates an adult fly (Alberts et al., 2002).

Considering the profound impact of dietary intake on the behavior of animals across a wide range of phyla, parameters such as growth, fecundity, and longevity, in addition to other behavioral and physiological assays, can be used to monitor the effect at larval, pupal, and adult stages (Ormerod et al., 2017). *Drosophila* can be a good model since, similarly to mammals, genetic studies have shown that they can regulate their metabolism, growth, and behaviour in reaction to the environmental conditions mediated via key hormones such as insulin, peptides with glucagon-like function, and steroid hormones (Koyama et al., 2020).

- **Genetic mutation in *Drosophila***

To screen for dominant mutations that suppress or enhance white variegation (Elgin & Reuter, 2013), the fly's variegated eye colour phenotypes, which are a marker of gene silencing (Y. Zhang et al., 2020), are evaluated. PEV is intensely studied in *Drosophila melanogaster*; and in particular with the mutant *white-mottled* using the *white* gene (Elgin & Reuter, 2013; Hoyer-Fender, 2020). It is described as the transposition of the euchromatic *white* gene into the vicinity of the pericentric heterochromatin that induced variegated expression of white due to the spread of heterochromatin (Figure 7) (Hoyer-Fender, 2020). PEV is among the phenotypic changes associated with epigenetic alterations in offspring influenced by environmental factors (Tando & Matsui, 2023). Research showed that nutrition and, in particular, changes in protein diet are among these factors. They can influence the epigenetic mechanisms, which regulate gene expression, which could influence productivity and productive traits in individuals and their offspring (Marín-García & Llobat, 2021).

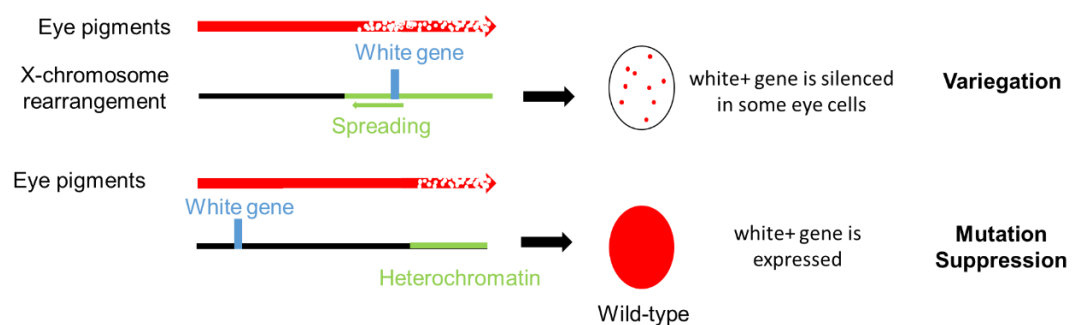


Figure 7. PEV in mutant vs wild-type *Drosophila*

- **Nervous system studies**

Since the brain is a vital organ with limited capacity for adaptation, it is essential to reveal its developmental coordination with the entire organism (Oliveira & Homem, 2023). The proliferative phase starts when the neuroepithelium begins expanding in the optic lobes as the larvae start feeding. Upon attainment of the “critical weight”, the generation of pulses of ecdysone marks the conversion of the neuroepithelium into neuroblasts and the beginning of the neurogenic phase (Maurange & Lanet, 2014)(Figures 8 and 9). On the other hand, the onset of pupariation, this hormone is also needed to trigger neuroblasts’ proliferation exit and consequently the end of neurogenesis (Oliveira & Homem, 2023).

Ecdysone is considered a humoral regulatory loop that determines organismal growth (Colombani et al., 2005) in general. It ensures that the brain has the correct number of neurons and glial cells (Oliveira & Homem, 2023). Nutrition and, in particular, amino acids can modulate their level via their action on *Drosophila* insulin-like peptides (Dilps) (Galagovsky et al., 2018). Furthermore, the level of this hormone can also be controlled by Sbm downregulation in glia. The research of Galagovsky et al., (2018) demonstrated that reduced levels of Sbm attenuate the ecdysone production, inducing a delay in the rise of PTTH levels associated with the generation of a smaller brain. This effect can be related to a reduction in cell proliferation with delayed consequences on neuron and glial cell numbers (Galagovsky et al., 2018).

Organ growth is recognized as being tightly regulated as a result of changes in environmental conditions, allowing it to achieve an appropriate final size. Although brain growth is generally regarded as unaffected by environmental factors, research showed that at least one compartment of the brain may exhibit a different growth pattern compared to the rest of the brain (Cobham et al., 2022). Research on *Platynereis dumerilii* larvae revealed that starvation conditions slow the development by reducing cell proliferation throughout the body. However, under normal conditions, resources were allocated to brain development, which manifested in the weak cell proliferation in the trunk segments, but an apparent rise in the nutrition of the newly larval stages does not have a neurological effect. For instance, during the development of cells in the head (Randel, 2025). For *Drosophila*, exposed to the same unfavourable conditions (starvation), the central nervous system thereof is more highly spared at the expense of other developing tissues (L. Y. Cheng et al., 2011; Franchet et al., 2025).

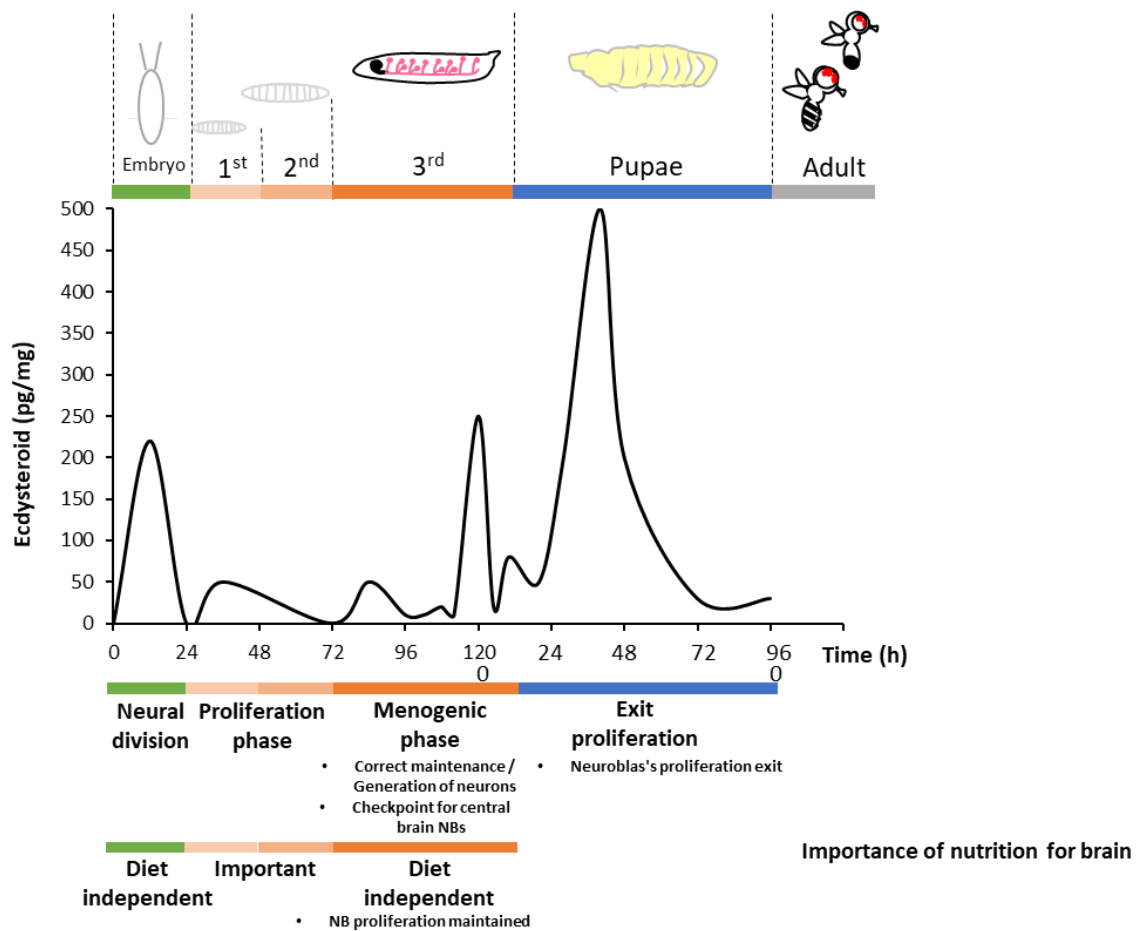


Figure 8. Brain progress and ecdysone concentration during larva development (Galagovsky et al., 2018), (Maurange & Lanet, 2014)

In contrast, starvation or nutritional fluctuations in general cannot continuously impair brain development. During embryonic and late embryonic stages, neural divisions remain unaffected by diet as the embryo is shielded from the environment and relies on its nutritional reserves (Maurange & Lanet, 2014). In the late larval stages, neural proliferation becomes insensitive to dietary nutrients due to the action of steroid hormone ecdysone, which modulates Delta/Notch signaling to promote the transition from symmetric mitoses to asymmetric neurogenic divisions (Lanet et al., 2013).

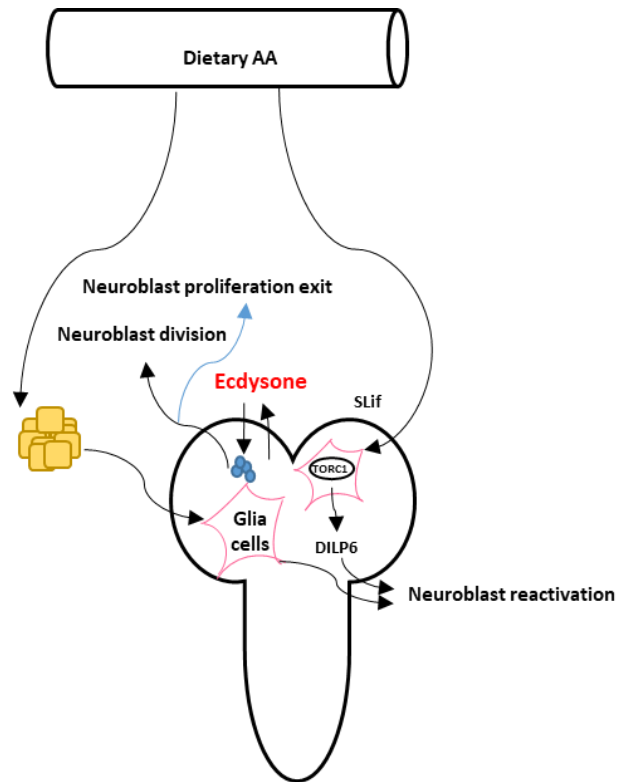


Figure 9. Nutritional and hormonal regulation during neuroblast development

2.5. Importance of legumes and cereal sequences on the immunological properties

As the world population grows along with rising affluence, urbanization, and evolving culinary preferences, protein production is becoming more diversified (Colgrave et al., 2021). Therefore, various alternative protein sources have been developed, varying from plants (i.e., grains, legumes, pulses, and nuts), fungi (i.e., mushrooms), algae, and insects to cultured (lab-grown) meat (J. Wu, 2022). In contrast, the potential health risks related to the consumption of some proteins remain a barrier to this development. In particular, proteins and peptides are recognized as responsible for food allergenicity (Hartmann et al., 2007). Among these proteins, tree nuts, peanuts, wheat, and soybeans are the most prevalent plant sources of food allergens, restricting their use and consequently limiting their consumption (Fadimu et al., 2023). While they possess these allergenic qualities, peptides and peptidomimetics can be immunomodulating agents, though blocking or stimulating the immune response (Gokhale & Satyanarayanajois, 2014). In particular, the immunomodulating peptides presented promising results against IgE-mediated allergic diseases (Lozano-Ojalvo & López-Fandiño, 2018). They influence intestinal homeostasis by maintaining and strengthening barrier function or modulating intestinal cell signaling to nearby immune cells and mucus production. Furthermore, they can stimulate both

innate and adaptive immune system cells while mitigating inflammatory responses (Lozano-Ojalvo & López-Fandiño, 2018). In particular, peptides can effectively defend against aberrant inflammation associated with chronic diseases like cancer (Mejia & Dia, 2009). For example, the anti-inflammatory peptides can inhibit the pro-inflammatory markers (Mejia & Dia, 2009). However, anti-proliferative peptides can present a promising prospect for developing drugs to cure cancer with fewer or no side effects (Shaik & Sarbon, 2022). Specific bioactive peptides generally possess unique structural characteristics that confer their particular functions (Mooney et al., 2012) or specific amino acid sequences in proteins that render certain proteins allergic (Goto et al., 2023). The identification of immunogens is needed before attempting to inactivate them (C. Liu & Sathe, 2018). In particular, revealing the structure of major allergens is a prerequisite for analysing surface-exposed areas of the allergen (Dall'antonia et al., 2014). 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). To enhance the efficiency of the detection of biological characterization of legume-derived peptides while minimizing time and cost, computational simulations and bioinformatics tools are emerging (Fonseca Hernandez et al., 2024).

Epitopes are recognised as a portion of a protein (C. Liu & Sathe, 2018) consisting of either contiguous or discontinuous specific amino acid residues on an antigen (He & Tao, 2015). Experimental techniques such as crystallographic and NMR-based approaches or predicted by computational methods are used to identify, analyse, and map conformational epitopes (Dall'antonia et al., 2014). Epitopes are classified into linear epitopes, which are recognized by their primary structure, and conformational epitopes. The conformational epitopes are composed of spatially adjacent, discontinuous amino acid residues and are related to the three-dimensional structure of proteins (F. Zhou et al., 2021). The type of the epitope can be a determining factor during the allergic reaction. In particular, antibodies can recognize either a linear epitope or a conformational epitope, unlike T-cell-receptor-that recognise exclusively linear peptides (C. Liu & Sathe, 2018).

Among the group of legumes, peanuts and soybeans are recognised to be part of the ‘Big 9’ allergen source (Abu Risha et al., 2024). Therefore, many countries require special labelling and manufacturing controls for foods containing these ingredients and their derivatives (Taylor et al., 2021). In contrast, although many other legumes (beans, peas, lentils, chickpeas, lupines, cowpea, pigeon pea, and fenugreek) are classified as “non-priority”, they are also potentially new important allergen sources, causing mild-to-severe allergic reactions (Abu Risha et al., 2024). For example, peas are not on the list of priority allergenic foods (Taylor et al., 2021) despite the recognition of vicilin, convicilin, and non-specific lipid transfer protein (nsLTP) as major allergens (X. Zhang et al., 2024). These results are further supported by the study of epitopes. For instance, Lima-Cabello et al, (2016) proved that peanut (Arah1), soybean (β -conglycinin), mungbean (Vig r 2), and pea (Piss 2) have a common epitope (T1). In addition, studies showed that the cross-reactivity of soybean allergens is likely to occur with allergens of other legume species (Wiederstein et al., 2023).

3. MATERIALS AND METHODS

3.1. Optimization of protein extraction

3.1.1. Materials

Es Mentor and Isidor varieties of soybean were used for protein extraction using a conventional method (alkaline-isoelectric precipitation) and using DES, respectively. Soybean cultivation and harvest took place at the University of Debrecen, Institutes for Agricultural Research and Educational Farm, Research Institute of Nyíregyháza (Nyíregyháza, Hungary). After that, the soybean grains were ground using Retsch SR2 (Retsch GmbH, Haan, Germany) into full-fat soy flour (0.5 mm), which was then stored at 4 °C.

For the conventional extraction method (alkaline-isoelectric precipitation), HCl and NaOH were used for pH adjustments (VWR Chemicals). Concerning the extraction using DES, choline chloride $[(CH_3)_3NCH_2CH_2OH]+Cl^-$ ($\geq 98.0\%$, high purity) and glycerol ($\geq 97\%$) were obtained from VWR Chemicals (Ohio, USA) and VWR Chemicals (Leuven, Belgium), respectively. All reagents were employed in experiments as received without additional purification. Distilled water was used as a solvent.

3.1.2. Methods

3.1.2.1. Extraction using data analysis and conventional method

Manual data mining served to retrieve information regarding the conditions for protein extraction. While creating this mini database, the pH of extraction and precipitation, and the extraction temperature were selected as factors. However, the purity of the protein, which reflects the crude protein content of the recovered protein, was considered as a response. Splitting the data into 70% for training and 30% for testing sets. The training data set is prepared for machine learning algorithms training in order to generate accurate predictions. The test data set assesses the model's performance (Pant et al., 2021).

Various machine learning models can predict the purity of protein isolate. Python (Jupyter) models were developed using libraries such as Pandas, Matplotlib, Numpy, Sklearn, and Seaborn. To discover the dataset's patterns, numerous predictive models, each distinguished by its approach: Bayesian or frequentist, parametric or nonparametric, data mining algorithm, or statistical model, among others (Shmueli, 2010). For this research, a group of algorithms was identified: linear regression, a polynomial model

frequently utilized for Design of Experiments (DOE), and machine learning algorithms. Machine learning, a key and crucial part of data science, is classified into two main classes: unsupervised and supervised (Rabbani et al., 2022). Supervised learning is divided into classification and regression (Arslan et al., 2021). In this study, regression models such as linear regression and logistic regression are applied to predict numerical values (Y. Wang et al., 2023).

- Linear regression: is among the most popular fundamental regression models (Hope, 2020). Additionally, it is also recognized as the simplest method for illustrating function fitting (Kotu & Deshpande, 2019).
- Nearest Neighbours Regression model: Known for its nonparametric nature (Srisuradetchai & Suksrikan, 2024), the nearest neighbor method relies on finding a number of training data points closest in distance to the new point and predicting the label (Zhu et al., 2022). It is a simple and effective method for handling complex structured data, particularly in big data contexts (Srisuradetchai & Suksrikan, 2024).
- Voting regressor: In general, Voting regressor is a machine learning ensemble methodology (Erdebilli & Devrim-İçtenbaş, 2022). It is used to improve the system's performance (Erdebilli & Devrim-İçtenbaş, 2022) and the prediction accuracy by combining multiple regression models (S. Chen & Zheng, 2025). In this study, the linear regression and the Decision tree regressor were applied.

Validation of protein extraction process

To assess the performance of the model, the coefficient of determination (R^2) and Akaike's information criterion (AIC), which is considered one of the less computationally demanding methods for comparing model fit, were used (Z. R. Yang & Yang, 2014). k-fold cross-validation was used for validation. The data points were randomly divided into k ($k=8$) groups. In each iteration, one group served as the validation set, while the remaining $k - 1$ groups were combined to form the training set (Maleki et al., 2020). For assessing the model's performance, we used the coefficient of determination (R^2) and Akaike's information criterion (AIC). We validated the model using k-fold cross-validation, where the data points were randomly divided into k groups ($k=8$). In each iteration, one group served as the validation set, while the remaining $k-1$ groups formed the training set. The full factorial design is regarded as one of the widely used

experimental designs in manufacturing companies (Antony, 2023). 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). These pH values were measured using the pH meter FiveEasy FE20 (Mettler-Toledo AG8603, Switzerland) and maintained at 8 and 10, corresponding to the low level (-1) and the high level (+1), respectively. A total of four randomized experiments based on two variables ($2^2=4$) were performed in triplicate according to the extraction procedure described in Figure 10.

The protein extraction procedure was conducted under mild conditions using the alkaline and isoelectric precipitation of protein, following the adjusted procedure of Johansson et al, (2022). An amount of 20 g of soybean flour was dispersed in distilled water at a 1:10 (w/v) ratio. The adjustment of pH was carried out using sodium hydroxide to dissolve protein from the soybean flour. The mixture was then stirred at 100 rpm for 1 hour using the incubator shaker (Infors HT Ecotron, Switzerland). Afterwards, centrifugation at 4020 RCF (for 30 min at 25 °C using Hettich EBA 21 D-8532 Tuttingen (Germany) separated the protein from insoluble material like carbohydrates. The supernatant was then filtered using a polyester sieve (88 µm) to remove the semi-solid cream layer from the soluble proteins. Protein precipitation was carried out by adding diluted HCl solution, followed by stirring at 100 rpm for 1 hour at room temperature (Figure 10).

Afterward, the obtained protein isolate was neutralized (pH 7), frozen, and then freeze-dried. Lyophilisation was conducted using an industrial-scale lyophilization equipment developed at the University of Debrecen. During the process, the vacuum chamber pressure was kept at 1013.25 Pa. The condensing chamber temperature was -40 °C. The sample temperature was monitored with a built-in core thermometer. The process was completed once the temperature reached 42 °C with a minimum duration of 24 hours. After that, the dried protein was weighed and stored for further analysis.

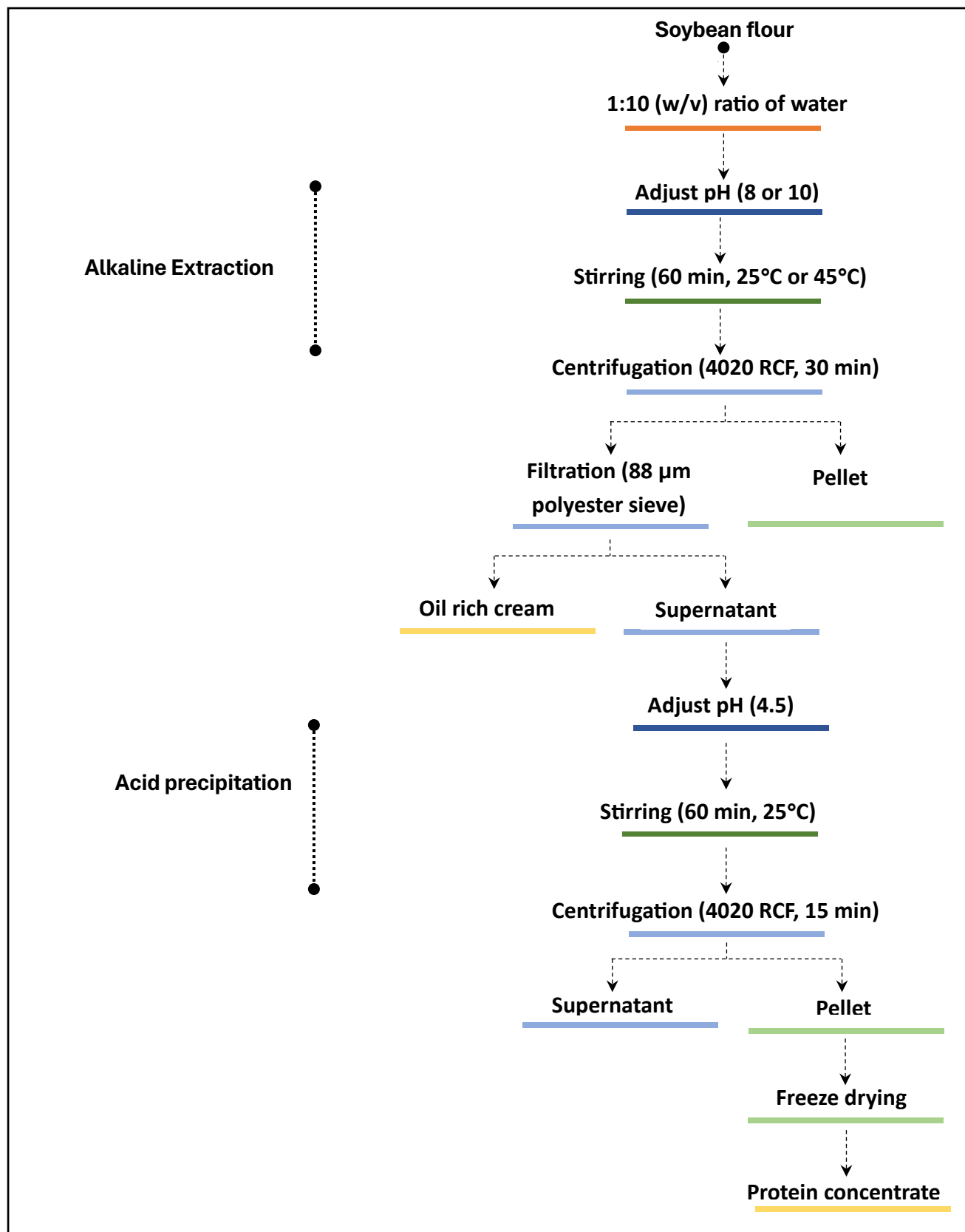


Figure 10. Experimental procedure for extracting soybean protein from Es Mentor variety using alkaline extraction and acid preparation

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 (%) = $\frac{\text{Weight of protein isolate} \times 100}{\text{Weight of soybean flour}}$ (Eq. 1) (Chiodza & Goosen, 2024)

Protein yield (%) = $\frac{\text{Mass of protein in protein isolate} \times 100}{\text{Mass of protein in the soybean flour}}$ (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 = $\frac{(\text{Measured value} - \text{Predicted value}) \times 100}{\text{Measured value}}$ (Eq. 3)

Percentage prediction error = $\frac{(\text{Predicted value} - \text{Measured value}) \times 100}{\text{Measured value}}$ (Eq. 4)

3.1.2.2. Extraction using Deep Eutectic Solvent (DES)

3.1.2.2.1. Preparation of Deep Eutectic Solvents

The preparation of the DES solvent followed the method of Yue et al, (2021) with some adjustments. Choline chloride and glycerol were combined in a molar ratio of 1:2 with water to get a DES/water ratio of 0:100%, 30:70%, and 60:40%. Then, continuous mixing with a magnetic stirrer at 25 °C was continued until the formation of a homogeneous and transparent liquid, approximately after 20 min.

3.1.2.2.2. Protein extraction using DES/water mixture

The extraction of protein 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 (Figure 11). After that, the samples were stirred at 160 rpm (Infors HT Ecotron, Switzerland) for a specific duration and temperature (Table 3), then centrifuged (Hettich, Germany) at 4020 RCF for 20 min to collect the supernatant. All experiments were conducted in triplicate.

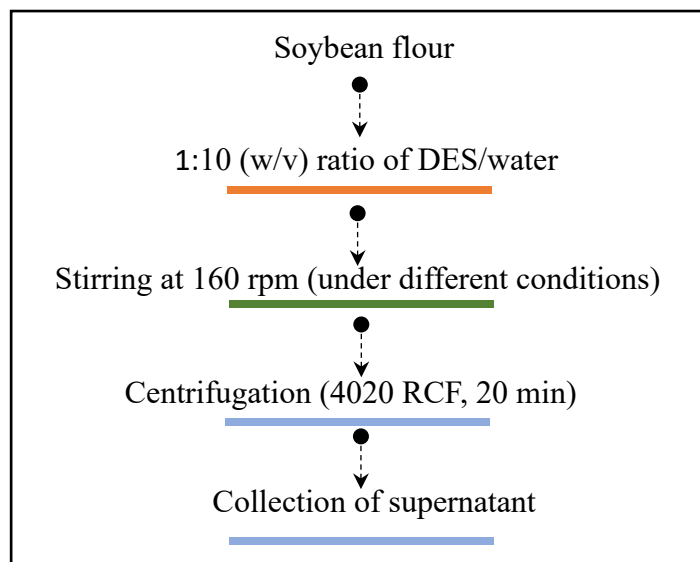


Figure 11. Experimental procedure for extracting soybean protein from the “Es Tribor” variety using DES/Water solvent

- **Characterization of DES**

The following parameters of the studied mixtures containing 30:70%, 50:50% and 60:40% of DES (ChCl:Glycerol): water was studied:

- The pH was measured with a digital pH meter from SevenEasy pH Mettler Toledo (China) at 25 °C.
- The electrical conductivity was assessed using the FiveGo portable conductivity Meter made by Mettler Toledo GmbH (China)
- The density (g/ml) was calculated using mass (g) and volume (ml) at 23 °C (Hewage et al., 2024).

3.1.2.2.3. Optimization of protein extraction

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

Response Surface Methodology is a fundamental mathematical tool for optimizing processes (Cano-Lamadrid et al., 2023). In particular, Box–Behnken design is one of the most frequently used methods for a second-order response surface model (T.-F. Zhang et al., 2011). In comparison with other designs, the Box-Behnken design excludes experiments where all the factors are set at their extreme level. However, the experiment combinations are located at the midpoints of the edges and the centre of the process space, and these designs are typically rotatable (Fini & Gharehghani, 2024). While constructing the design, to reduce the cost, a method based on the classification of parameters into groups, introducing Group Moment Matrix to estimate all the parameters in every group,

was used (T.-F. Zhang et al., 2011). In this research, part of the experimental runs were conducted using the Behnken design (Figure 12).

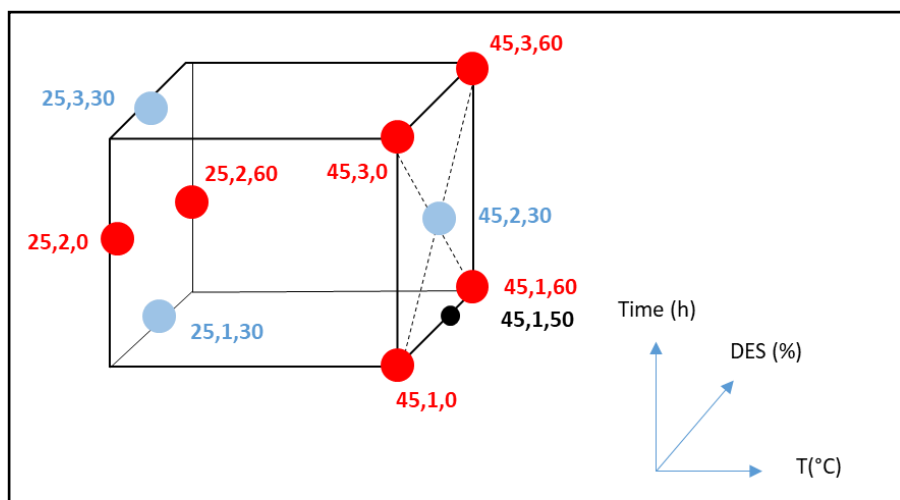


Figure 12. Applied combination of factors including extraction time (Time (h)), DES (%), and extraction temperature (T (°C))

To optimize the soybean protein extraction, a three-level-three-factor Box-Behnken design was used. The variables time and DES(%) were fixed at three levels. However, only two levels were chosen for the factor temperature to preserve mild reaction conditions.

Table 3. Minimum and maximum limits of variables

FACTORS/LEVELS	Minimum	Maximum
Time (h)	1	3
Temperature (°C)	25	45
DES(%)	0	60

While Box-Behnken is recognized for its ability to reduce the operational time and associated cost (Veza et al., 2023), minimizing further experimental runs may be advantageous.

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

- **Model validation**

The dataset was partitioned into training and testing subsets. While the training subset is used to develop the model, the testing subset, referred to as the validation dataset, serves to evaluate the accuracy of the developed model (Kotu & Deshpande, 2015). 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. It divides the datasets into K samples (usually 5 or 10) at random (Pachouly et al., 2022), and was used with K set to 10.

3.2. Evaluation of protein quality

3.2.1. Sample preparation

3.2.1.1. Protein extracted 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 (X. Peng et al., 2016). To evaluate the effect of high temperature treatment, stirring was completed at two different temperatures (45 °C and 65 °C) at 160 rpm (Infors HT Ecotron, Switzerland) for 1 hour. Similarly, to the research of Cao et al, (2023), absolute ethanol was mixed with the solution and kept at 4 °C overnight to precipitate the protein, then centrifuged at 4020 RCF for 15 min and freeze-dried. As a reference, an extraction using just water was conducted.

3.2.1.2. Protein isolate modified with enzymatic hydrolysis

The modification of soybean protein isolate purchased from the market (Brand Natura, Hungary) 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 (Guangdong Hongyuan Food Ingredient store, China) at 37 °C for 90 min with continuous agitation. Throughout hydrolysis, the solution's pH was continuously adjusted to pH 8.0 with 1 mol/L NaOH. To stop the reaction, the mixture was heated in boiling water for 10 min to inactivate the enzyme. After cooling to room temperature, the hydrolysate was centrifuged at 4020 RCF for 15 min. The resulting precipitate was freeze-dried.

3.2.1.3. Protein isolate subjected to physical modification

The physical modification of the soybean protein isolate purchased from the market (Brand Natura, Hungary) 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 min in a bath. After cooling, the solution was sonicated at a frequency sweeping between 50-60 Hz for 30 min using an ultrasound processor model Olympus Endosonic supplied by the manufacturer (KeyMed, UK). After that, 300mM of NaCl was added and centrifuged for 15 min at 4020 RCF. Finally, the collected pellet was freeze-dried.

3.2.2. Evaluation analysis

3.2.2.1. SDS-PAGE analysis

SDS-PAGE was carried out in accordance with the study by Domokos-Szabolcsy et al., (2022) with some modifications. 10 mg of the homogenised protein extract was dissolved in 250 µL of (2× Laemmli). The sample solution was then mixed with the buffer. After incubation at 95 °C for 5 min, followed by centrifugation (13000 rpm, 20 min) 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 min and 160 V for 59 min, 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).

3.2.2.2. Fluorescence spectroscopy

The intrinsic fluorescence spectra of the extracted soybean protein were measured with an F-8500 Spectrofluorometer (Jasco, Oklahoma, United States). The protein was dispersed in ultrapure water in order to attain a final concentration of 4 mg/ml and 0.01 g/ml, corresponding to samples extracted with DES: water and treated with enzymes and physical modification, respectively. The protein samples were then filtered (syringe filter FPFE, 0.22 µm). During the acquisition of the excitation-emission matrix measurement,

the excitation and emission wavelengths used ranged from 200–550 nm and 210–750 nm, respectively.

3.2.2.3. Light microscope

5% (w/v) and 10% (w/v) of protein extracted with DES, and treated protein (enzymatic hydrolysis and physical modification), respectively, were stained using Fast Green FCF (Thermo Scientific, A16520.06). After vortexing, the samples rested for 20 min. 20 µL of the samples was placed onto the slides. Then, they were covered with the coverslip without the use of excessive pressure.

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.

3.3. Assessment of the nutritional effect of soybean protein

3.3.1. *Drosophila melanogaster* as a model for nutritional value assessment

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

The evaluation of nutritional effects by *Drosophila melanogaster* *in vivo* digestion experiment was conducted using the 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. Fly stocks were maintained as overlapping bottle generations at 25 °C and 60–70% humidity during stock maintenance in complete darkness. Experiments were performed under identical conditions. Gathered from population cages containing about two hundred females and males (5-day-old), embryos were placed on zero-nutrient media (0M) plates supplemented with yeast paste. The 0M is prepared using 1% of charcoal vegetal activated powder (VWR, 87126.230) mixed with 1% of agar (VWR, No. 20767.298) in water, boiled for 5 min before cooling to 45 °C and aliquoting into vials containing 3 ml.

Fifty embryos (0-2 hours old) were collected under a microscope and introduced into the corresponding vial. To prepare the NM and HS media, 70 g of commercial yeast paste was mixed with 1 L of water until homogeneous, 52 g (for NM-diet) or 256.72 g (for HS-diet) of D(+)-Saccharose (VWR, 27483.363), and 30 g of wheat flour. After that, 10 g of agar powder was added after thorough mixing. The boiling continued for over 30 min,

and the culture media reached a final volume of 1 L. Next, the culture media were cooled in a water bath to 50 °C, and 1 g of Nipagin (Ethyl 4-hydroxybenzoate 99%, A13172) purchased from Thermo Fisher was mixed thoroughly.

The different media were supplemented with soybean protein isolate “Es Mentor”, which was extracted according to the conventional alkaline-isoelectric precipitation or L-lysine 98% (Thermo Scientific, 303341000). The supplementation was done at five different concentrations: 0.25%, 0.5%, 1%, 2%, and 4% (w/v). For each diet, at least five replicate vials were tested. Vials were monitored every 24 h for 20 days, and the number of pupae and flies was recorded.

To evaluate whether lysine could contribute to the development of flies together with soybean protein, different percentages of lysine were added to the 0M. These experiments were done according to Table 4, with three different concentrations of SP and lysine (2%, 1%, 0.5%).

Table 4. Mixture composition and percentage

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

3.3.1.2. Pupae length and width measurements

Pupal size (length and width) was determined according to the method of Enriquez et al, (2022). The pictures of at least 10 pupae raised in 0M supplemented with SP were taken using the optical microscope Olympus Corporation (Model SZX2-ILLK, Japan). Pupae from NM were used as controls. The pictures were taken after 9-10 and 6 days after the embryo's implantation for 0M treatment and NM (control), respectively. The length and width were measured using the freely accessible online software ‘Imagej’ (*ImageJ*, n.d.).

3.3.1.3. 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. 10 male fly heads were mixed with 300 μ l of chloroform and 300 μ l of NH_4OH (0.1%) and then centrifuged (7000 rpm, 3 min). 200 μ l of the top layer was mixed with 500 μ l of NH_4OH (0.1%) and transferred for spectrophotometric analysis (485 nm).

There are two classes of pigments in *Drosophila* eyes: the brown ommochromes and the red drosopterins, along with pteridine and sepiapterin. They have been identified as the compounds responsible for the typical eye colour phenotype (Wiederrecht et al., 1981). In particular, drosopterin and isodrosopterin are considered the major pigment components (Yim et al., 1993). Knowing the extinction coefficient of drosopterin ($25100 \text{ M}^{-1} \text{ cm}^{-1}$) at 485nm, this coefficient was used to determine drosopterin concentration (Evans & Howells, 1978). Duplicate readings were averaged and expressed as drosopterin concentration. The heads of the *D. melanogaster* strain Viking GFP (vkg-GFP) (white eyes, curly hair, yellow body), which does not have eye pigmentation, were used as a negative control. The strain was obtained from the Bloomington stock center (y[1] w[*]; P{w[+mC]=PTT-un}vkg[G00454]).

3.3.1.4. 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. The 3rd instar larvae were grabbed using a bacteria inoculating loop and moved to PBS solution on an excavated glass slide. The larvae are pinched and held from the head around the mouth parts with forceps and the posterior part. After peeling the larval cuticle, all the organs (fat body, internal organs, and trachea) are removed. Cleaning activities were conducted with due diligence to avoid damaging the brain, which is attached to the mouthpart and imaginal discs. At about 72 hours, 3rd instar larvae were collected from 0N media enriched with 2% and 4% of SP. Larvae reared in an NM diet were collected as a control. 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.

3.4. Assessment of immunological properties of cereal and legume protein sequences

3.4.1. Dataset preparation and evaluation for epitope characterization

To characterize plant proteins with allergenic and bioactive functionality derived from cereals and legumes, a small database composed of protein sequences was constituted and analysed.

3.4.2. Dataset preparation

A group of 263 selected sequences of allergens and bioactive peptides (immunomodulatory/anti-inflammatory/anti-proliferative) from cereals and legumes (Figure 13) was 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.

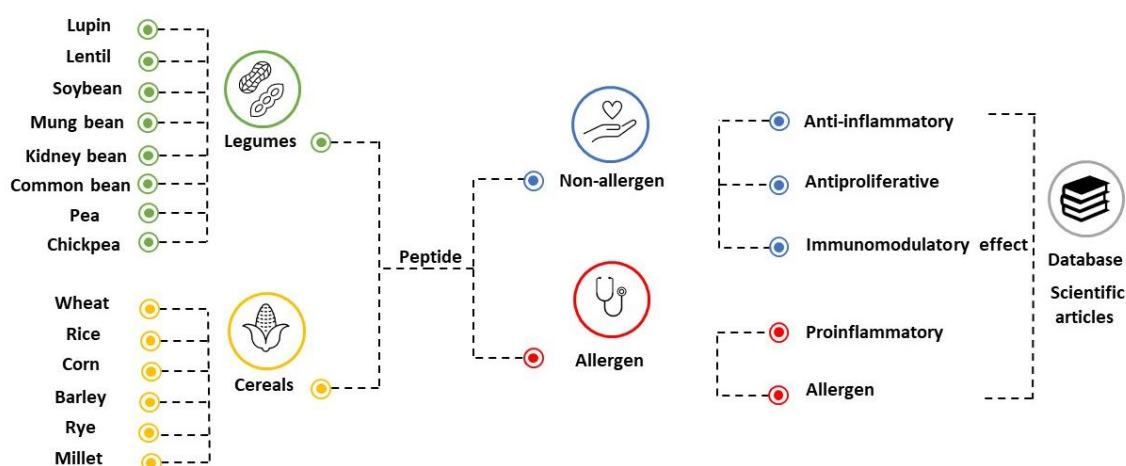


Figure 13. Allergenic and non-allergenic peptides derived from legumes and cereals.

Different plant sources and the specific function of the studied protein sequences

1963 allergenic epitopes from different sources (legumes, fish, and cereals), and 32 immunosuppressive epitopes were collected from the IEDB-AR database (IEDB.org: View Results and Refine Search). (*IEDB Analysis Resource*, n.d.).

The characterization of peptides includes the amino acid count (prevalence frequencies of 20 amino acids), molecular weight, 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, and 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.

3.5. Statistical analysis

All the results were expressed as means $\pm$ standard deviations of experimental triplicate from each independent experiment.

For choline chloride extraction, data related to the model were statistically analyzed using Minitab (version 21). Similarly, for the data obtained through the fly experiment. Mean differences between sample treatments were assessed with one-way ANOVA followed by Tukey's comparison test, and the significance of differences was defined at $p < 0.05$.

For protein extraction, Python (Jupyter) and R were used. The 3D fluorescence spectra and the related charts were illustrated using GraphPad Prism version 9.5.0 (730).

For peptide and allergen characterization and analysis, the data were collected using an Excel sheet. Jupyter Notebook using different libraries (Matplotlib, Protopy, Pandas, Numpy, Seaborn, Sklearn), classes (SequenceMatcher), and packages (Bio.SeqUtils, Peptides) was used for data visualization and characterization. On the other hand, the cosine_similarity metric was used to test the similarity between epitopes. 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. The Principal Component Analysis (PCA) was used to minimize the dimensionality of the dataset and hence enhance the interpretability of the results.

4. RESULTS

4.1. Optimization of protein extraction

To improve the soybean protein extraction efficiency, two different approaches were conducted (using the conventional method and with DES solvents) and validated with experimental analysis.

4.1.1. Extraction using data analysis and conventional method

4.1.1.1. Understanding the data

Based on the collected data, during protein extraction from the soybean flour, many factors can influence the purity of the protein isolate, such as raw material and chemicals.

4.1.1.2. Relation between predictors and dependent variables

The correlation matrix (heat map) in Figure 14 shows the linear relationship between factors (temperature, pH of extraction, pH of precipitation) and the purity of the protein, indicating the strength and direction of each pairwise correlation. Reflecting the strength of the correlation, higher scores denote a stronger correlation (Hazra et al., 2024). For instance, a score below 0.4 is considered weak; 0.4-0.7 indicates a moderately strong correlation; and above 0.7 is regarded as strong (Y. Lin et al., 2023).

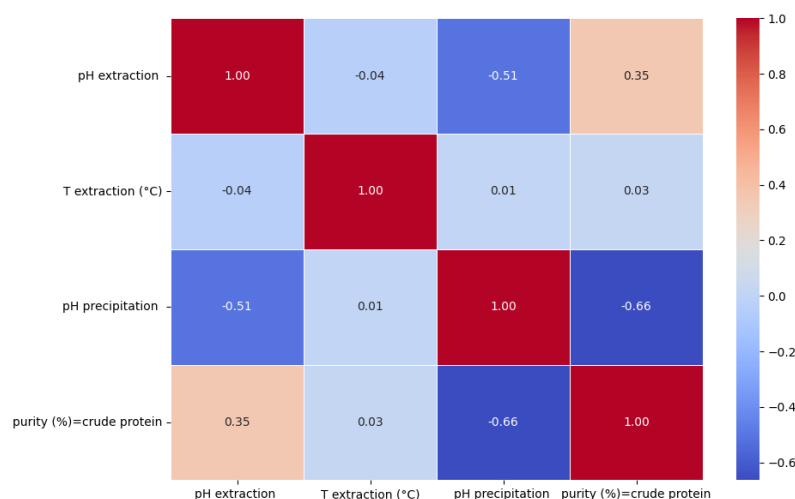


Figure 14. Correlation among extraction factors (temperature, pH of extraction, and pH of precipitation) and the response (purity of protein). The colour of cells represents the strength of the correlation between feature-to-feature and feature-to-response. The strength of the correlation is indicated in the colour scale: red (positive correlations), blue (negative correlations).

The sign of the correlation indicates its direction: positive values suggest a favourable correlation, while negative values signify an inverse one (Khalil et al., 2024). A low positive correlation was found between the purity of the protein isolate (response) and the pH of 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.

4.1.1.3. Optimization of soybean protein isolate extraction conditions

4.1.1.3.1. Model fitting

To evaluate the regression algorithms, methods such as the coefficient of determination (R^2), root mean squared error (RMSE), Akaike's information criterion (AIC), and the Bayesian information criterion (BIC) can be used (Maleh et al., 2021). Considering the results from various models, the voting regressor yields achieves the highest prediction accuracy, and the lowest mean squared error (RMSE) (Table 5). The high R^2 indicates a strong explanatory power (H. Liang et al., 2024). The model fits 83% of the training data and 81% of the test data set. As shown in many studies, similar R^2 values of the training and test data suggest the model is not overfitting (Cobb et al., 2021).

Table 5. Statistical evaluation of the models

Model	R2 score		RMSE		AIC	Equation
	Training	Testing	Training	Testing	Training	
Voting_reg	0.83	0.81	9.07	9.61	349.02	
Polynomial model (3 rd degree)	0.69	0.64	12.21	13.37	350.29	$y = -1274.465031 + 155.670318 \cdot \text{pH extraction} + 486.419176 \cdot \text{pH precipitation} - 16.317953 \cdot (\text{pH extraction})^2 - 0.941313 \cdot \text{pH extraction} \cdot \text{pH precipitation} - 84.476774 \cdot (\text{pH precipitation})^2 + 1.178910 \cdot (\text{pH extraction})^3 - 3.084978 \cdot (\text{pH extraction})^2 \cdot \text{pH precipitation} + 4.950677 \cdot \text{pH extraction} \cdot (\text{pH precipitation})^2 + 2.217404 \cdot (\text{pH precipitation})^3$
Nearest Neighbors Regression (K=5)	0.68	0.74	12.41	11.42	338.46	
Linear Regression	0.41	0.49	283.67	249.68	378.75	$y = 114.0666 + 2.0245 \cdot \text{pH extraction} - 10.6792 \cdot \text{pH precipitation}$

AIC, an information criteria based on penalized likelihood (Dziak et al., 2020), is more suitable for selecting the best model for predicting future observations than BIC (Chakrabarti & Ghosh, 2011) and is the asymptotic equivalent of cross-validation (Z. R. Yang & Yang, 2014). It showed a lower value for the voting regressor compared to the polynomial and linear regressions. This result suggests that the voting regressor is well-fitted, since a lower AIC (Leitner & Turner, 2017) or BIC (Bayesian information criterion) value indicates a better fit (Mohammed et al., 2015; Schreiber, 2008).

The distribution of results (Annex 1) is not normal, which explains the low determination coefficient in both testing and training data. The research of Halimi et al, (2015) linked the optimal model with the normal distribution of residuals. In all results, residuals are centered around zero (Annex 2), randomly, and are normally distributed (Annex 3). This result was validated by the quantile-quantile comparison (Annex 3), indicating that the residuals of the Voting_reg model follow the reference distribution, since it is linear.

Cross-validation is a widely used method in machine learning for training and evaluating the model's performance (Seraj et al., 2023). The implementation of this method can be achieved using various data sources for training-evaluation steps or dividing a large dataset into distinct subsets using split-half (50% training, 50% test) or imbalanced-split (i.e., 70:30 train-test, or 50:25:25 train-test-validation) sampling techniques (Delgadillo & Atzil-Slonim, 2023). Using both methods (train-test (70:30) and 10-fold cross-validation), the Voting regressor achieved the highest R^2 (Figure 15).

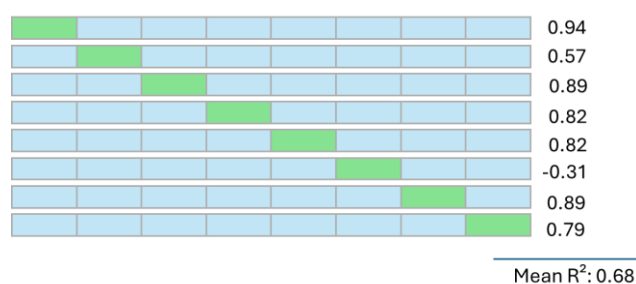


Figure 15. Cross-validation with $K=8$ performed on the Voting regressor algorithm. Each group was iteratively selected as the validation set, while the remaining $k - 1$ groups were combined and used as the training set. The R^2 metric was used to evaluate the proportion of variance in the dependent variable (protein purity), indicating the goodness of fit of the regression model.

4.1.1.3.2. Validation of protein extraction process

According to the results from the mini database, the alkaline extraction of soybean flour, presenting $37.76 \pm 0.32\%$ of protein (dry basis), resulted in a relatively pure protein isolate (Figure 16). This is likely because the alkaline solution can break down the plant cell matrix and solubilise the protein (Hadidi et al., 2023).

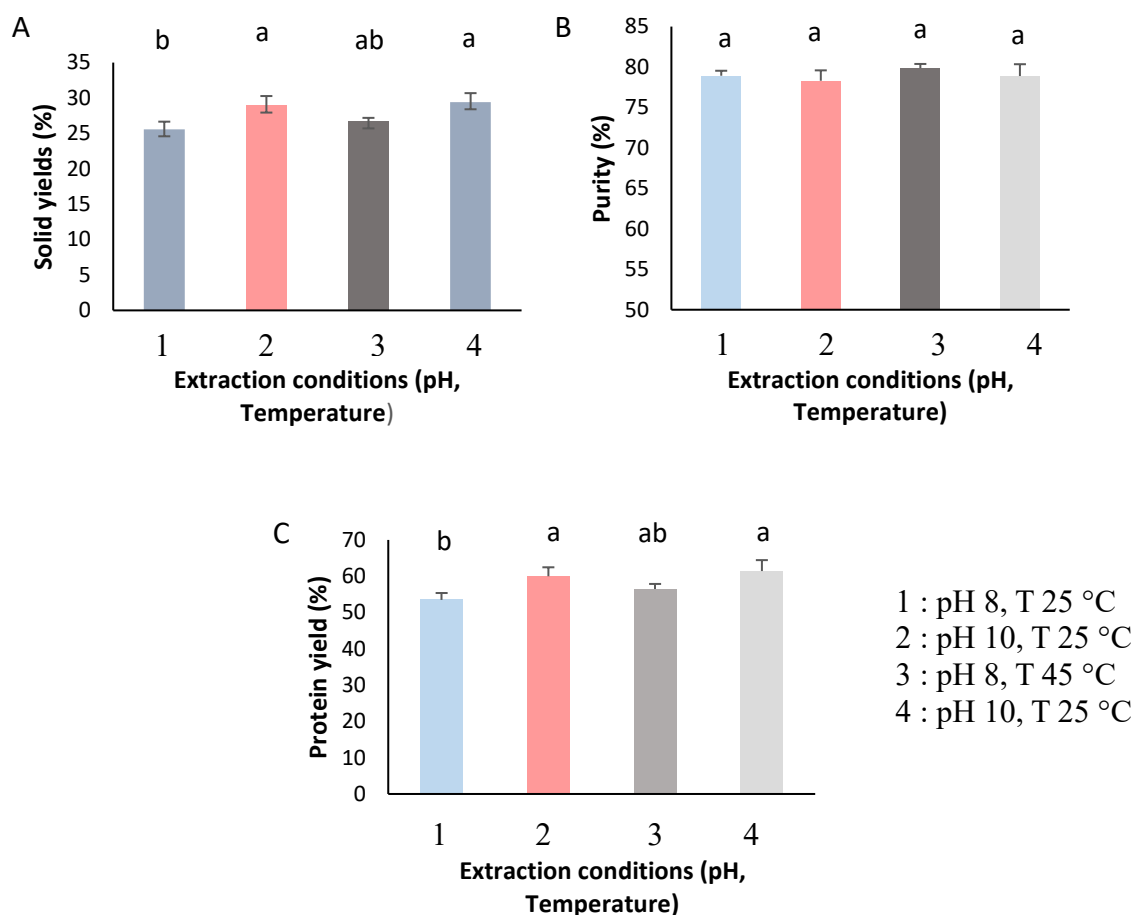


Figure 16. Soybean protein properties after extraction with different extraction pH (8 or 10) and temperature (25 °C or 45 °C). 15A: Solids yield of the extracted protein, 15B: Purity of the extracted protein, 15C: Yield of the extracted protein. The error bars indicate the standard deviations of the triplicate samples. Values with different superscripts differ significantly at $p < 0.05$

Additionally, the fat's low density and higher freezing point compared to water led to the formation of a creamy butter-like layer during centrifugation. This made fat removal easier, reducing the fat content (Chamba et al., 2015).

Mild fractionation at pH 8 and 25 °C produced an isolate with a protein content of $78.92 \pm 0.61\%$. A separate study achieved comparable purity (85.3%) after the extraction with the same process (Y. Peng et al., 2020).

The findings indicated that fluctuations in temperature and pH levels during the extraction process did not 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 17).

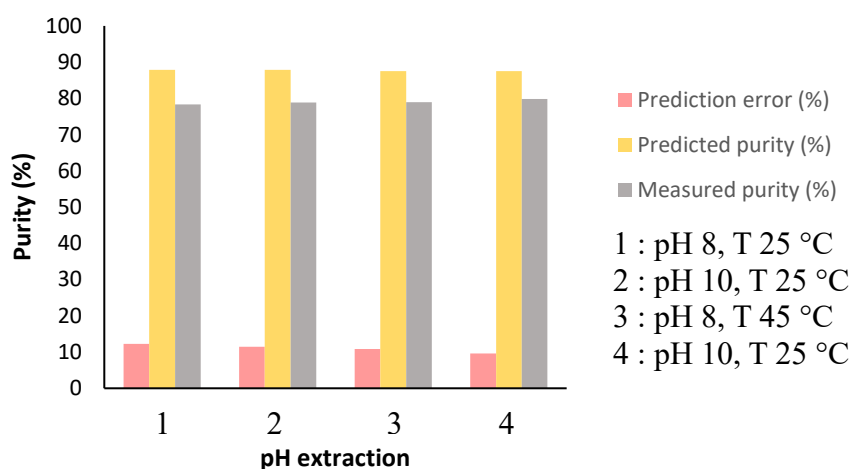


Figure 17. Variation between the predicted (estimated), measured (real) purity, and prediction error.

In agreement with the findings of Y. Peng et al, (2020), the purity of protein obtained through mild extraction methods was comparable to that of commercial protein ($79.98 \pm 0.51\%$), which is unexpected. Since oleosin proteins form a coating around oil bodies in soybeans (B. Chen et al., 2012) compared with conventional methods that include defatting treatment, a portion of the protein can be removed along with the lipid fraction. Previous reports have indicated that the cream layer obtained from soybeans and collected post-centrifugation of the soaked solution (pH 8.6, 4–6 °C, overnight) contained $45.8 \pm 2.9\%$ of oil and $27.6 \pm 1.4\%$ of protein (B. Chen et al., 2012). Consequently, the similarity observed between our results and commercial protein isolate may be attributable to differences in soybean varieties or the production process scales.

The variation of the extraction parameters presented a significant impact on the protein solids (Figure 16-a) and yield (Figure 16-c). The optimal result was obtained at (pH 10,

45 °C). Adopting a mild extraction process, 53.46 - 61.41% of the protein was recovered from the protein fraction in the soybean flour (Figure 16). The research of Y. Peng et al., (2020) showed similar results while using comparable extraction conditions, including cream removal ($55.8 \pm 6\%$). Other techniques to produce hexane-free soybean meal, such as the gas-supported screw pressing, also revealed promising results (59.7%). The latter method produced significantly higher protein yields than those defatted with hexane (51.6%) (Nazareth et al., 2009). Compared to studies that used conventional techniques, the use of natural chemicals and full-fat soybean flour to adjust the pH of the solutions resulted in a low protein yield (43.62%) compared to the experiment with the synthetic chemical and defatted soybean flour (51.16%) (Chamba et al., 2015). Additionally, other factors, such as the adapted agriculture strategy, may affect the protein yield. For example, organic soybeans had higher protein than the commodity soybeans due to improved handling during identity-preserved storage or variations in soybean variety (Nazareth et al., 2009).

4.2. Extraction using DES

4.2.1. Selection of factors for the optimization of protein extraction

To determine the degree of efficiency of protein extraction using DES: water, the diffusion of nitrogen originating from soybean flour was examined. The determination coefficient (R^2), which is used to assess the model's accuracy (Kouzehgar & Eslamian, 2023), (Table 6) indicates that the RandomForest, and the polynomial model showed a good fit to the data. This finding was validated by the low root mean square error (RMSE), which indicates better simulation results (Kouzehgar & Eslamian, 2023).

Table 6. Models Summary

Models	RMSE	R^2	MAE	Rel_RMSE
Linear	0.0525	0.620	0.0452	0.137
Lasso	0.0523	0.625	0.045	0.137
Ridge	0.0526	0.618	0.0433	0.138
RandomForest	0.012	0.987	0.0115	0.0313
Polynomial	0.0282	0.977	0.0206	0.0738

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 18), as the Kjeldahl method measures the overall nitrogen levels (Prandi et al., 2019). The study of Tenhunen et al, (2018) demonstrated that while treating the pulp with DES, the N residuals found originated from strongly bound choline chloride residuals (Tenhunen et al., 2018).

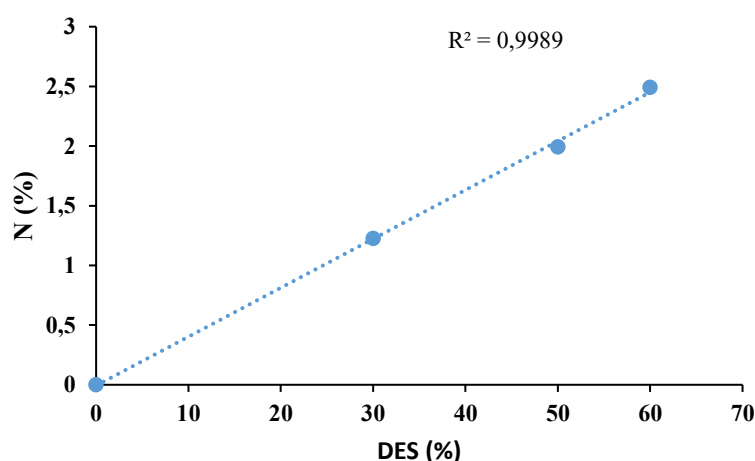


Figure 18. The N content of the DES: water solutions (0:100%, 30:70%,50:50%, and 60:40%). A strong linear relationship between the percentage of DES and the N content, as indicated by the coefficient of determination R^2

When the percentage of DES is around 60%, the extraction revealed the lowest yield, giving a result lower than the effect of pure water ($0.43 \pm 0.01\%$). The water ratio can also be a responsible factor since water supplementation, even at a minimal ratio of 40%, can contribute to the inefficiency of this binary solution. It alters its physical properties by transforming its mixture from binary to ternary (Omar & Sadeghi, 2022).

Other researchers found that the integrity of DES was compromised as a result of DES hydrogen bond complex disruption and the elective solvation of the extracted components by water. Every type of DES has a threshold for the amount of water that can be incorporated before the total disruption of the DES complex, leading to the formation of an aqueous solution with individual solvated components (Passos et al., 2016). This effect was confirmed by the study of Ferreira et al, (2021). As in this study, using choline chloride: glycerol (at a 1:2 molar ratio) and 11 wt% of water did not perturb the DES

structure. At the molecular level, it induced the formation of cohesive interactions (glue-like effect) serving as a molecular bridge between DES components. In contrast, 35wt% of water disrupted the DES structure, making a transition to a DES-in-water solution (Ferreira et al., 2021).

Changes in the protein structure (protein conformation, local folding) and the solvation environment co-occur, induced by DES hydration, were also confirmed by the study of Sanchez-Fernandez et al, (2022a). This research demonstrated that the amino acid environment and secondary structure of two proteins (bovine serum albumin and lysozyme) and an antibody (immunoglobulin G) in 1:2 choline chloride: Glycerol display re-entrant behaviour as a function of solvent hydration. At low (<10 wt% H₂O) and high (>40 wt% H₂O) levels of DES hydration, a dome-shaped transition characterized by folded or partially folded protein structures is observed (Sanchez-Fernandez, Basic, et al., 2022a). The research of Vilková et al, (2020) confirmed the potential positive effect of water addition and the importance of the viscosity and polarity as DES features. In particular, the addition of water reduces the viscosity of the media, improving the mass transfer and thus the extraction efficiency. Conversely, high water content can weaken the interactions between the DES and the compounds, as well as among the DES components (Vilková et al., 2020).

Other research proved the efficiency of the DES binary solvents in extracting protein. For instance, oat protein extraction was successful using the binary mixtures with a lower concentration of ChCl (i.e., 1:2:1 and 1:3:1) except the one with 2,3-butanediol at 1:2:1 ratio (Yue et al., 2021). In contrast, the highest concentration of ChCl (1:1:1) was not successful. This effect was assigned to the high viscosity of the solvent/flour mixtures (Yue et al., 2021).

The density is affected significantly by the addition of DES. It gives an idea about the solvation ability of DESs (Abdollahzadeh et al., 2022) (Figure 19).

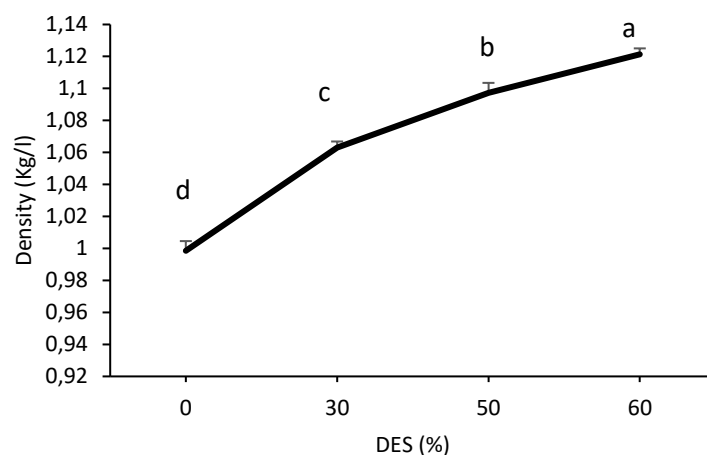


Figure 19. Density of the DES:water mixtures

The error bars indicate the standard deviations of the triplicate samples. Values with different superscripts differ significantly at $p < 0.05$

With the addition of water, an adverse effect was obtained for the conductivity (Figure 20). As a result of the decrease in viscosity, it has been proven that a low amount of water may increase the conductivity by up to 100 times for some NADES. Although the decreased density could signify an enhancement of mass transfer, water dilution may make the interactions weak gradually and ultimately vanish completely at approximately 50% (v/v) water supplementation (Dai et al., 2015).

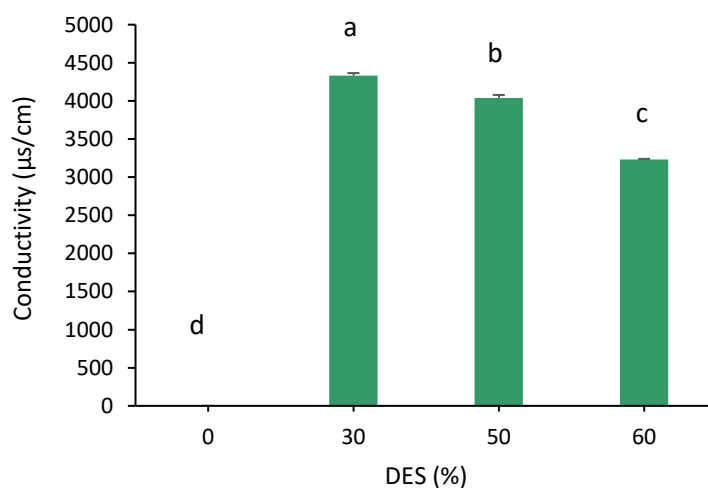


Figure 20. Conductivity of the DES: water mixtures

The error bars indicate the standard deviations of triplicate samples. Values with different superscripts differ significantly at $p < 0.05$

Alongside the liquid-liquid extraction, choline chloride-based DES aqueous solutions demonstrated a high absorption capability for capturing NO₂ while achieving a lower viscosity (C.-C. Chen et al., 2021).

The inefficiency of soybean protein extraction using (ChCl: Glycerol) may derive from a combination of factors in relation to the characteristics of these solvents, such as their type, structure, properties (Ali et al., 2015), and extraction conditions, such as temperature. In particular, an extraction under 210 °C using choline chloride and oxalic acid removed N from hydrochar with an efficiency of 69.49% compared to conditions without DES (Z. Xu et al., 2023). The research of Ali et al., (2015), highlighted the impact of DES structure on the removal ability for N-compounds and the extraction mechanisms (Ali et al., 2015). The extracted compound's nature and properties are important parameters. Extracted under the same conditions, ChCl: Glycerol (1:2) had an extraction efficiency of 51.4% and 78.9% for pyridine (basic N-compounds) and carbazole (non-basic N compounds), respectively (Ali et al., 2015). In our study, the amino acid that comprises the soybean protein is characterized by its variability. It has acidic with a pronounced negative charge amino acids (31.63 %), basic (24.05%), which have a positive charge, and attract oppositely charged molecules, in addition to non-polar proteins (37.06%) (Rijavec & Zupin, 2015). On the other hand, used as HBD, glycerol derives from triglycerides (Yeong et al., 2012) that are frequently referred to as neutral lipids since they do not present free acidic or basic groups (Penfield & Campbell, 1990). This may be considered as one of the factors contributing to the reduced binding affinity to proteins.

The acidity of the HBD component in DES has a critical role in denitrogenation performance. Insufficient acidity results in weak interaction with basic N compounds. However, too-high acidity can induce a chemical reaction between DES and basic N-compound, which can trouble extractant regeneration. Excessive acidity of HBD and insufficient hydrogen bond basicity of HBA can also lead to weak extraction of non-basic N-compound (Ali et al., 2015).

In this research, the limited effectiveness in extracting proteins can be related to the neutrality of glycerol. Originating from triglycerides (Yeong et al., 2012), commonly known as neutral lipids since their molecules do not present free acidic or basic groups

(Penfield & Campbell, 1990). This result is further supported by the pH measurement of the used solvent (Figure 21), proving its neutrality.

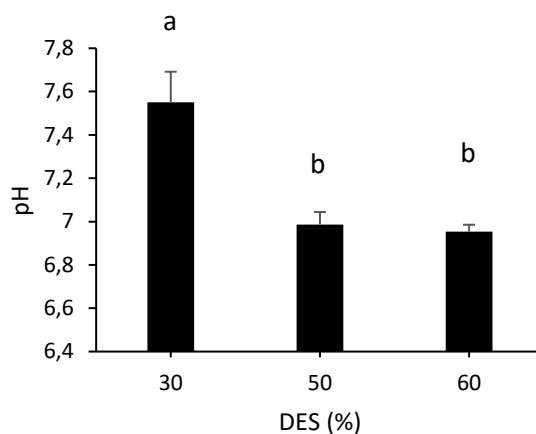


Figure 21. pH of the DES: water mixtures

The error bars indicate the standard deviations of the triplicate samples. Values with different superscripts differ significantly at $p < 0.05$

4.2.2. Model validation

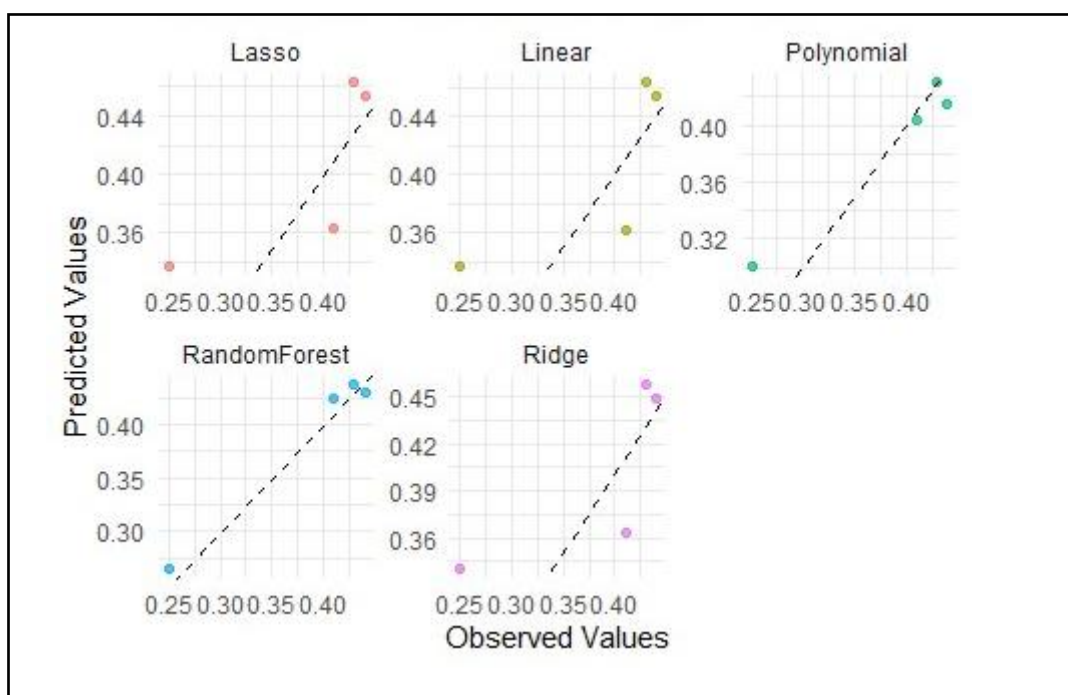


Figure 22. Observed and predicted values estimated by: Lasso, linear, polynomial, RandomForest, Ridge

The predicted nitrogen from soybean flour was compared to the experimental values (Figure 22). To validate the model's accuracy, the normal probability plot of the residuals should show that most residuals fall along a straight line (Nisbet et al., 2018), indicating normal distribution. Residual plots are useful for detecting non-linearity (James et al., 2013). Since RandomForest was the used model, the residuals display a U-shape, indicating the non-linearity of the model (Figure 23).

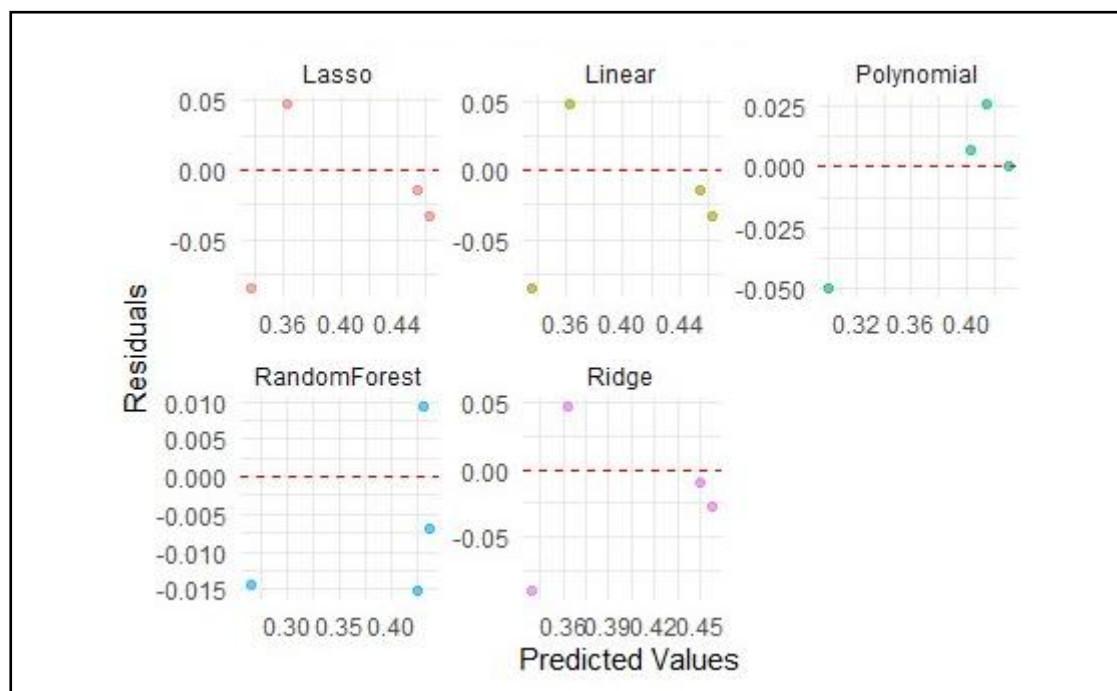


Figure 23. Residual plot estimated by: Lasso, linear, polynomial, RandomForest, Ridge

Following the RandomForest model, the most important factor is the percentage of DES in the solvent (Figure 24). When this ratio attains or surpasses a critical percentage value of 50%, the protein extraction efficiency approaches that of pure water. This likely results from the water supplementation's role in disrupting hydrogen-bonding interactions between DES components. The research of Dai et al, (2015) found that water dilution disrupted hydrogen bonding between the natural DES, completely eliminating these interactions at 50% of water (Dai et al., 2015), transforming the system into an aqueous electrolyte-like mixture (Sapir & Harries, 2020).

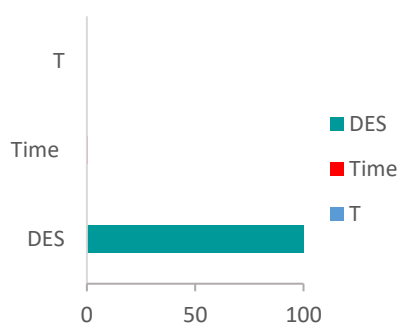


Figure 24. Factor importance of the studied factors (DES, Time, and Temperature (T))

4.2.3. Characterization of the extracted protein

4.2.3.1. Protein extracted using the DES-water solvent

- **Molecular weight profiling of the protein isolate through SDS-PAGE**

As shown in Figure 25, protein bands of both protein isolate samples 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.

These findings are consistent with the literature. The different molecular weights of β -Conglycinin fractions correspond to α' (58–83 kDa), α (58–77 kDa), and β (42–53 kDa), with isoelectric points of 4.90, 5.18, and 5.66–6.00 (Mulalapele & Xi, 2021). On the other hand, the main soybean globulin, glycinin, has a molecular weight that ranges between 22,300, and 37,200 Da (Catsimpoolas et al., 1971). The data reveal the absence of the basic (B) subunits of the 11S protein after the extraction using DES solvents at temperatures of 45 °C and 65 °C. This result contrasts with the results obtained after the extraction using the alkaline/isoelectric precipitation method. This effect can be attributed to the nature of the DES solution and, in particular, its pH. Glycinin is soluble at pH 7.6 with a high ionic strength (0.5 M NaCl), and its solubility decreases when the pH is lowered (Keerati-u-rai & Corredig, 2011). The comparison of these results with earlier studies that used only water for extraction at 90°C for one hour, a band corresponding to the basic fraction of 11S was observed (Hsiao et al., 2015). While the acidic glycinin chain is considered the primary contributor to allergic reactions, the basic glycinin chain can also interact with IgE, potentially triggering an allergic response (Mulalapele & Xi, 2021). For these reasons, the lack of recovery of this fraction of protein may lessen the allergenic properties of the extracted protein.

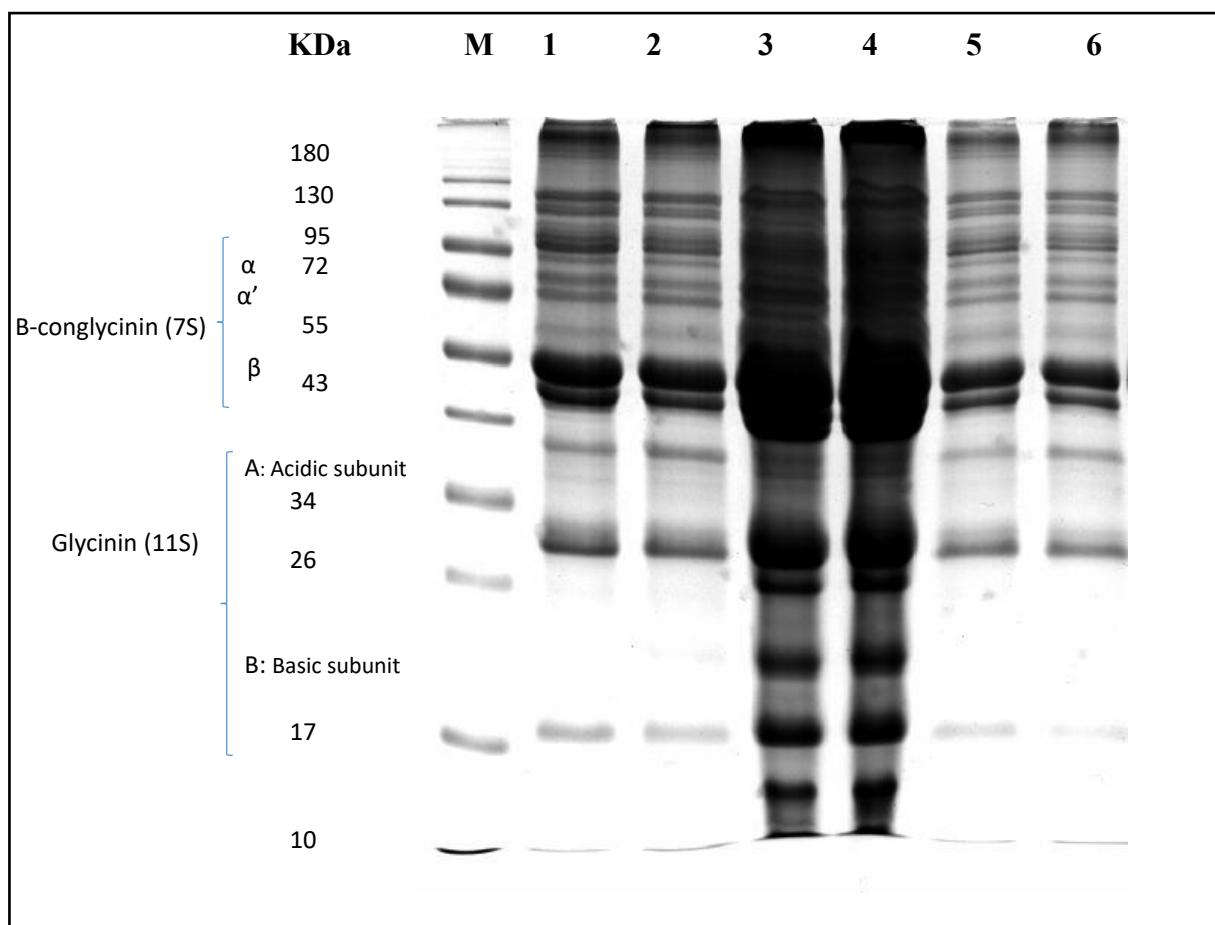


Figure 25: SDS-PAGE visualization of the soybean protein isolate extracted with DES:water (60%:40%) and isoelectric-precipitation method. Lane M: Marker protein; Lane 1,5: Soybean protein (Isidor) extracted using DES:water at 45 °C, and Lane 2,6: Soybean protein (Isidor) extracted using DES:water at 65 °C. Lane 3,4: Soybean protein (Es Mentor) extracted using alkaline-isoelectric method. Native soy globulin: α' , α , and β represent subunits of β -conglycinin; A, and B represent the acid and basic polypeptides of glycinin (N. Chen et al., 2016; G. Shen et al., 2020)

The findings of this research give proof that DES can be used effectively for soybean protein extraction, further substantiating previous research outcomes. For instance, the DES solution (ChCl:Glycerol) successfully allows the extraction of protein from rice brain containing albumin and globulin fractions (Ramlee et al., 2024).

- **Morphological analysis of the protein isolate via light microscopy**

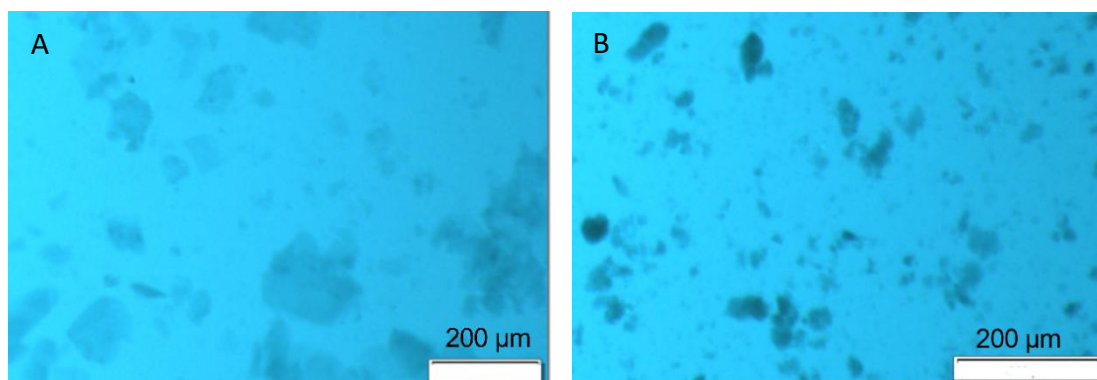


Figure 26. Light microscopy images of soybean protein, stained with Fast Green FCF and extracted with A. water at 45 °C, and B. DES 60% at 45 °C. Scale bar = 200 μm

A prior investigation into whey protein isolates revealed that ethanol treatments at various concentrations (20%, 40%, 60%, and 80%, v/v) induce alterations in the structural, morphological, and functional properties of native isolates (Y. Feng et al., 2021). Although ethanol was employed to facilitate protein precipitation in both extraction methods (DES (60%) or water), the isolate extracted with DES exhibited greater susceptibility, evidenced by a smaller particle size (Figure 26). According to the study by Guo et al, (2023), the initial increase, followed by the decrease in the stability of the enzymes with an increase in water content in DESs, is related to the secondary structural changes (Guo et al., 2023). Likewise, the research of Chen et al., (2021) demonstrated that DES extraction during heating resulted in an irreversible denaturation of soy protein after the use of a pure DES mixture (Q. Chen et al., 2021)(Figure 27).

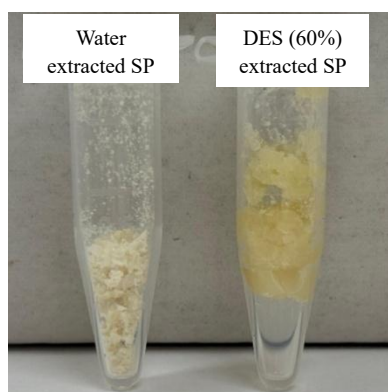


Figure 27. DES extracted soybean protein isolates and water extracted soybean protein

Concerning the effect of DES on the denaturation of the extracted protein, conflicting results were supported by researchers. The study of (Sanchez-Fernandez et al., 2022 b) showed that the storage of lysozyme for 40 days in a 1:2 choline chloride : glycerol solution at room temperature retrieved its native conformation and activity upon rehydration (Sanchez-Fernandez et al., 2022 b). On the other hand, the research of Maity et al., (2020) showed that choline chloride in DES influences water molecules. It increases water density around the protein backbone compared to other systems, helping to preserve the protein's native structural fold.

- **Fluorescence-based characterization of the protein extract**

In this research, the excitation-emission matrix for both proteins exhibits fluorescence at Ex/Em wavelengths of 220–270/340 nm (Figure 28). Our findings are consistent with previous research indicating that fluorescence at Ex/Em of 230–285/335 nm is predominantly linked to protein (Saito et al., 2021). To investigate the effect of solvent on the fluorescence, the fluorescence intensity of the protein extracted under mild conditions (45°C) using 60% DES or water.

Previous research has shown that the emission resulting from the excitation at 220-230 nm (same as 280 nm) is associated with the excitation of specific amino acids, particularly tyrosine and tryptophan. Evaluating the intensity of fluorescence can provide valuable information about the local environment surrounding the aromatic side chains (Bortolotti et al., 2016). Additionally, the research conducted by Kadyan et al, (2019) proved that the quantum yield of fluorescence of L-Trp in the glycine is significantly higher than in water across the entire temperature range (Kadyan et al., 2019).

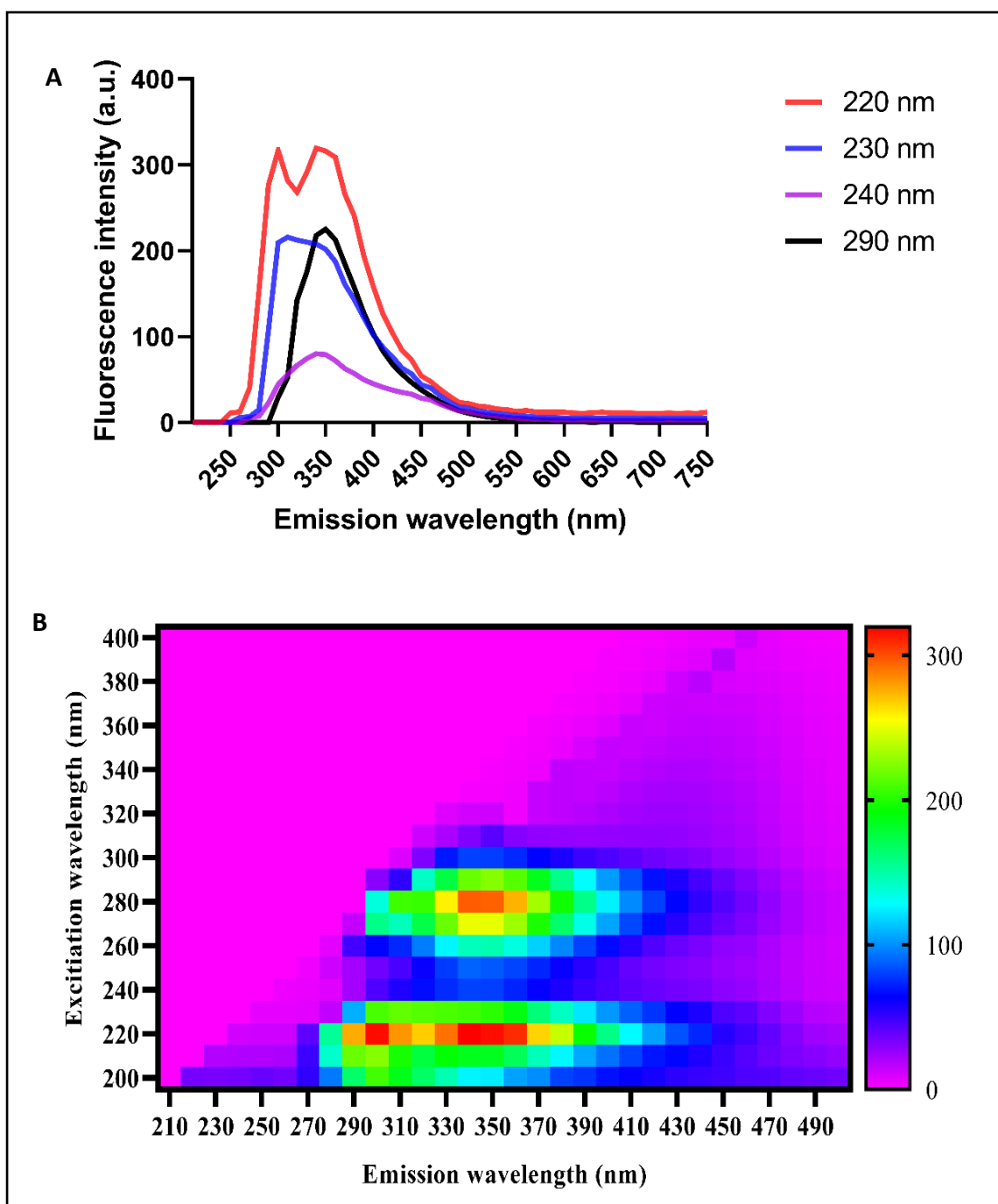


Figure 28. Fluorescent characterization of soybean protein extracted with 60% of DES at 45 °C. A: 2D emission-excitation-intensity fluorescent spectrum, B:3D fluorescence spectra

In contrast, the fluorescence intensity of the protein extracted using water was found to be 4 times higher than that of the protein extracted with 60% of DES (Figure 29).

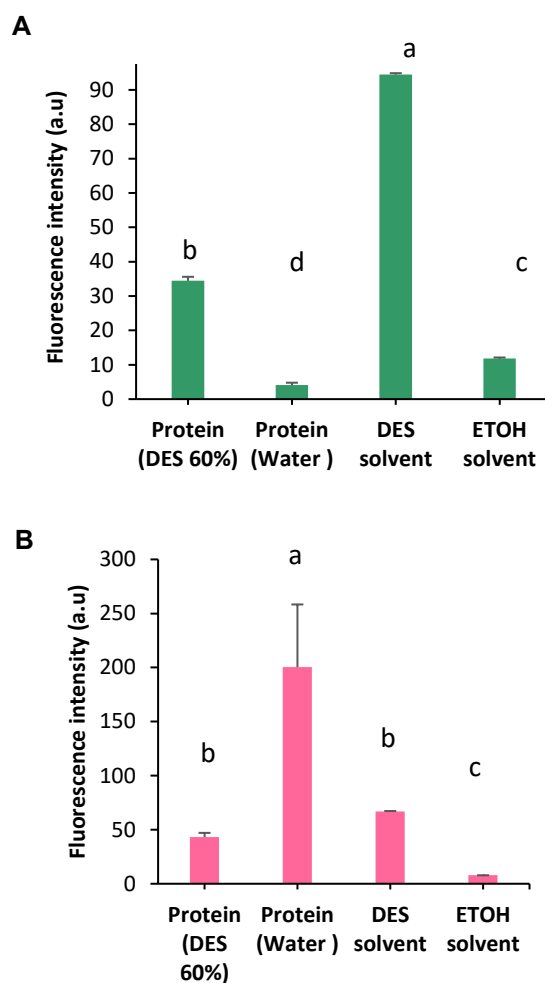


Figure 29. Fluorescence intensity of protein (extracted with 60% of DES, water, and precipitated using ethanol) and of the pure solvents used during the protein extraction at the excitation wavelength/emission wavelength of A: 220/350 nm and B: 280/350 nm (solvent containing 60% of DES and water). The error bars indicate the standard deviations of the samples. Values with different superscripts differ significantly at $p < 0.05$

From a qualitative perspective, there is no evidence indicating the presence of impurities such as oil (Ex/Em of 350/500 nm) (Saito et al., 2021), oxidised lipid (Ex/Em of 365/450 nm) (Figure 29) (J.-H. Liang, 1999). Furthermore, the potential absence of Maillard reaction conjugates can be supported by the low fluorescent intensity within the excitation wavelength range of 340–370 nm and the emission wavelength range of 420–440 nm (Figure 30) (L. Sun et al., 2023).

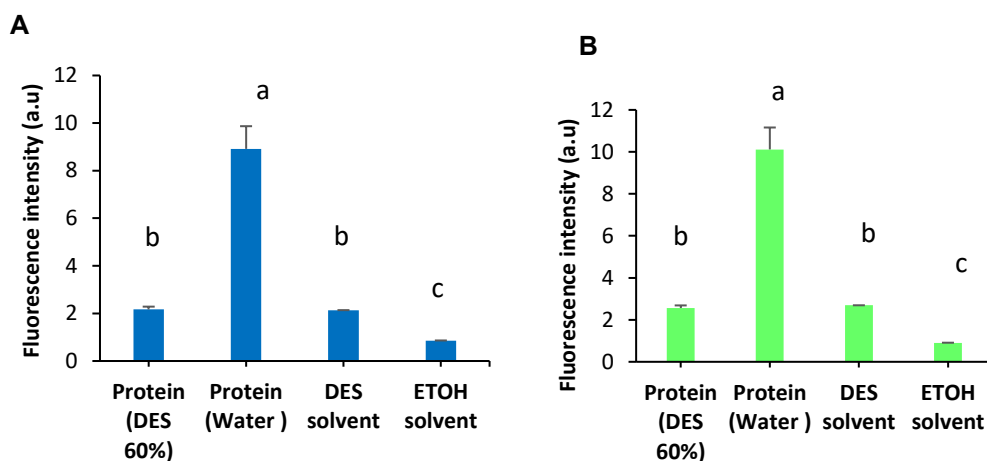


Figure 30. Fluorescence intensity of protein (extracted with 60% of DES, water, and precipitated using ethanol) and pure solvents used during the protein extraction at the excitation wavelength/emission wavelength of: A: 365 /450 nm, B: 350/450 nm. The error bars indicate the standard deviations of the samples. Values with different superscripts differ significantly at $p < 0.05$

4.2.3.2. Characterization of the treated protein isolate

- Molecular weight profiling of the protein isolate through SDS-PAGE

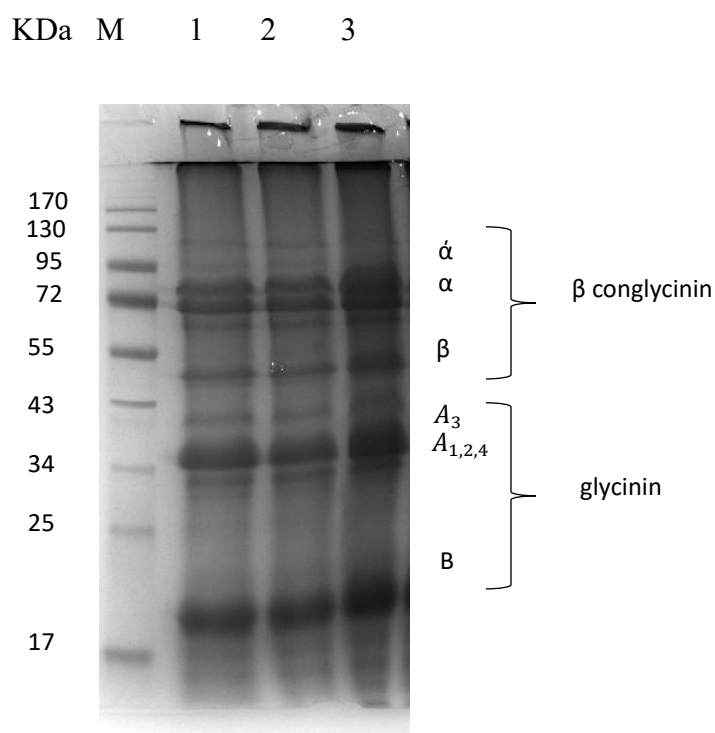


Figure 31. The electrophoretic pattern of native and modified proteins. M: Protein marker (kDa). Lane 1: Heat and ultrasound-treated protein. Lanes: 2: Enzyme (trypsin) treated proteins, 3: Native protein

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 (Figure 31). There was no significant difference between the bands of the native and the modified proteins. In particular, the bands corresponding to ~34 and 95 kDa had the highest intensity. Similar findings were reported by the research of Badar et al, (2024), indicating that control samples showed aggregates with a relative molecular mass above 49.7 kDa.

The similarity in electrophoresis patterns between the native and modified proteins may be due to the addition of salt, which neutralized the effect of the other treatments. Research has shown that preheating induces denaturation of soybean protein isolate, which can lead to changes in its subunits, especially the decline in the content of the 11S globulin subunits (Li et al., 2020). However, the increased concentration of salt reduces electrostatic repulsion on the protein surface and increases hydrophobicity, allowing protein assembly via hydrophobic interactions (Meng et al., 2024).

On the other hand, researchers found that sonication treatment of soybean protein isolate (H. Hu et al., 2013; Badar et al., 2024; Karki et al., 2009; Rahman et al., 2020; X. Wu et al., 2024) did not cause changes in the bands of protein subunits. Similarly, the study of (L. Jiang et al., 2014) showed that the ultrasonication (low-frequency (20 kHz) at various powers) of black-bean protein isolate did not cause significant changes in protein electrophoresis profiles, indicating that the ultrasonic treatment did not alter the primary structure of the proteins (L. Jiang et al., 2014). However, the combination of the ultrasound treatment with the enzymatic treatment induces the process of breaking the soybean protein isolate subunits into smaller units (X. Wu et al., 2024).

In this study, the hydrolysis using trypsin did not induce any changes in the subunits of the protein. A similar result was obtained in the research of Wu et al, (2024), which proved that the use of pepsin hydrolyzed α' , α , β , and 11 S subunits. However, a larger fraction of subunits of the alkaline protease-treated soybean protein isolate sample remained unhydrolyzed (X. Wu et al., 2024). This effect can be related to the degree of hydrolysis of the used enzyme (Ahmadifard et al., 2016).

- **Morphology profiling using microscopy and particle size**

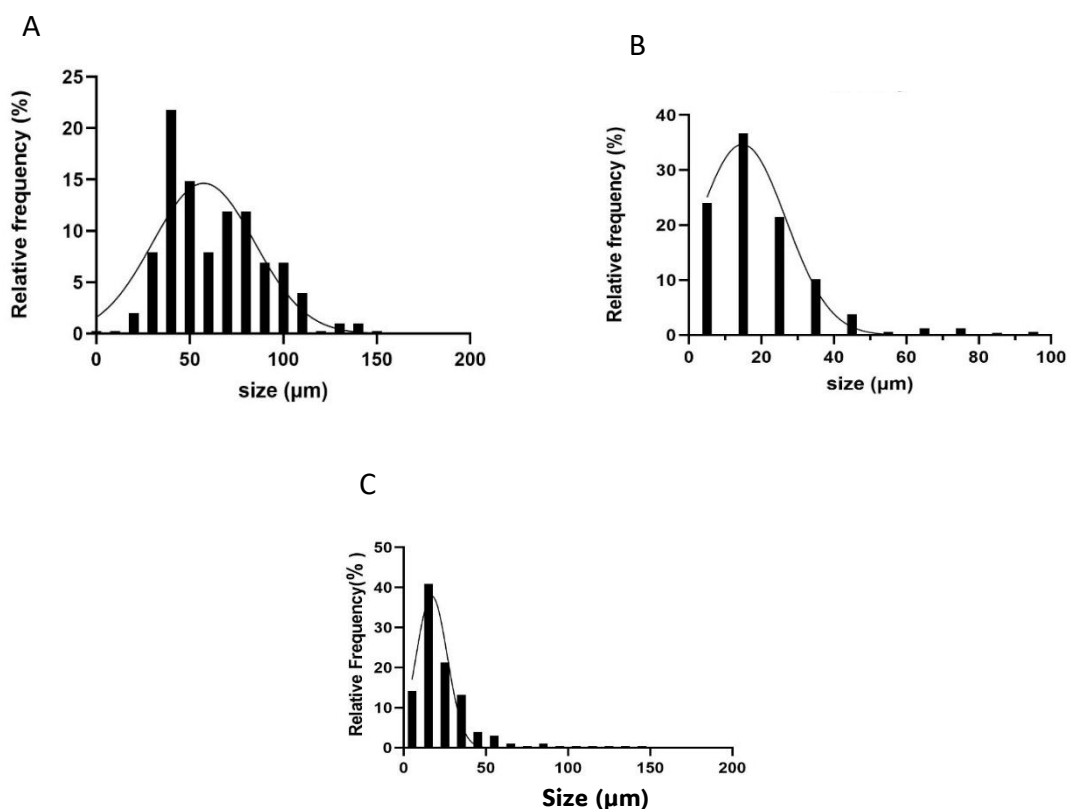


Figure 32. Particle size distribution of the native and modified soybean protein isolate
A: Native soybean protein isolate, B: Soybean protein isolate treated with trypsin, C:
Soybean protein isolate treated with physical treatment

The particle size analysis revealed that the size of the native soybean protein isolate varies between 40–50 μm (Figure 32). Supposed to be in the nanoscale with an average size of around 60–80 nm (P. Shen et al., 2022), the large particle size can be associated with the effect of the extraction process of the native protein. Similar results were found by (T. Zhang et al., 2020) that the protein was turned into spheroidal gel particles (2–20 μm) under the action of heating at 95 °C for 30 min. The formation of the microparticles can be attributed to the considerable number of active sites for aggregation that contributed to the network formation inside the particles (T. Zhang et al., 2020).

Compared to the native soybean protein isolate, which has a spheroidal shape, the treated one with physical treatment has an irregular shape in an aggregated form. This effect can be related to the synergetic effect of the heat treatment, sonication, and salt addition. A previous study showed that the addition of different types of salt (CaSO₄, MgSO₄, or ZnSO₄) at 0–15 mM induces the pre-aggregation of the soybean protein isolate (X. Wang

et al., 2022). In contrast, the heat treatment (95 °C for 30 min) did not generate a significant particle aggregation (T. Zhang et al., 2020).

These changes, when applied under formulation conditions and high protein concentration (~100 mg/ml), may be associated with denaturation since, as an irreversible reaction, protein denaturation is frequently coupled with aggregation and precipitation (Schön et al., 2017). The reduced particle size of the treated soybean protein isolate aligns with the fluorescence results, confirming the denaturation of the treated ones.

- **Fluorescence-based characterization of the treated protein**

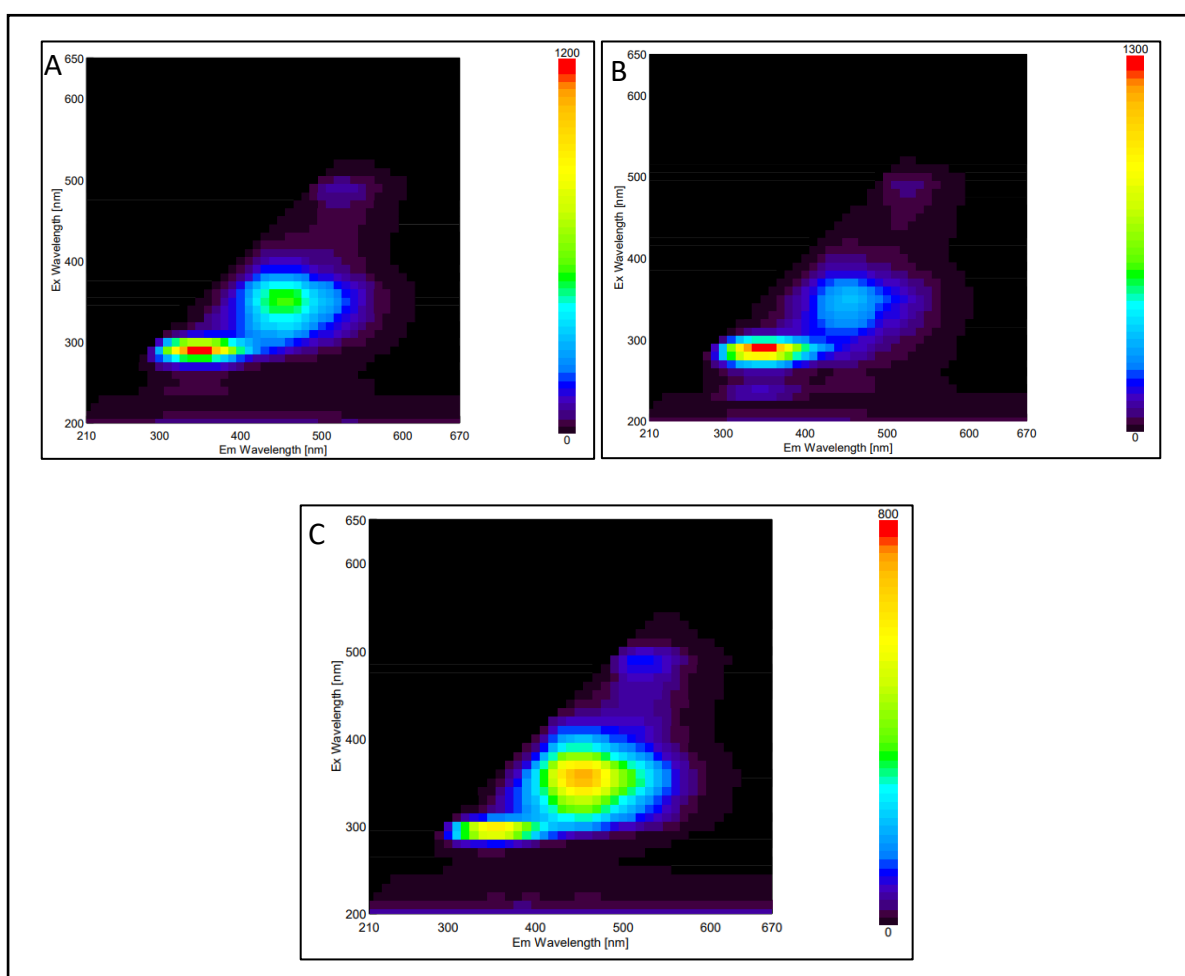


Figure 33. Fluorescence of treated and native soybean protein extract

A: Enzymatic treatment, B: Physical modifications, C: Native soybean protein isolate

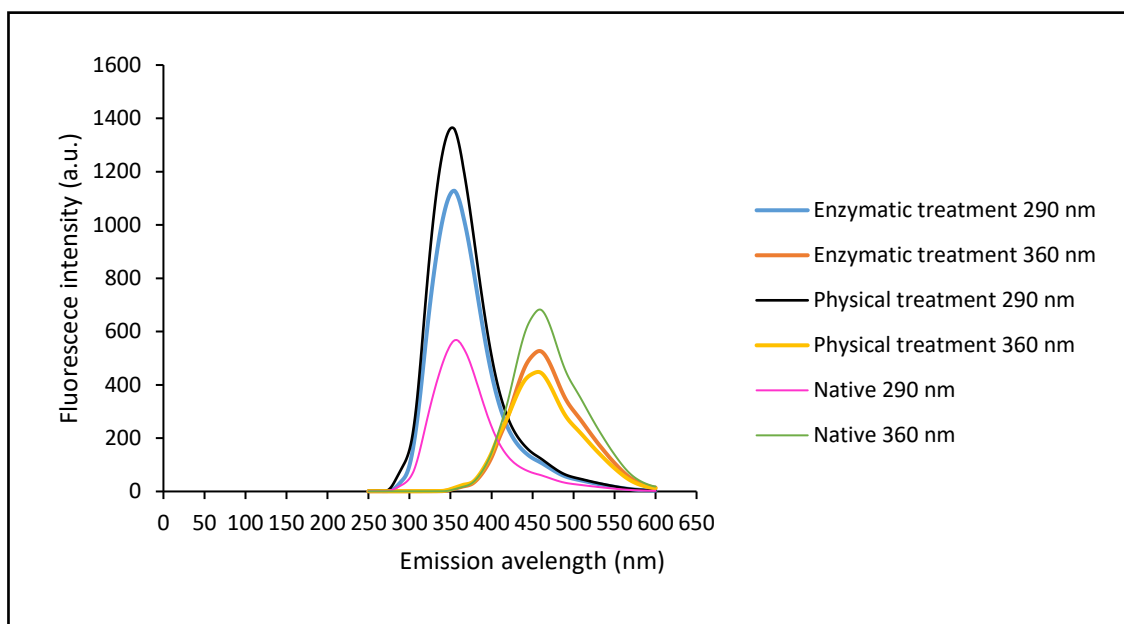


Figure 34. Fluorescence of native and modified soybean protein extracts gained by trypsin treatment, or heat-and ultrasound treatment, compared to the native soybean protein

In general, the emission fluorescence spectroscopic technique can reveal unique information about the sensitivity of intrinsic fluorescence to the microenvironment change around proteins (C. Sun et al., 2016) and, in particular, the polarity of hydrophobic amino acids, reflecting the change in tertiary structure (Z. Ma et al., 2022).

Following the enzymatic and physical treatment, the fluorescence intensity of the emitted light at 357 nm increased significantly (Figures 33, 34, and 35) (Annex 4). This result is consistent with the study of L. Huang et al, (2020) showed that the sonication treatment of the soybean protein isolate increased the fluorescence intensity when excited at 280 nm compared to the control protein (L. Huang et al., 2020). Native and treated isolates subjected to Ohmic heating, with different electrical frequencies (50 Hz–20 kHz) and moderate electric field intensities (up to 20 V/cm), presented a maximum emission at 332 nm when excited at 280 nm (Pereira et al., 2021). Comparing the treatments revealed that the decrease in intrinsic fluorescence and the absence of peak shifts (no changes in the environment's polarity of the fluorophores), suggests that the frequencies between 50 and 500 Hz cause structural rearrangements of the protein fold, leading to fluorescence quenching of the aromatic residues (Pereira et al., 2021).

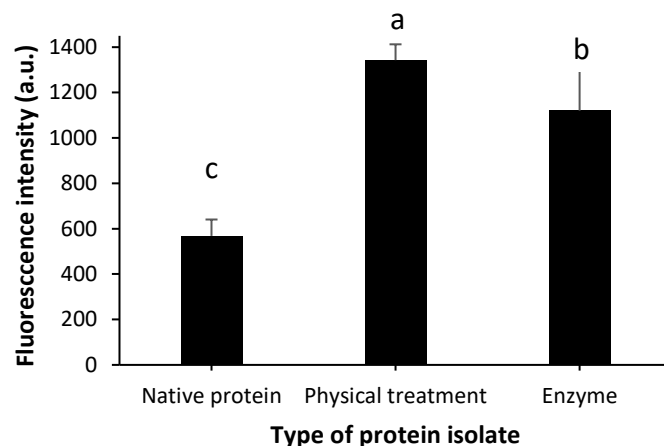


Figure 35. Fluorescence of native and treated soybean protein extract (excitation:290 nm, emission: 357nm). The error bars indicate the standard deviations of the samples.

Values with different superscripts differ significantly at $p < 0.05$

Similarly, other studies linked increased fluorescence intensity with the thermal treatment of various protein isolates, such as zein (C. Sun et al., 2016), and pea protein isolate (Chao & Aluko, 2018), under treatments of 95°C and 100°C (pH 7), respectively. These conditions expose 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; A. Zhang et al., 2021), increasing the fluorescence intensity (A. Zhang et al., 2021). Trp is selectively excited at 295 nm, since other residues that have a little absorption when excited at this wavelength, for an emission spectrum detected at 355 nm (Yammine et al., 2019). It is characterized by its high sensitivity to its local environment (Lakowicz, 2006). Trp emission spectra are susceptible to changes since conformational transitions, subunit association, substrate binding, or denaturation can affect the local environment surrounding the indole ring (Lakowicz, 2006). Since the sample was subjected to three different physical treatments, and 3 of them are going in just one direction, which is the increase in the intensity of fluorescence, the observed effect can be the synergic effect of these treatments. In this case, the phenomenon of salt-enhanced fluorescence, which relates the increase in intensity to the salt concentration, can also be behind the effect on the protein isolate. Research revealed that the addition of salt is behind the conformation change. In particular, some amino acid residues, such as arginine and asparagine, which are responsible for the suppression of fluorescence, have their 3D structural features changed, leading to a fluorescence intensity enhancement (F. dos Santos Rodrigues et al.,

2023). In contrast, this effect varies depending on the type and the quantity of salt. For example, the introduction of Ag⁺ ions to BSA, induces a decrease in the fluorescence of Trp along with a shift in emission wavelength towards the longer wavelength (Senthil et al., 2020). In contrast and similar to our study, the fluorescence intensity of whey protein isolate increased with increasing NaCl concentration. This effect can be associated with the improved solubility of the protein and exposure of more tryptophan residues. Additionally, the tertiary structure of protein was also modified to change the charge density on molecular surface (Zhong et al., 2022).

Likewise, the fluorescence intensity of the enzyme-treated soybean protein isolate was higher than that of the native isolate. Similar results were found using different enzymes, such as papain (Z. Ma et al., 2022).

In our study, no distinct red or blue shift occurred for the fluorescence spectrum of soybean protein isolate upon the different types of modification (Figure 36). Likewise, L. Huang et al, (2020) showed that ultrasonic pre-treatments did not shift the maximum wavelength of the emission fluorescence. In contrast, a redshift in the spectrum of soybean protein isolate due to an enzymatic treatment was reported by the study of Z. Ma et al, (2022). The appearance of red shift to wavelengths >350 nm was related to the exposure of the tryptophan residues to a hydrophilic environment (Chao & Aluko, 2018). On the other hand, the blue shift in the case of degraded glycinin hydrolysate of soybean protein isolate was related to the change in the microenvironment of Trp residues that became less polar (Xia et al., 2020).

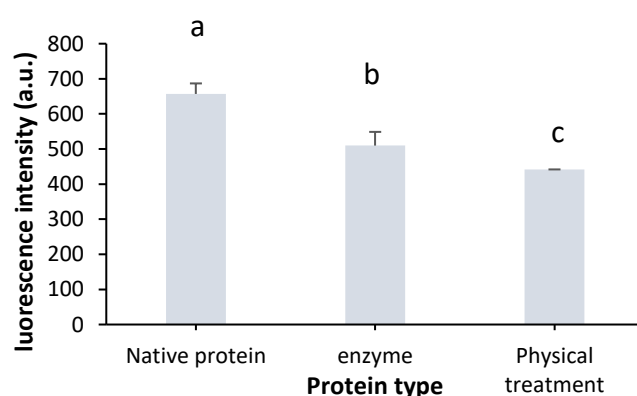


Figure 36. Fluorescence of native and treated soybean protein extract (excitation:360 nm, emission: 450 nm). The error bars indicate the standard deviations of the samples.

Values with different superscripts differ significantly at $p < 0.05$.

The native soybean protein isolate has $80.61 \pm 0.61\%$ of protein, the rest, which represents 20% includes non-fluorescent components such as lipids, membranes, and saccharides, and components with weak fluorescence, such as the DNA (Lakowicz, 2006). In addition to the protein, the fluorescence can be generated by an in vitro reaction of proteins with oxidized lipids. The generated fluorescent and cross-linked proteins are similar to those in the aging cells or tissues (Kikugawa & Beppu, 1987). 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 (J.-H. Liang, 1999). For these reasons, the reduction of the fluorescence intensity (Ex:360 nm /Em: 440 nm) of the treated samples (Figure 37) can be related to the removal of the oxidised oil by the addition of water and centrifugation.

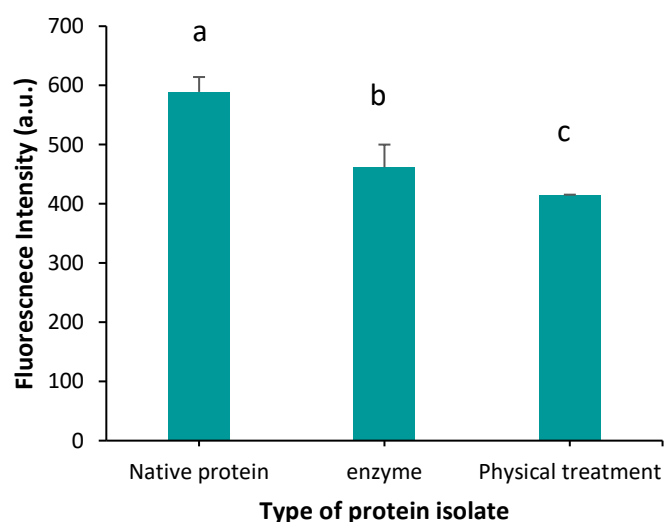


Figure 37. Fluorescence of native and treated soybean protein extract (excitation:360 nm, emission: 440 nm). The error bars indicate the standard deviations of at least triplicate samples. Values with different superscripts differ significantly at $p < 0.05$.

As shown in Figure 37, 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.

The decrease in the intensity can be explained by their elimination after washing out with water, since, in general, the solubility of proteins can be improved through the Maillard reaction with polysaccharides (Dumitraşcu et al., 2023). This effect can be related to the high solubility of protein-sugar conjugates. For instance, an enhancement in the solubility was observed in the study of Kutzli et al, (2020) and Nooshkam & Madadlou, (2016)

wherein conjugates of pea protein isolate-maltodextrin and whey protein isolate-sugar were investigated.

4.3. Assessment of the nutritional effect of soybean protein and lysine

4.3.1. *Drosophila* model

4.3.1.1. Effect of the soybean protein and lysine on the development of *Drosophila*

Assessing the quantity and quality of the diet is crucial to preventing protein-energy malnutrition, which contributes to poor growth, diminished mental development, and illness (Neumann et al., 2002). To examine the performance of a specific diet, a standard diet that promotes normal growth and a stressful diet can be taken as references.

- **Effect of addition of SP on 0M**

The addition of 5-fold in sucrose showed a 2-day delay compared to NM (Figure 38). The effect of a high sugar diet on the delay in pupation was confirmed by the study of (Rovenko et al., 2015) and (Loreto et al., 2021), which also reported a reduction of viability in fruit fly larvae hatched from a 30% sucrose-containing medium (Loreto et al., 2021).

These findings highlight the essential role of the larval phase in overall development as the pupal stage undergoes substantial metabolic reorganization without involving new nutrients. Consequently, the quality and quantity of resources accumulated during the larval stage are essential for successful pupal development (Boatta et al., 2024). In addition to *Drosophila*, during the larval stage, feeding *B. dorsalis* and the first generation of housefly *Musca domestica* with a high-sucrose diet exhibited a longer developmental period (Boatta et al., 2024; L. Wang et al., 2023). The delay in the developmental time of pupae reared on high-sugar diets has also been previously demonstrated in *Drosophila* (Rovenko et al., 2015). It can be attributed to the increased levels of hemolymph and body glucose, which make the hyperglycemic phenotype the reason for developmental delay and increased pupa mortality in glucose-fed flies (Rovenko et al., 2015). On the other hand, the effect of sucrose is dose-dependent. For example, feeding the Egyptian Cotton Leafworm, *Spodoptera littoralis*, with a high-sugar diet (5% of sucrose) did not affect the larval development time (Doğan et al., 2021).

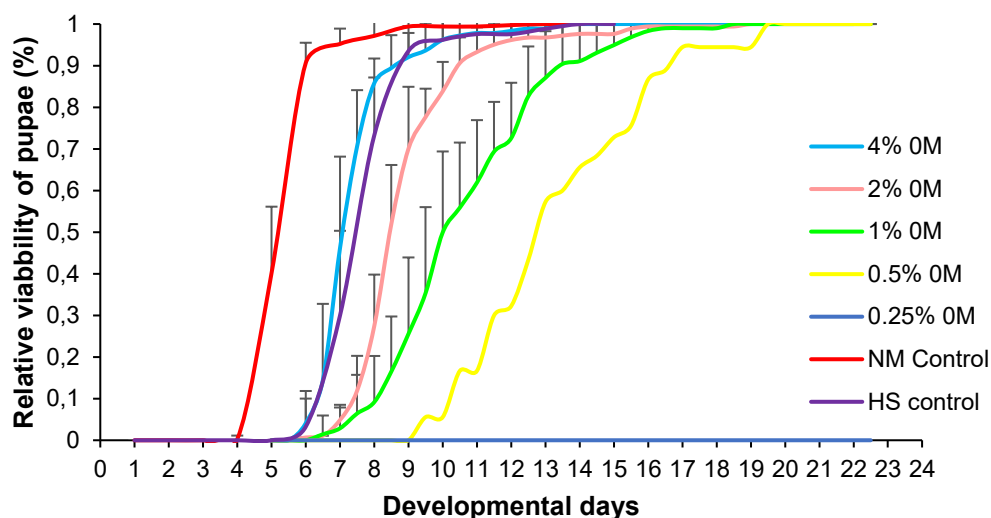


Figure 38. Progress of pupae's emergence after the addition of different percentages of SP in OM over time. The error bars indicate the standard deviations of at least triplicate samples

These results highlight the critical role of the larval phase during development. During this phase, the quality of the accumulated resources is important to ensure a successful pupal development (Boatta et al., 2024). In general, in holometabolous insects, growth ceases at the onset of metamorphosis and depends on achieving the critical size in the final larval instar (Stieper et al., 2008).

The addition of soybean protein isolate that contains $79.89\% \pm 0.7$ of crude protein to the OM environment promoted normal development of larvae and pupae at 2 and 4% of SP in OM (Figure 38). A delay comparable to HS conditions was observed when 4% of SP/ was added to OM. A more pronounced delay occurred with the reduction of the percentage of supplemented SP. The cumulative distribution curve illustrates a delay in hatching associated with an intermittent apparition of the pupae (Figure 39). Conversely, low SP levels (0.25% and 0.5% OM) did not ensure the proper development of the larvae and, thus, fly hatching (Figure 39).

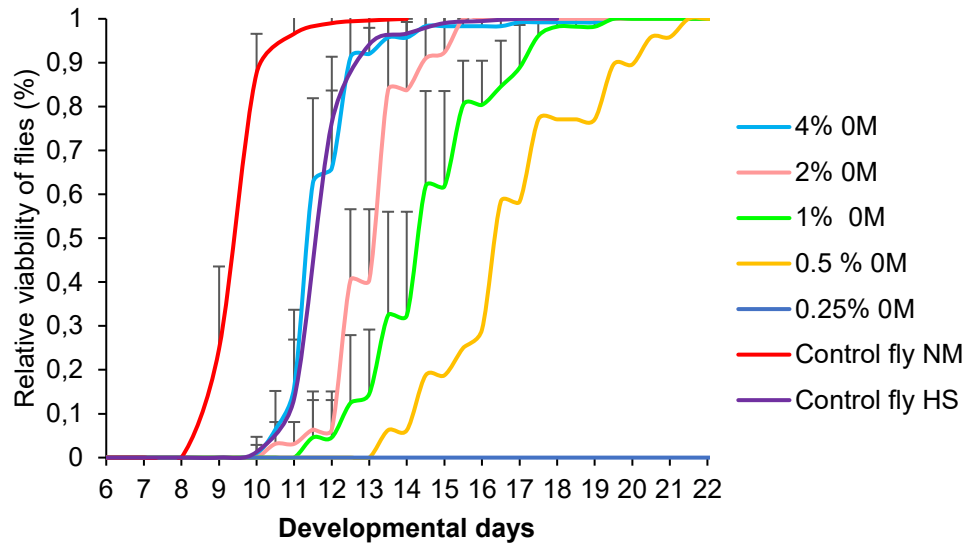


Figure 39. Progress of fly's emergence after the addition of different percentages of SP in 0M over time. The error bars indicate the standard deviations of at least triplicate samples

This effect is consistent with established knowledge that the starvation effect is always associated with the prolongation of larval life (Tatum & Beadle, 1939) and that *Drosophila*'s larvae tend to target their consumption to P:C ratios that minimize development time, which correspond to the ratios 1:2 (M. A. Rodrigues et al., 2015).

In 0M, 0.25 and 0.5% of Lysine (data not shown), 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 Lys on 0M environment**

Compared to the samples supplemented with just SP, the viability of the larvae reared in media containing 75% and 50% of lysine (total concentration of 2% 0M and 1%0M) dropped to 10% and 13.7%, respectively. This indicates that the lysine addition did not support larval development in reaching the pupation phase. Further experiments confirmed that the larval viability decreased further, preventing any larvae from pupating, with the increase of the percentage of the supplemented lysine to reach 50% and 75%. These results indicate that lysine addition disrupts larval development in a protein-amino

acid-restricted diet. The study of (Xiao et al., 2023) highlighted the negative effect of lysine supplementation in a lysine-sufficient diet since it can decrease the protein quality of the diets and modify the serum concentrations of many amino acids, affecting the growth and the utilisation of nutrients due to amino acid imbalance or antagonism (Xiao et al., 2023). These results align with our study since soybeans contain well-balanced essential amino acids except sulfur-containing ones (Qin et al., 2022), with a moderate lysine content, representing 6.8% (Nill, 2016). On the other hand, an experiment conducted on broilers demonstrated lysine's non-crucial role in animal growth. In particular, lysine restriction during various feeding phases exerts negligible effects on the broilers' production performances due to their compensatory growth (Soglia et al., 2021).

The larvae's viability decreased to 10% and 13.7% when the supplementation of SP was reduced from 100% to 75% and 50% by adding lysine while maintaining total concentrations of SP and lysine at 2% 0M and 1% 0M, respectively. This indicates that added lysine did not support larval development towards pupation. Further experiments confirmed that as lysine supplementation increased to 50% and 75%, larval viability continued to decrease, completely preventing pupation. These findings indicate that lysine incorporation disrupts larval development in a diet limited in protein and amino acids. Xiao et al. (2023) also documented the adverse effects of lysine supplementation in lysine-sufficient diets, which can diminish protein quality and alter serum amino acid concentrations. The excess of lysine negatively impacts growth and nutrient utilization due to amino acid imbalance or antagonism (Xiao et al., 2023). This correlates with our findings, as soybeans provide a well-balanced profile of essential amino acids, with the exception of sulfur-containing ones (Qin et al., 2022), and possess a moderate lysine content of 6.8% (Nill, 2016). Conversely, an experiment with broilers showed that lysine is not essential for animal growth, as its restriction at different feeding phases had minimal impacts on production performance, due to compensatory growth (Soglia et al., 2021).

- **Effect SP on the NM environment**

In the NM, when the protein-to-sugar ratio was 1.34, adding 0.25, 0.5, 1, and 2% of SP did not change the development time compared to the control (Figure 40). In contrast, the incorporation of 4% SP significantly slowed the larvae and pupae's development time, delaying the fly's appearance (Figure 39).

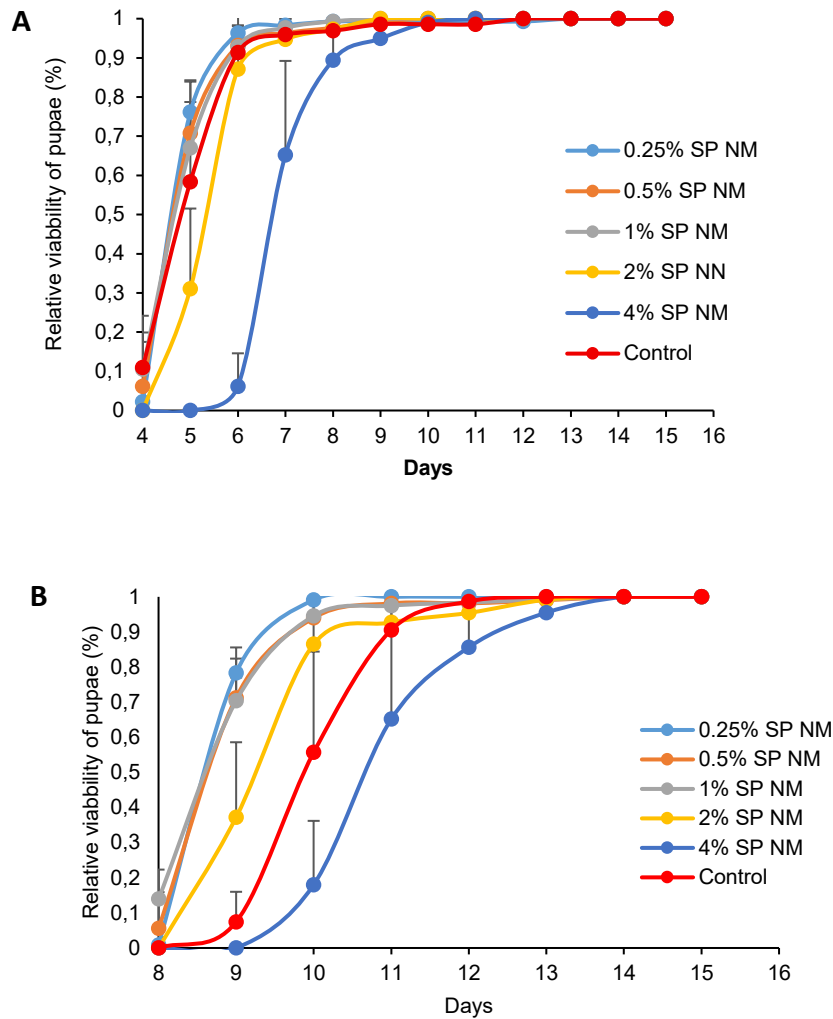


Figure 40. Progress of: A. pupae and B. flies emergence after the addition of different percentages of SP in NM over time. The error bars indicate the standard deviations of at least triplicate samples

A study on *H. vigintioctopunctata* larvae demonstrated the detrimental effect of protein deficit on larval survival compared to carbohydrate deficits (Z.-L. Wang et al., 2018). Preserving a constant protein content while increasing the proportion of carbohydrates affected the larval developmental duration, preventing them from completing the larval stage. In contrast, when the carbohydrate content remained constant, a higher percentage of protein content induced a shorter larval development time (Z.-L. Wang et al., 2018).

The increase in the sugar relative to protein content in the media can significantly affect the time, mainly by prolonging the larval feeding period (Matzkin et al., 2011). For *Drosophila*, this effect was significant solely for the low protein-sugar ratio diet containing a protein-to-carbohydrate ratio equivalent to 2.5%:24.6% where the pupation

was delayed by 1 day. However, for *D. mojavensis*, a diet with an equal protein: sugar ratio containing a protein-to-carbohydrate ratio equal to 4.3%:21.2%, delayed pupation by 6 days. This can be attributed to the need for some minimum protein level to be consumed for larvae to pupate, which seems less restrictive for *D. melanogaster* (Matzkin et al., 2011). The delay in the development of *Drosophila* following the increase of sugar content, SP amount in NM (4%) or in the 0M (Figure 40) can be related to disruption or reduced expression of ecdysone hormone (molting hormone) such as 20-hydroxyecdysone (20-E) that is required from the early embryonic stage to the late larval, pupal, and adult stages (A. Kumar et al., 2013). These Ecdysteroids (Pan et al., 2021), juvenile hormones are essential to regulate insect molting and metamorphosis (Hiruma & Kaneko, 2013). Four different classes of ecdysteroids, which are obligatorily derived from four types of dietary sterol precursors, can be produced by *Drosophila*. *Drosophila* makes makisterone A, from plant sterols and epi-makisterone A, which supports only larval development from ergosterol, the major yeast sterol. However, they prefer to selectively utilize scarce ergosterol precursors to make a novel hormone 24, 28 dehydromakisterone A, and trace cholesterol to synthesize 20-hydroxyecdysone (Lavrynenko et al., 2015).

- **Effect of Lysine on NM Environment**

During the larval stage, *Drosophila*'s larvae must adapt their metabolism to attain a critical body size to pupate (Manière et al., 2020). Nutrients such as amino acids serve as a major regulator to control metabolism and growth. They exert mainly an effect on the fat body and the brain to control larval growth, body size, developmental timing, and pupariation (Manière et al., 2020). With normal diet conditions, the supplementation of lysine in low concentration (0.25 and 0.5 g/100 g NM) did not significantly influence the appearance time of the first pupae (Figure 41).

Similar results were obtained for other insects, such as *Hermetia illucens* (Koethe et al., 2022). Koethe et al. (2022) found that a lysine content of 0.3% in the substrate promotes optimal larval growth of *Hermetia illucens*. In contrast, Spranghers et al. (2024) found there is no notable growth improvement beyond a lysine content of 0.53 g/100 g of diet (Spranghers et al., 2024). This leads to the conclusion that the optimal dose may also vary depending on the insect species. For example, for *Tenebrio molitor*, the limit is set at 0.35 g/100 g diet (Spranghers et al., 2024).

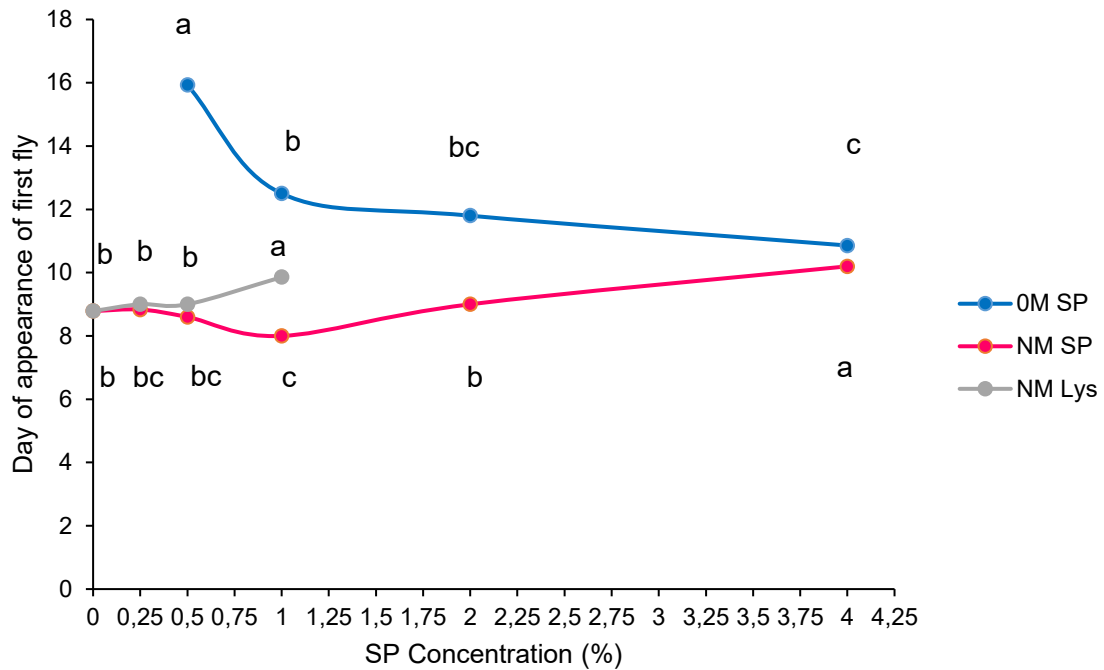


Figure 41. Time required for the appearance of the first fly in NM and 0M containing different percentages of SP and lysine. Values with different superscripts differ significantly at $p < 0.05$

Under normal nutrition conditions, the high concentration of supplemented lysine (2% and 4%) hindered the development of larvae and the pupae hatching (Figure 42).

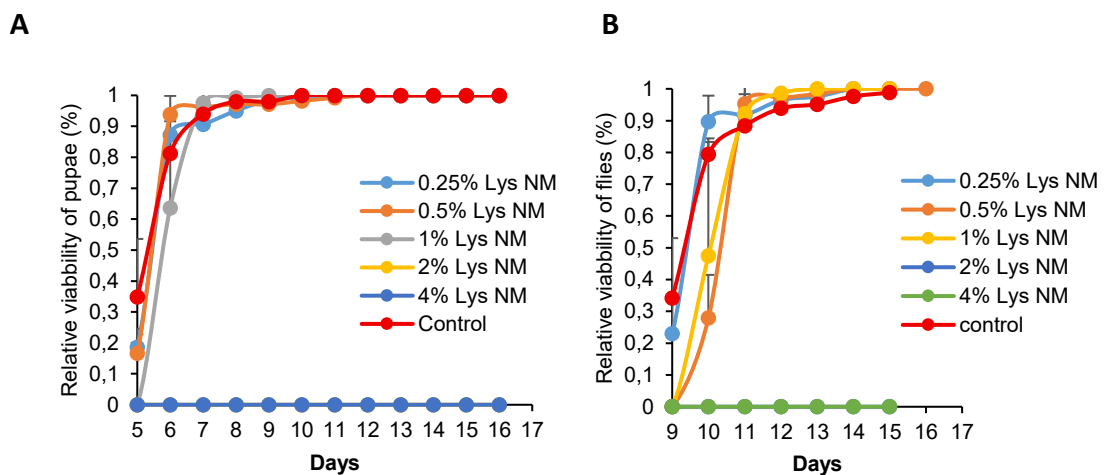


Figure 42. Progress of A. pupae and B. flies emergence after the addition of different percentages of Lysine in NM over time. The error bars indicate the standard deviations of at least triplicate samples

In particular, out of the 250 embryos of the group supplemented with 2% in NM, only a few pupae (1-3) could develop, but their development was delayed compared to the control group. Similar results were obtained by other insects, such as *H. illucens*. The addition of lysine did not improve their larvae's growth, development, or nutritional values. In contrast, it reduced larval size, survival, and pre-pupal rates, especially when high amounts of lysine were provided (Koethe et al., 2022). For *Drosophila*, these concentrations appear toxic, suggesting their potential application as pesticides similar to phenylalanine, which have been known for many years in synthesizing compounds associated with plant defence (Y. Chen et al., 2023). The toxicity of lysine to some groups of bacterial cells was proved. However, the specific mechanisms behind this effect remain unclear. For instance, some cyanobacteria, such as *Microcystis aeruginosa*, cannot efficiently export and degrade lysine (Kim et al., 2023)

An excess of dietary L-lysine, either as the hydrochloride or free base, depressed the growth rate and produced toxicity symptoms in 14- to 21-day-old chicks fed adequate protein diets (Jones, 1961). On the other hand, the supplementation of lysine to juvenile *GIFT tilapia* improved growth performances and feed utilization. It enhanced the relative expression of MyoD and Myogenin and influenced haematological and biochemical responses. However, dietary lysine supplementation did not affect the whole-body chemical composition and amino acid profile (Prabu et al., 2020).

- **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).

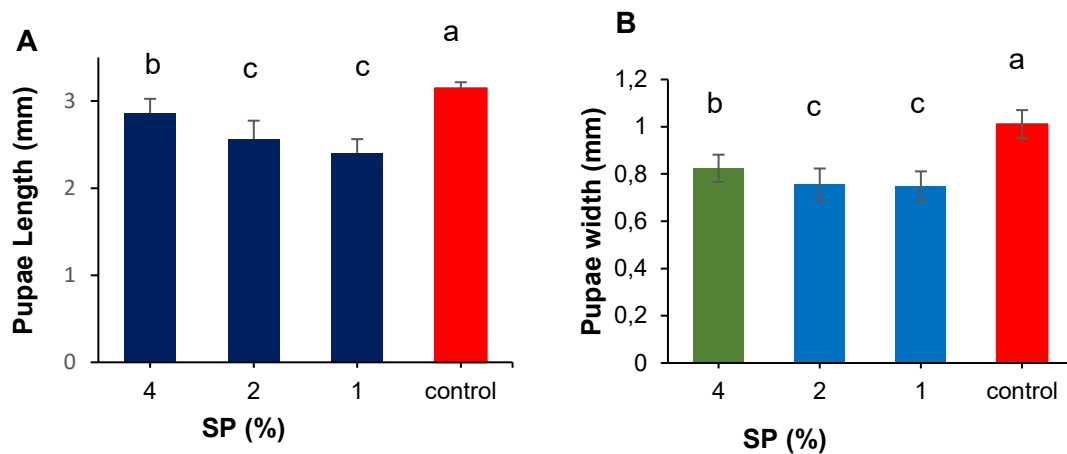


Figure 43. Pupal length (in mm) (A) and width (B) reared on different diets. Different letters denote significant differences based on estimated marginal mean comparisons (p value < 0.05). Control: NM. Values with different superscripts differ significantly at $p < 0.05$. The error bars indicate the standard deviations of the samples

The analysis of the pupae size that emerged in 0M with a variable SP concentration showed a significant difference in the average length and width (Figure 43). The largest pupae were found in NM (control), while the smallest pupae were observed in the 0M supplemented with (1% and 2% SP in 0M). At low SP concentrations (0,5%, 0.25% of SP in 0M), a limited number of pupae appeared, making it challenging to generalise the findings about their size. A restricted nutritional diet that mediates starvation (larvae fed on the standard medium for 2 or 3 days, then maintained without sugar or yeast), crowding, or reduced nutrient content (no sugar media) led to a decrease in the pupae size compared to the high-sugar and standard medium. Likewise, a decrease in yeast content in the *Drosophila* larval diet prolonged developmental time and reduced body size, mimicking the effects of larval crowding (Klepsatel et al., 2018). On the other hand, an experiment conducted on *Manduca sexta* (Sphingidae: Lepidoptera) showed that the larvae fed on low-water and low-protein diets (replacing 40% of the casein in the standard diet with α -cellulose) initially grew slowly, but accelerated in the fifth instar, resulting in pupae and adults sized comparably to control animals (Ojeda-Avila et al., 2003).

In this study, the reduced protein content created a crowding-like environment condition, which can be considered a complex multifactorial stressor that could elucidate the effect of these treatments on pupae size. A cannibalistic behaviour displayed by larvae was noticed under these conditions: by eating deceased flies and devouring pupae, making them flabby. Under normal conditions, *Drosophila melanogaster* larvae are known as

herbivores and feed on a non-carnivorous diet. However, under nutritionally challenged conditions, they can exhibit cannibalistic behaviour by consuming a diet composed of larger conspecifics and also their own embryos (Ahmad et al., 2015). This predation on conspecifics' behavior is viewed as an adaptive behavior, which can rapidly evolve in response to nutritional conditions (Vijendravarma et al., 2013). On the other hand, in *Drosophila*, the cannibalism behavior was associated with other factors, such as intraspecific competition that may cause stress and affect survival and reproduction (Bezerra Da Silva et al., 2019). Conditions such as malnutrition evolved an enhanced propensity towards cannibalism, which led to the plasticity of larval mouth parts (Vijendravarma et al., 2013).

4.3.1.2. Effect of diet on *Drosophila*'s eye color

Environmental conditions and diet can trigger epimutations and thus influence gene activity (Hoyer-Fender, 2020). 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 %) was also identified (Figure 44,45,46, 47 and 48).

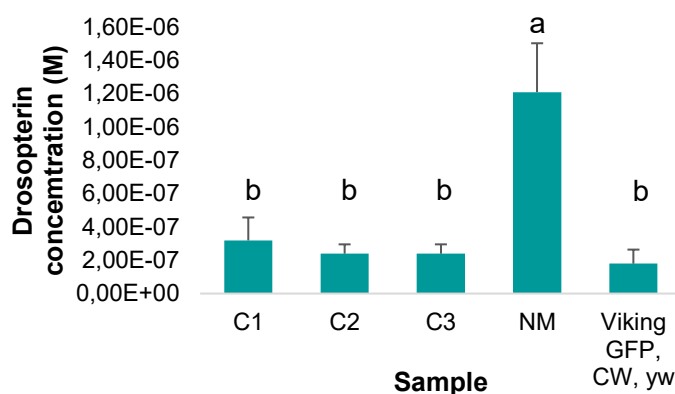


Figure 44. Drosopterin concentration of W^{m4h} *Drosophila* eyes reared in C1: 0M with 4% SP isolate, C2: 0M with 2% SP isolate, C3: 0M with 1% SP isolate, NM (Control), and of *Drosophila* mutation Viking GFP (white eyes, Curly wings, yellow body) reared in NM. The error bars indicate the standard deviations of the samples.

A confirmatory analysis indicated that SP isolate, independent of its concentration, was able to modify PEV phenotype and influence heterochromatin formation in cells. This effect contradicts the conclusion reached by the research of Tatum & Beadle, (1939),

which stated that the prolongation of life induced by starvation is possible under certain conditions without any modification of eye color (Tatum & Beadle, 1939). In contrast, Michailidis et al. (1988) found that the extent of eye-colour variegation in w^{m4h} *Drosophila melanogaster* depends on development time. They demonstrated that when the pH varies between 2 and 6, the development becomes slow and exhibits a more pronounced variegation than those reared at higher pH (Michailidis et al., 1988).

Researchers have proved that environmental constraints, such as an unbalanced prenatal nutrition (protein-restricted), induce persistent, gene-specific epigenetic changes that alter mRNA expression (Lillycrop et al., 2005).

In particular, it has been shown that amino acids by themselves can modify the expression of target genes (Chaveroux et al., 2010), which illustrates the organism's need to adapt to a changing nutritional environment (Chaveroux et al., 2010). On the other hand, research on dietary supplementation that was long presumed to be purely beneficial, revealed its potential to induce an unintended deleterious influence on the establishment of epigenetic gene regulation in humans (Waterland & Jirtle, 2003). It has been proven that the dietary methyl supplementation of a/a dams with extra folic acid, vitamin B₁₂, choline, and betaine alters the phenotype of their A^{vy}/a offspring (Waterland & Jirtle, 2003). On the other hand, a conclusion drawn from a study using Tranylecypromine, which is used to treat major depressive disorder, suggested that food compounds potentially affect chromatin modification, and hence gene expression, and that the alteration is putatively inherited not only parentally but transgenerationally (Hoyer-Fender, 2020). In these studies, different models were used, including mammals such as mice (Waterland & Jirtle, 2003) and rats (Lillycrop et al., 2005), and fruit flies using *Drosophila melanogaster* w^{m4} strain (Hoyer-Fender, 2020).

The results showed that within the same group of flies, the variability of the intensity of the eye colour and the distribution of the red colour was demonstrated. Although they share a common genotype, epigenetic differences can explain this difference. For humans, research found that although twins are epigenetically indistinguishable during the early years of life, older monozygous twins exhibited remarkable differences in their overall content and genomic distribution of 5-methylcytosine DNA and histone acetylation, affecting their gene-expression portrait (Fraga et al., 2005).

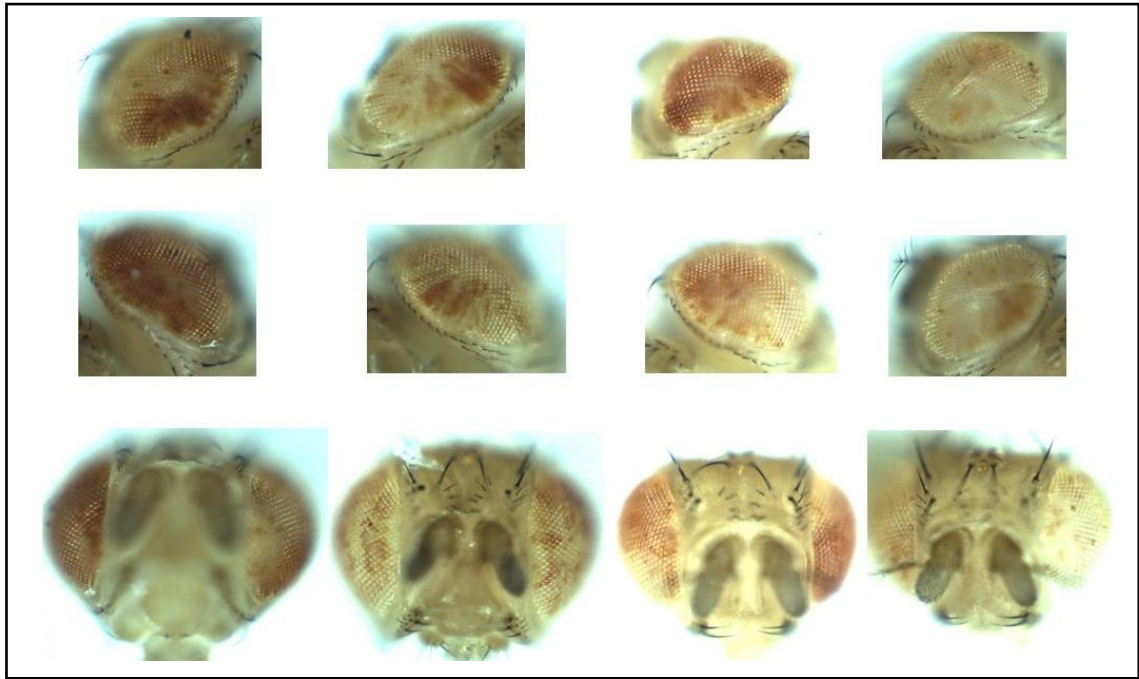


Figure 45. *Drosophila* Heads reared in 0M supplemented with 4% of SP

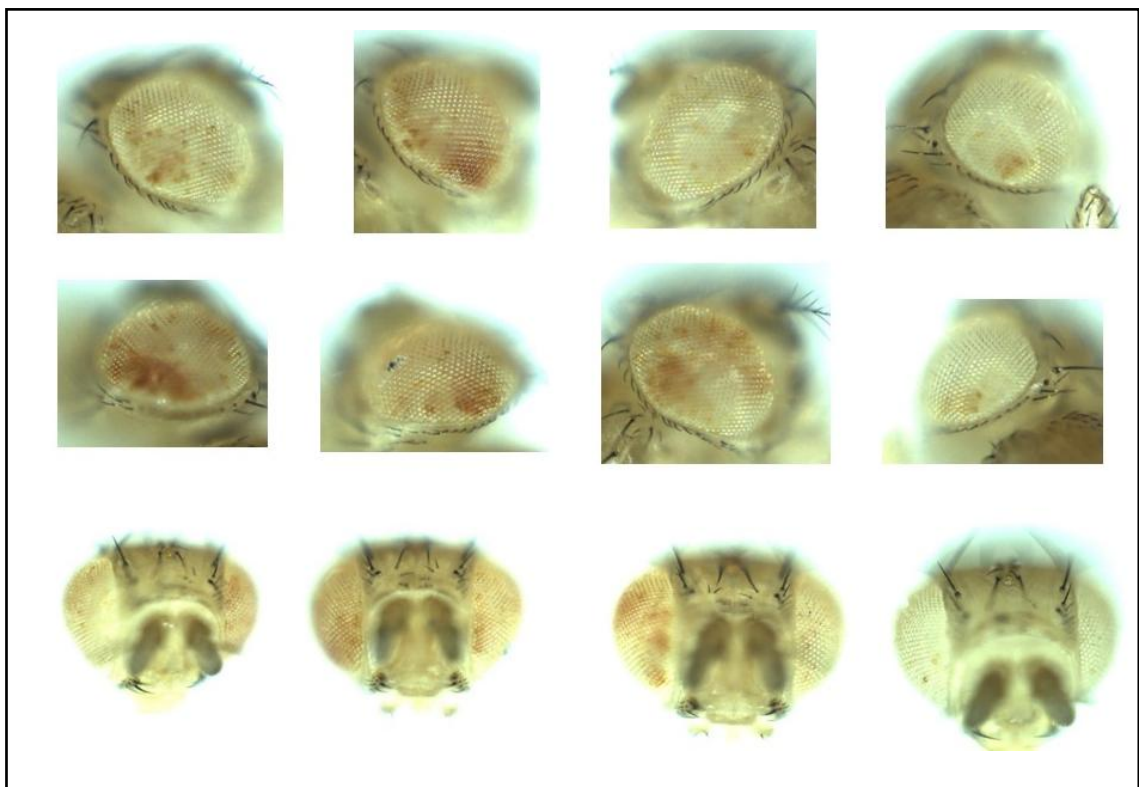


Figure 46. *Drosophila* Heads reared in 0M supplemented with 2% of SP

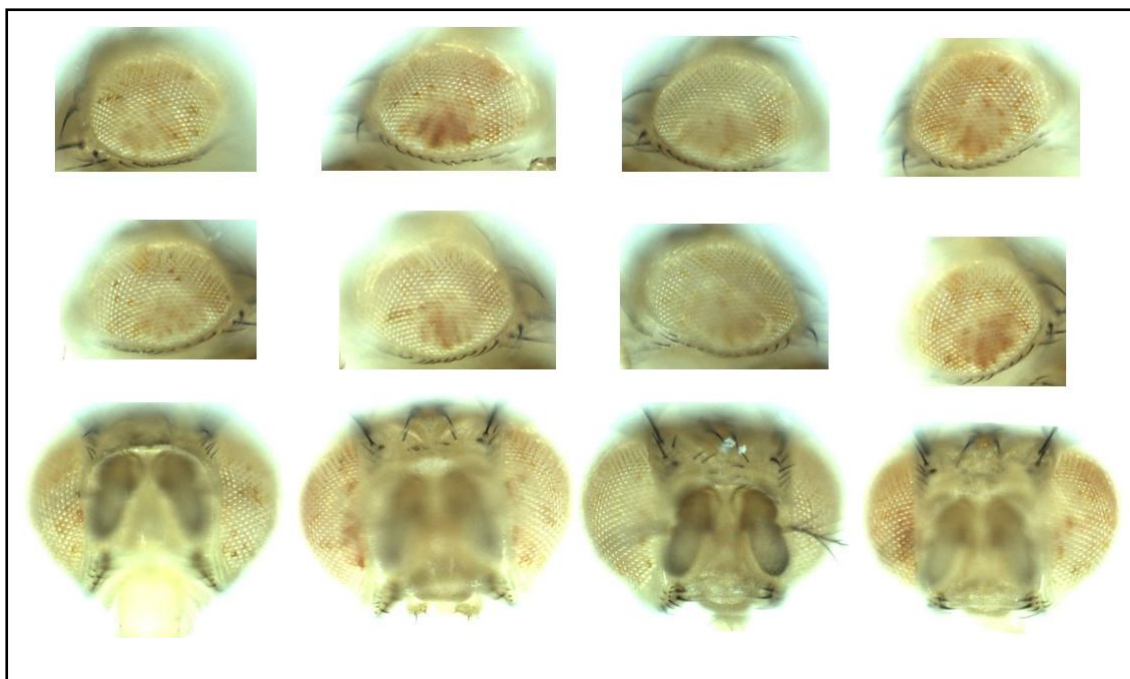


Figure 47. Drosophila Heads reared in 0M supplemented with 1% of SP

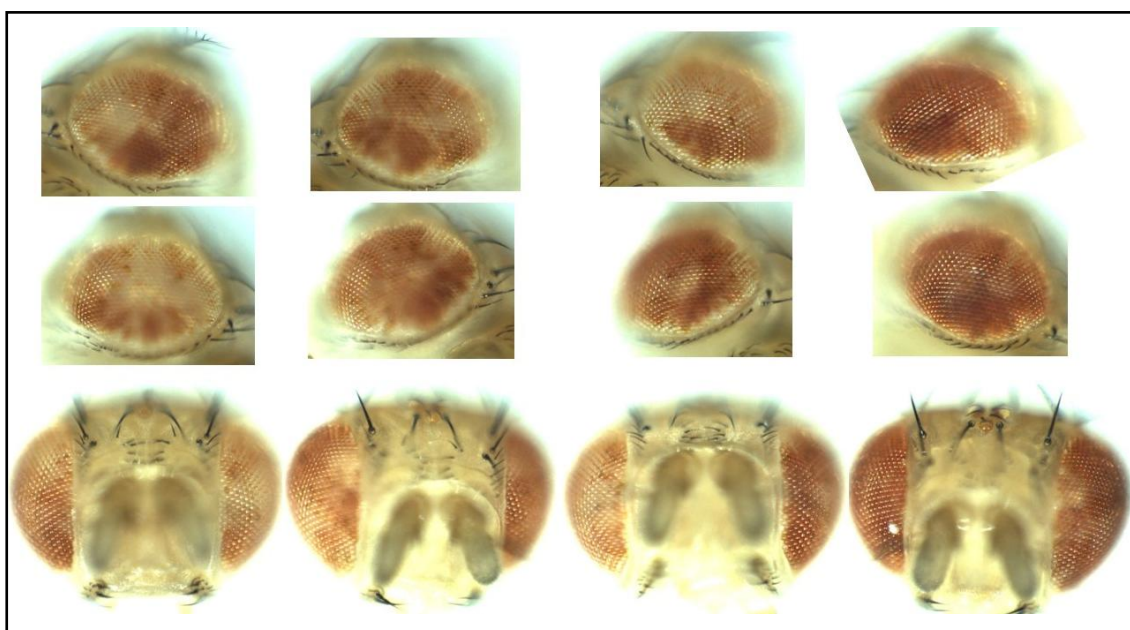


Figure 48. Drosophila Heads reared in NM

4.3.1.3. Effect of the protein-restricted diet on the brain size

Our research revealed a variation in the shape of the brain (Figure 49). 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 50). This disparity can be attributed to the amino acid profile of SP isolate, as amino acids influence the fat body and brain, regulating larval growth, body size, developmental

timing, and pupariation (Manière et al., 2020). In particular, the dietary amino acids in newly hatched *Drosophila* larvae triggered the quiescent cells to reenter the cell cycle (Britton & Edgar, 1998). The role of these amino acids in ensuring the cell cycle re-entry of neural progenitors occurred through the pathway of fat body, glia, and neuroblast relay (Sousa-Nunes et al., 2011). In addition, glial cells can directly detect amino acids by TORC1 via the Slif amino acid transporter and secrete DILP6 that promotes the neuroblast reactivation and Jelly belly (Jeb) secretion, allowing *dilp5* gene expression and DILPs release (Manière et al., 2020). In general, for insects, arginine, histidine, isoleucine, leucine, lysine, methionine, phenylalanine, threonine, tryptophan, and valine are the essential amino acids (EAA) (Behmer, 2008). However, the rest are considered non-essential and can be synthesised or derived from the EAA (Behmer, 2008). Yeast is the primary source of nutrients for *Drosophila*, providing amino acids and proteins (Steck et al., 2018). Since a yeast-based diet provided proper brain development, comparing its amino acid composition to the SP isolate reveals a richer profile of EAA than in the SP isolate, except for Phe, His, and Arg (Figure 51).

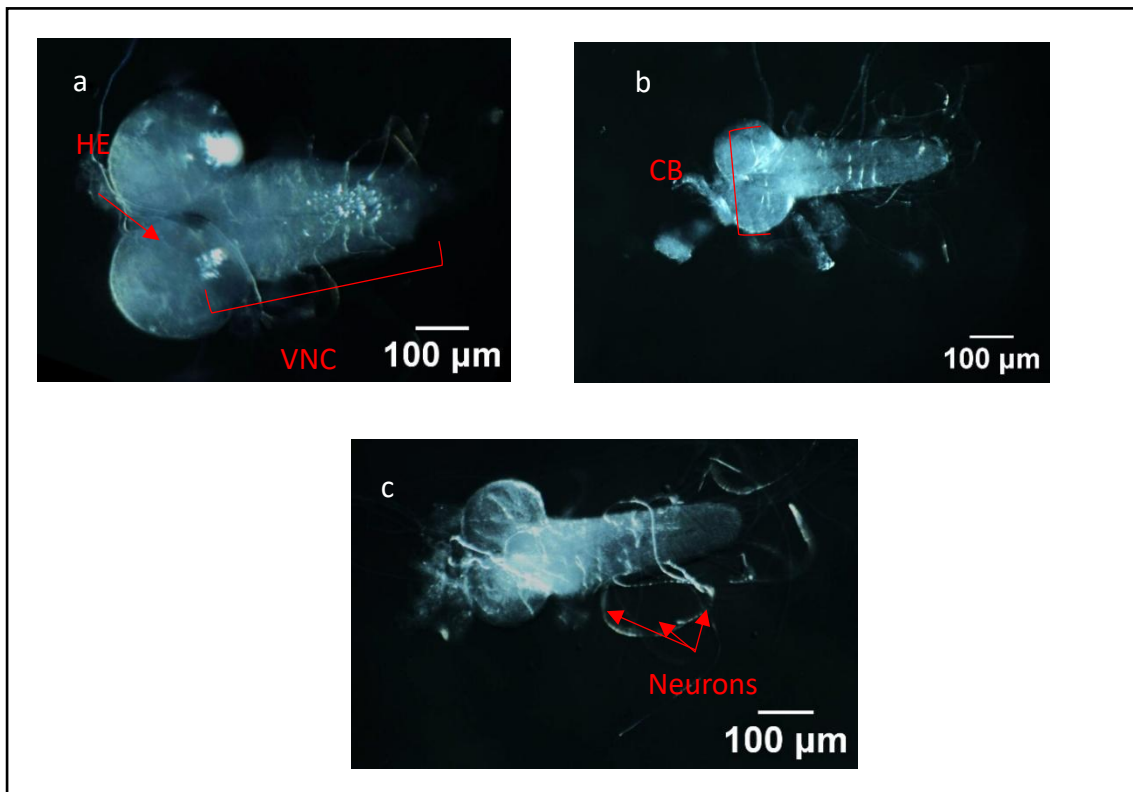


Figure 49. Brain of 3rd instar larvae cultivated in NM (control), 0M supplemented with 4% and 2% of SP. HE: Hemisphere, VNC: Ventral nerve cord, CB: Central brain.

Different superscript denotes the significant difference at $p < 0.05$

The importance of the amino acid was proven by the study of Britton & Edgar, (1998), which demonstrated that larvae grown in an amino acid-free medium did not grow. To assess whether a particular amino acid served as a signal for the cell cycle, each of the eleven essential amino acids was introduced individually. None of the amino acids alone restored the larval growth or cell cycle activation, suggesting that protein synthesis is necessary for cell cycle activation and growth (Britton & Edgar, 1998). This independence can be explained by the distinct behaviour of larvae (attraction, aversion, neutral) towards different amino acids (Croset et al., 2016). Furthermore, this preference for specific amino acids may differ depending on the growth stage, such as in first-instar and third-instar larvae (Kudow et al., 2017).

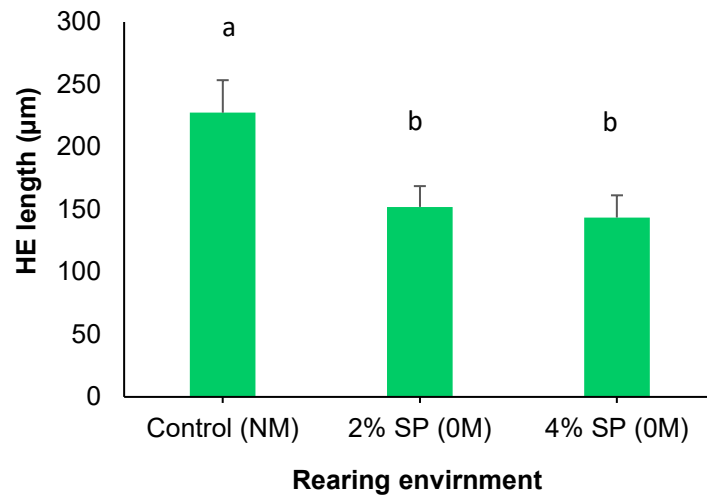


Figure 50. Length of the HE of the brain of 3rd instar larvae cultivated in NM(control), OM supplemented with 4% and 2% of SP. The error bars indicate the standard deviations of the samples. Different superscript denotes the significant difference at $p < 0.05$

Therefore, attributing the primary effect to a particular amino acid can not be supported; rather, the overall protein content of the SP incorporated is more critical. A previous research on *Drosophila* proved that extreme larval crowding (60 larvae cm^{-3}) impaired the volumetric development of the Antennal lobe, Optic lobe, and Central complex of adults (X. Wang et al., 2017). This effect can be related to the high larval density because of the excessive utilisation and interference competition for food and space.

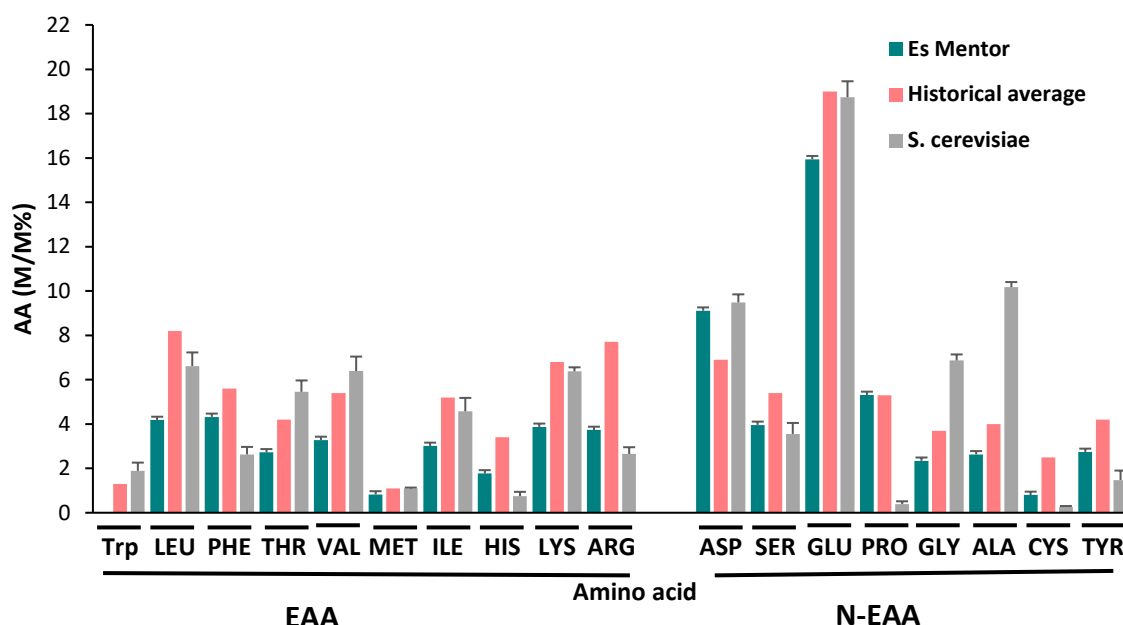


Figure 51. Amino acid composition of soybean (Es Mentor variety) extract (prepared by conventional alkaline-isoelectric precipitation) ($\pm 15\%$), the recorded historical average of amino acid composition in soybean protein (Nill, 2016), and the average of the amino acid composition of yeast extracts obtained from 3 strains of *S. cerevisiae* PFC107, PFC121, and PFC401 (Demirgul et al., 2022). EAA: Essential amino acid, N-EAA: non-essential amino acid. The error bars indicate the standard deviations of the sample

On the other hand, ecdysone can also regulate brain development by inducing the neuroblast division and their termination (Oliveira & Homem, 2023). The secretion of this hormone can be controlled by the PTTH secretion, which is downregulated by the amino acid transporter Sobremesa (Sbm)(Manière et al., 2020). The research of Galagovsky et al., (2018) indicated that the larvae with lowered Sbm levels in glia exhibit reduced brain growth alongside increased body size, but do not show a reduction in insulin signalling or production (Galagovsky et al., 2018). Therefore, we can conclude that a dysfunction in the secretion of the hormone or its cessation may induce a dysfunction in the development process.

4.3.2. Allergenic and bioactive peptide characterization

4.3.2.1. Size of the peptide

The comparison between the 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 characterize the bioactive peptides compared to the allergic ones (241 amino acids) (Figure 52).

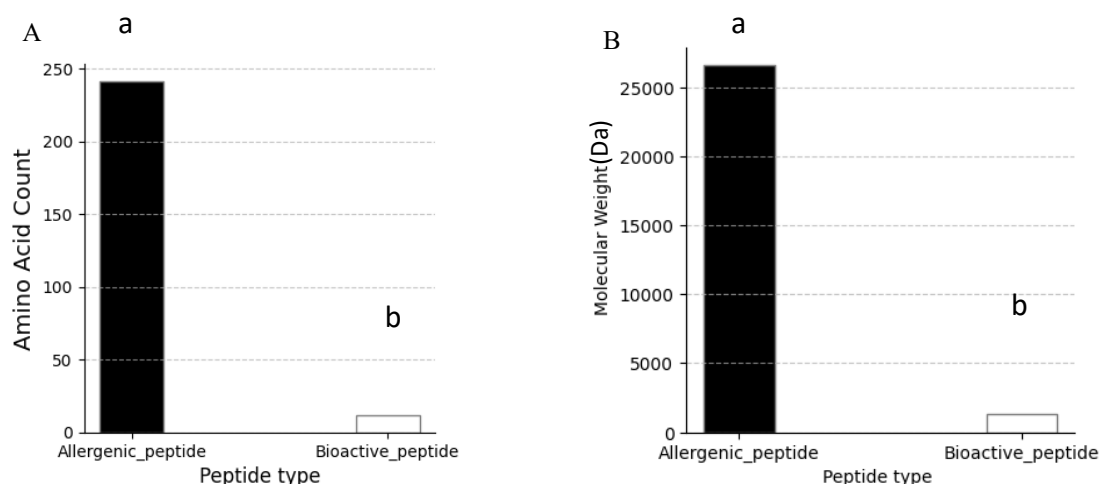


Figure 52. Amino acid count (A) and molecular weight (B) of the allergic and bioactive peptides. The mean of the amino acid count and molecular weight in each group of peptides. The statistical analysis was conducted using one-way ANOVA. Value with different superscripts differs significantly at $p < 0.05$.

Based on previous research, bioactive peptides are typically composed of 2-3 and 20 amino acid residues per molecule (Rivera-Jiménez et al., 2022, Chew et al., 2018, Naeem et al., 2022). However, longer bioactive peptides were rarely identified (Zaky et al., 2022). In particular, the functional property of a peptide may pertain to a specific segment of the peptide. In lunasin, which is a soybean-derived polypeptide that has marked anti-cancer activity, it is recognized that residues 1–22 have no known biological functions, whereas residues 23–31 make up a helical structure that is structurally similar to the protein that binds to histones H3/H4 (Wan et al., 2017, Guha & Majumder, 2019). The peptide sequence length can be related to its bioactive characteristic due to its reliance on the degree of hydrolysis during digestion. Polypeptides with a chain length varying from 3 to 51 amino acids can be absorbed by the gastrointestinal tracts of adult animals. Their intact form can enter the circulation and produce biological actions as long as they remain intact (Roberts et al., 1999). Several studies have connected the molecular weight of peptides

with their bioactive characteristics. There is a statistically significant difference between the molecular weight of the allergic and immunomodulatory peptides (Figure 52b), which can be associated with the ability of these peptides to be easily absorbed in the digestive tract to exert biological effects (Vo et al., 2013). Other studies confirmed the correlation between immunomodulating activity and molecular weight, indicating that lower molecular weight and positively charged peptides released from soy protein effectively stimulated immunomodulating activity (Kong et al., 2008). Another research related the low molecular weight (under 1 kDa) and the short chain length (2 to 20 residues) with the peptide bioactive activity. These characteristics increase the probability of being absorbed in the intestine and penetrating cells compared to the larger-sized peptides (Rivera-Jiménez et al., 2022). For example, the anti-inflammatory peptides from edible microalgae *Spirulina maxima* demonstrated that those with a low molecular weight of less than 3 kDa exerted the most inhibition on histamine release (Vo et al., 2013).

- **Composition of the peptides**

Peptides with allergic effect showed the dominance of Q (glutamine) with more than 11% occurrence (Figure 53). However, bioactive peptides showed a higher occurrence of E, P, G, K, S, and Q amino acids (>5%).

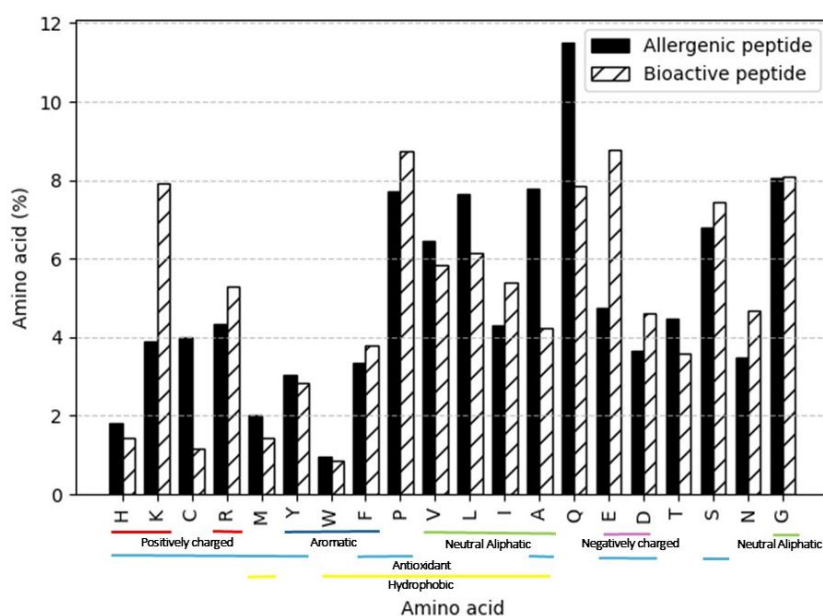


Figure 53. Prevalence of each amino acid in the allergic and immunomodulatory/anti-inflammatory peptides. The amino acid properties such as hydrophobicity, polarity, and side chain structure aliphatic, aromatic) are represented with different colours (Ren et al., 2013), (Aftabuddin & Kundu, 2007; L. Wang et al., 2016)

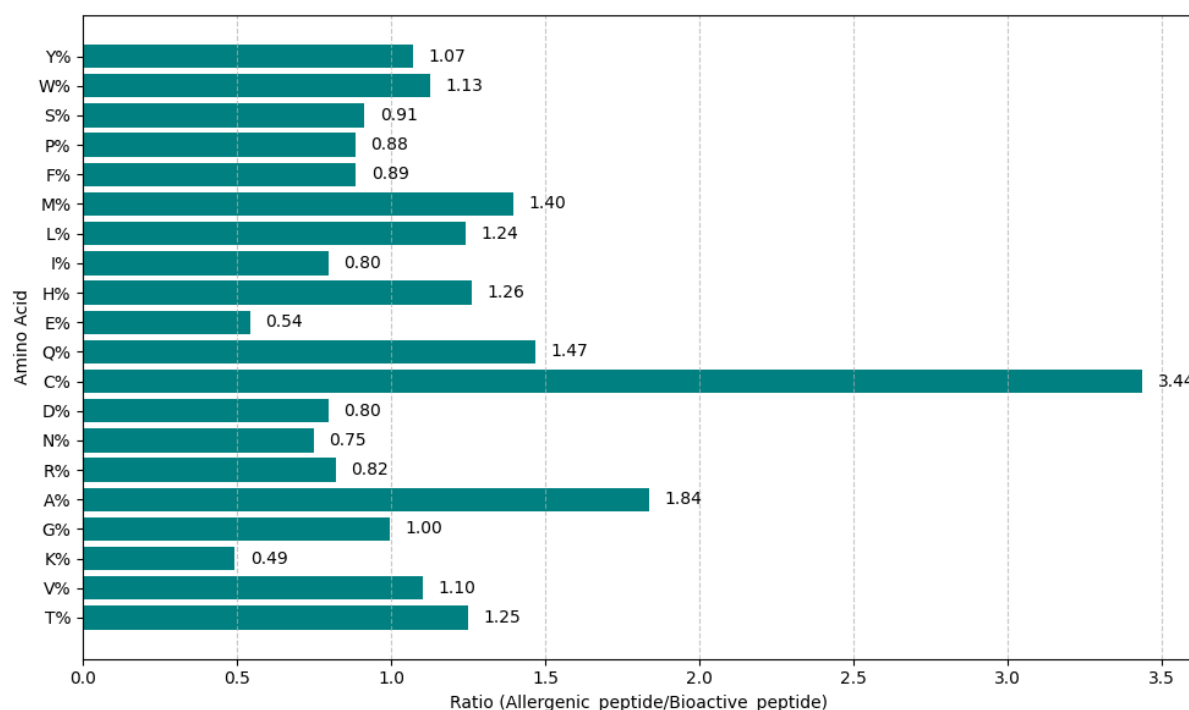


Figure 53. Average ratio of each amino acid in the allergic peptide/amino acid in the immunomodulatory peptide

- **Thiol amino acid (Cysteine)**

Sulphur-containing amino acids, such as methionine (M) and cysteine (C), proved to have a critical role in health maintenance (Masella & Mazza, 2009). Although they are characterized 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 (Figures 54 and 55). The occurrence of cysteine was 3.4 times higher than in the bioactive peptides.

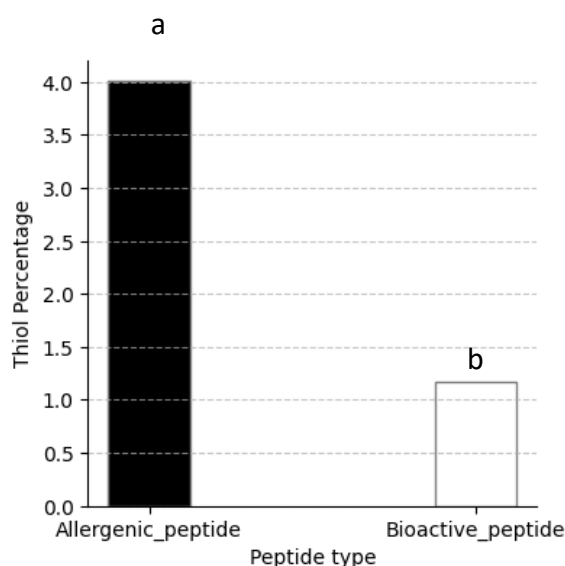


Figure 5. Thiol in allergic and bioactive peptides

The values represent the mean value. The statistical analysis was conducted using one-way ANOVA. Value with different superscripts differs significantly at $p < 0.05$.

- **Amine amino acid**

In contrast with the thiol amino acids, the lysine (K) ratio was twice as high as that in the bioactive peptides as in the allergic ones (Figure 56). According to the literature, lysine-rich content is a characteristic feature of many bioactive peptides (Tu et al., 2009). Lysine has been identified in the majority of anti-inflammatory peptides from various plant sources. Along with arginine and histidine, these amino acids are essential to boost the anti-inflammatory activity of plant-derived bioactive peptides (W. Liu et al., 2022). The research of (J. Cheng et al., 2020) proved that L-lysine effectively promoted M2 microglial polarization and reduced the inflammatory response both *in vitro* and *in vivo*, indicating its potential neuroprotective role in intracerebral hemorrhage injuries.

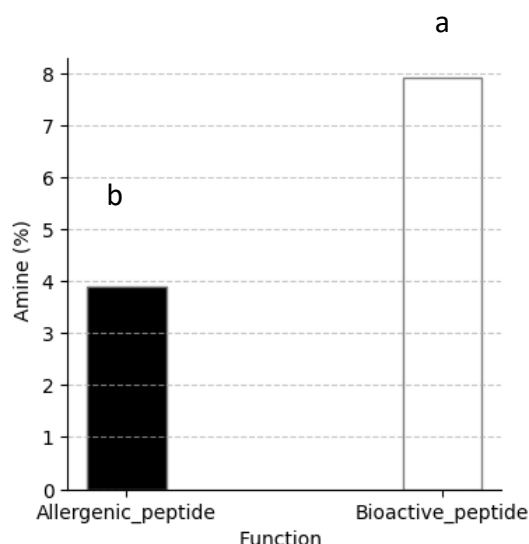


Figure 55. Frequency of prevalence of the amine amino acid (K)

The values are the mean. The statistical analysis was conducted using one-way ANOVA. Value with different superscripts differs significantly at $p < 0.05$.

The potential of the lysine/leucine-rich peptide L-K6 demonstrated significant cytotoxicity against various cancer cell lines. Additionally, their cytotoxicity was reduced after the substitution of the leucine residues (L-K6V1) and deletion of the lysine residues (L-K5V1) (C. Wang et al., 2017). An experiment on piglets demonstrated this effect, showing that lysine restriction in their diet influenced the inflammatory responses in the kidney, liver, and spleen via mediating serum antibody volume and inflammatory cytokine (H. Han et al., 2018). Another study indicated that modifying lysine, arginine, and cysteine residues in antigen epitopes in processed shrimp might be responsible for the hypo-IgE-binding activity (M. Liu et al., 2021).

- **Amino acid E (Glutamic acid)**

The results showed that the proportion of glutamic acid (E) in the bioactive peptides is 1.8 times that in the allergic peptides. This amino acid is recognized for its therapeutic uses, including as an endogenous anticancer agent, in conjugates with anticancer agents, and in its derivatives, which may also serve as potential anticancer agents (Dutta et al., 2013). Recent research demonstrated the potential of using glutamic acid as a ligand to react with low-molecular-weight Hyaluronic acid to mitigate inflammation-triggered skin symptoms and aging. The findings indicated that it effectively supported the repair of damaged skin by enhancing the compactness of the stratum corneum and increasing the live cell layer (F. Jiang et al., 2024).

- **Amino acid residues**

The structure and properties of proteins rely on the charge distribution of their side chains, which are essential to their electronic structure, physicochemical properties, and biological functionality (Soriano-Correa et al., 2015). Among the amino acids in the residue on the C and N-termini of the bioactive peptides, the results indicated that K, R, and G are the most prevalent amino acids in the residue on the C and N-termini of the bioactive peptides (Figure 57).

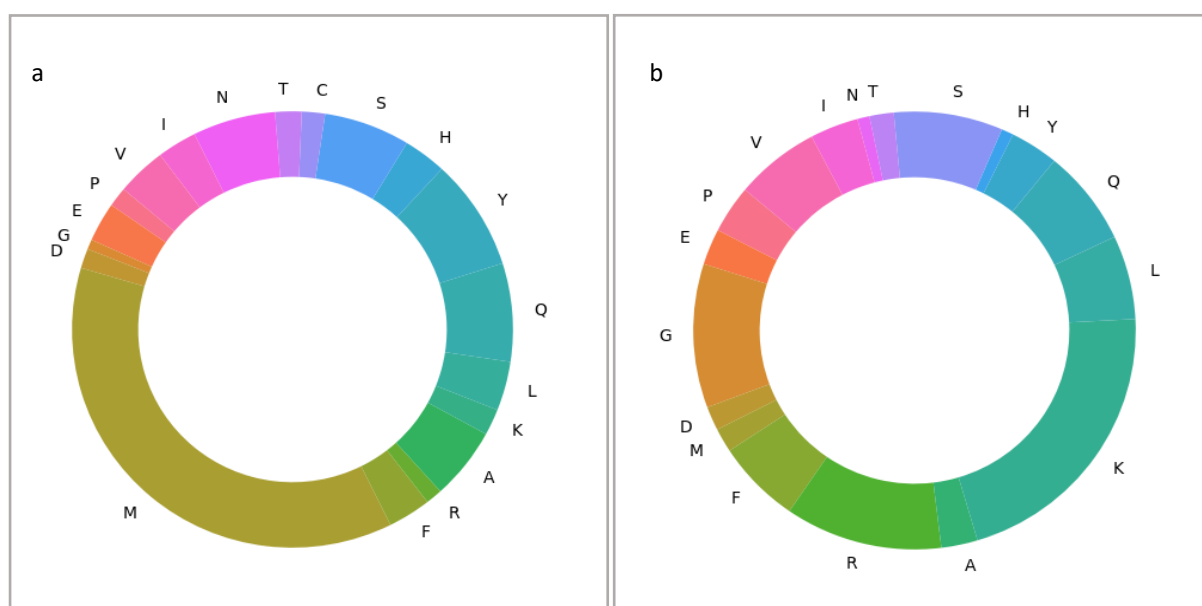


Figure 57. Frequency of amino acid in the termini of the sequence of (A) Allergenic peptide and (B) bioactive peptide

Colours representing the ratio of each amino acid

a: Allergenic peptide, b: Bioactive peptide

Research literature confirms arginine's link to its immunomodulatory effects, indicating that peptides containing arginine in the N- or C-terminal regions can be considered potential immunomodulatory peptides (Q. Yang et al., 2020). In addition, the immunomodulatory effect can be attributed to the charge of these amino acids. For example, positive correlations were obtained between the immunomodulating activity and the content of positively charged peptides (Kong et al., 2008). In particular, positively charged amino acids, such as lysine, arginine, and histidine, demonstrated their necessity to enhance the anti-inflammatory activity of plant-derived bioactive peptides (W. Liu et al., 2022).

- **Aliphatic_Index**

The results show that there is no statistically significant difference between the aliphatic index of the immunomodulatory and allergic peptides (Figure 58). As demonstrated in previous research (Rehman et al., 2020), the obtained results (Aliphatic_Index >65) show that both types of peptide are stable at high temperatures. The same effect was observed with allergenic proteins derived from different milk sources (cow, goat, camel, and mare milk), suggesting their good thermal stability (Zhao et al., 2023).

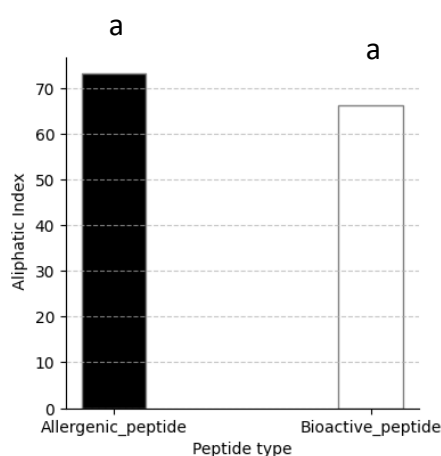


Figure 58. Average aliphatic index in allergic and bioactive peptides.

The values are the mean. The statistical analysis was conducted using one-way ANOVA. Value with different superscripts differs significantly at $p < 0.05$.

- **Peptide Hydrophobicity**

Our results demonstrated a statistically significant difference between the hydrophobicity of the allergic and immunomodulatory peptides (Figure 59).

Peptide hydrophobicity contributes to the determination of the protein function. The immunomodulatory functions of Human Cathelicidin LL-37-derived peptides with an identical positive charge and amino acid composition are directly correlated with the predicted α -helical propensity and hydrophobicity of the peptide (Hemshkhar et al., 2019).

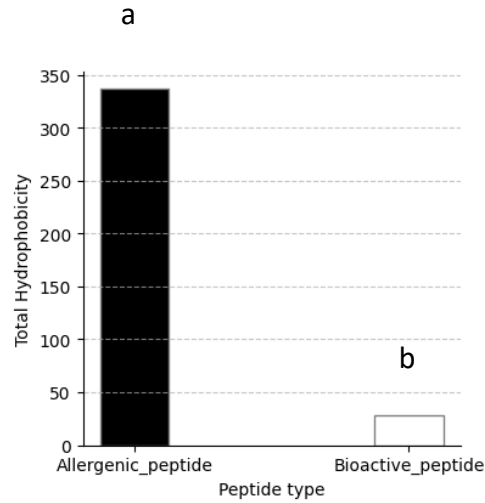


Figure 59. Average total hydrophobicity in allergic and bioactive peptides
The values are the mean. The statistical analysis was conducted using one-way ANOVA. Value with different superscripts differs significantly at $p < 0.05$.

Previous research proved that the presence of hydrophobic (V, I, P) and positively charged (H, R, K) amino acids seemed to represent a common characteristic of the anti-inflammatory peptides (Rivera-Jiménez et al., 2022).

- **Aromatic amino acids**

The results revealed no significant difference between the prevalence of the aromatic amino acids in the allergic and the bioactive peptides (Figure 60).

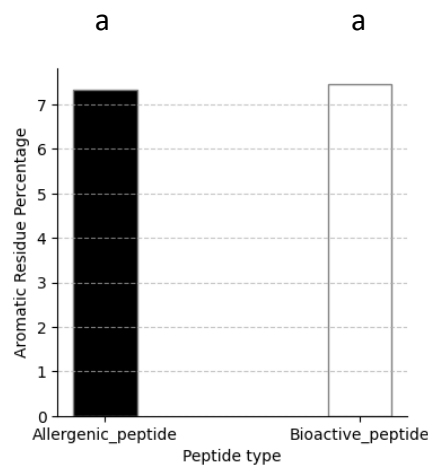


Figure 60. Average total aromatic amino acids in allergic and bioactive peptides
The values are the mean. The statistical analysis was conducted using one-way ANOVA. Value with different superscripts differs significantly at $p < 0.05$.

Precious results emphasized the significance of the position of the amino acid within the sequence. Rakesh and his collaborators (2019) showed that the aromatic amino acids have less anti-inflammatory activity than D and T, when substituted in the second position (Rakesh et al., 2019). In contrast, Feng et al, 2024 demonstrated that the aromatic residues at the penultimate position are preferable to exert potent anti-inflammatory activity (J. Feng et al., 2024).

4.3.2.2. Relationship between peptide characteristics and functional properties

The heatmap illustrates the correlation between all the studied 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)(Figure 61).

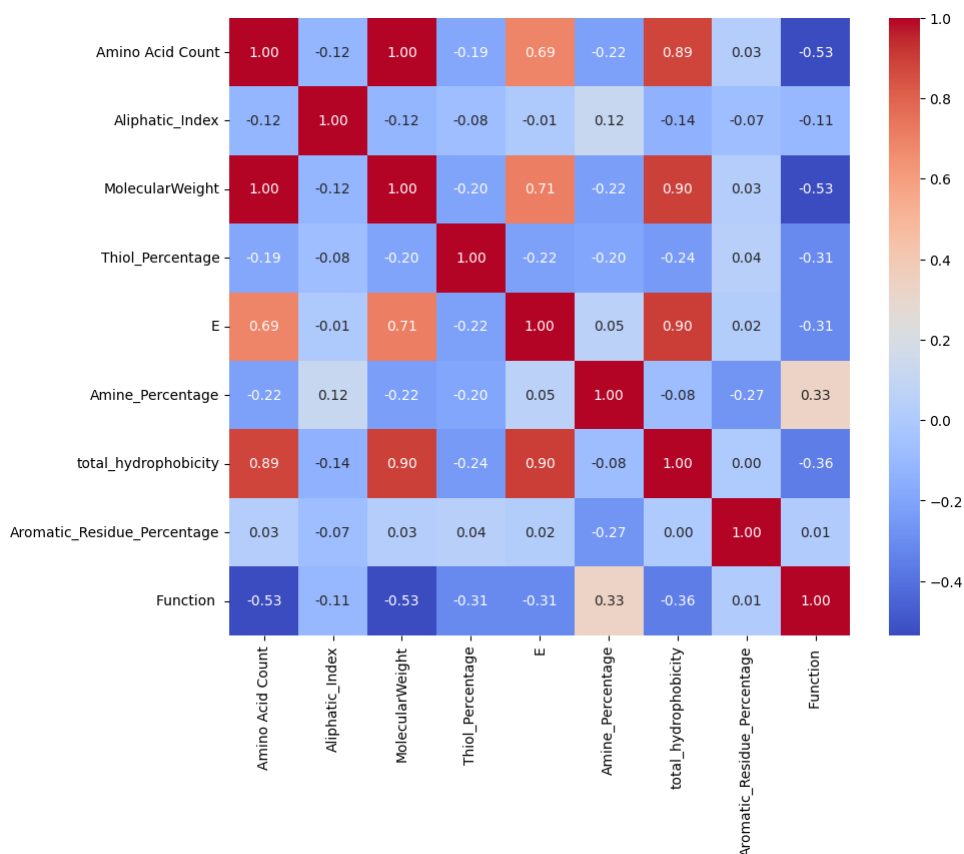


Figure 61. Correlation of all the descriptors and the nature of the peptide (allergenic/bioactive) referred 'Function'.

The colour of cells is associated with the strength of the correlation between feature-to-feature and feature-to-response. The strength of the correlation is indicated in the colour scale: red (positive correlations), blue (negative correlations).

The correlation map reveals 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 the amine amino acid (K) and the response. Conversely, the prevalence of thiol, E, and total hydrophobicity demonstrated a weak negative relationship. On the other hand, the aromatic amino acids show almost no correlation with the type of the peptide, indicating their independence.

4.3.2.3. Epitopes detection from bioactive sequences

The structure of the protein significantly influences its allergenicity. Modifications in the structure by amino acid mutation or other factors can alter the IgE binding capacity of allergens, thereby reducing their potential allergenicity of proteins (X. Zhou et al., 2023).

Defined as antigenic determinants are regions of proteins that can trigger the allergic reaction (Kogay & Schönbach, 2019), epitopes that induce allergy should be identified (C. Liu & Sathe, 2018). Concerning bioactive peptides, it is also essential to identify, characterize, and check the occurrence of the epitope in the sequence. In particular, Tran and his collaborators (2022) demonstrated that extending short peptides into longer multi-epitope peptides overcomes immune tolerance, inducing an immune response compared to a single short peptide (Tran et al., 2022). Among the bioactive peptides, the possible epitope responsible for anti-inflammatory reaction can be identified by detecting the pattern in the amino acid sequence of the bioactive peptide (Annex 6).

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) (Figure 62). However, the relative prevalence of W in the epitope sequence was positively correlated with the epitope's allergenicity loss. The richness of the epitope with the specific amino acids W and Y was demonstrated by a previous study, indicating the general preference for charged and polar amino acids (Rubinstein et al., 2008). Similarly, in another research W residue was recognized as crucial for immunoreactivity of a diagnostically relevant peptide epitope of *Aspergillus fumigatus* (Kamal et al., 2005), and as a critical amino acid, while the identification of the nephritogenic T cell epitopes, (W136I137xL139W140) (S.-Y. Hu et al., 2017).

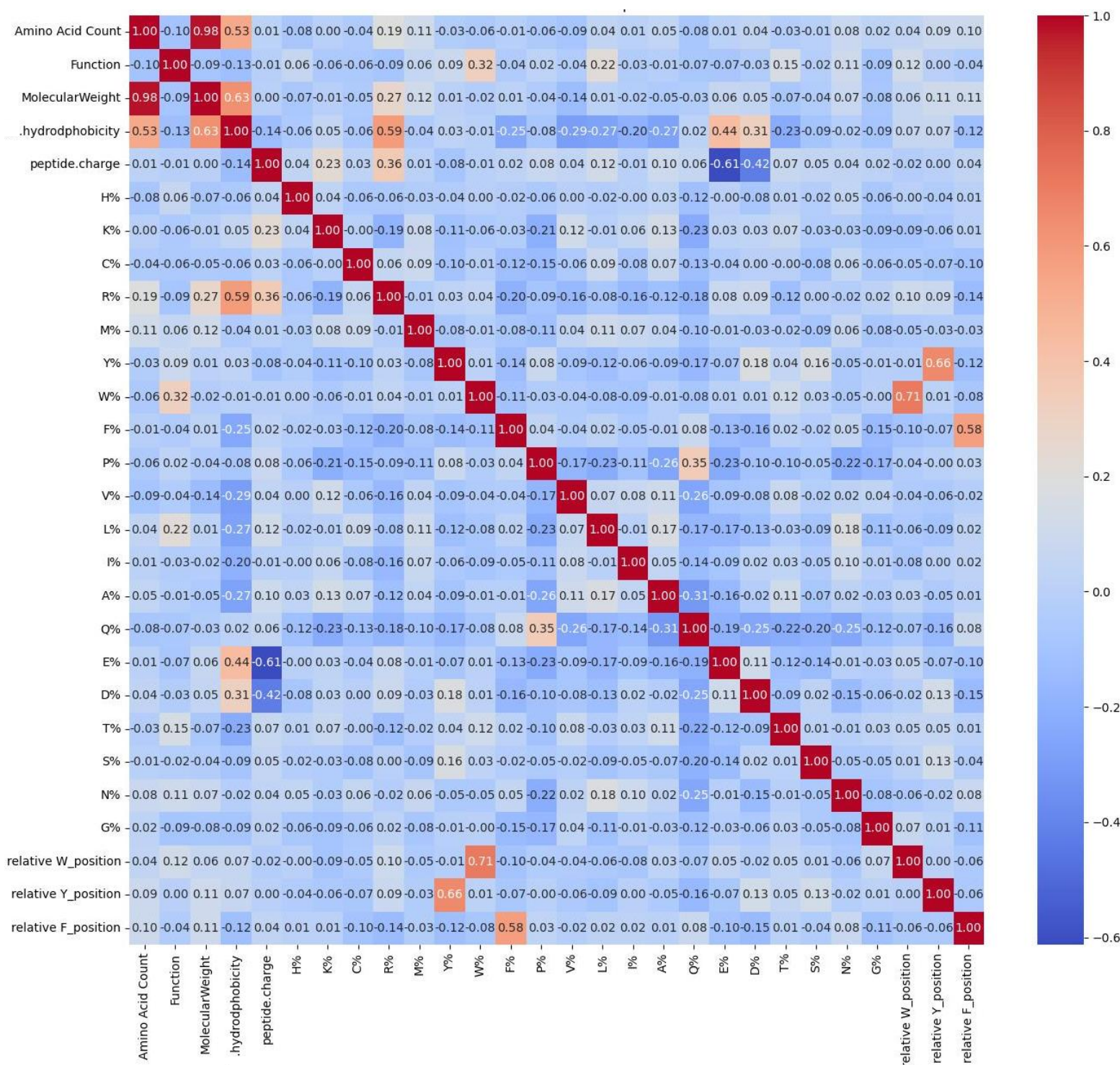


Figure 62. Correlation of all the descriptors and the epitope/immunosuppressive sequence referred to as Function. The colour of cells is associated with the strength of the correlation between feature-to-feature and feature-to-response. The strength of the correlation is illustrated in the colour scale: red (positive correlations), blue (negative correlations).

The results of the PCA, as illustrated in Figure 62, show that PC1 and PC2 explain about 86% of the variance. In particular, the feature representing the prevalence of W in the epitope's sequence contributed more to PC2, which accounted for 25% of the cumulative variance. However, the epitope's hydrophobicity, molecular weight, and amino acid count showed a higher contribution to PC1, with 61% (Figures 63 and 64).

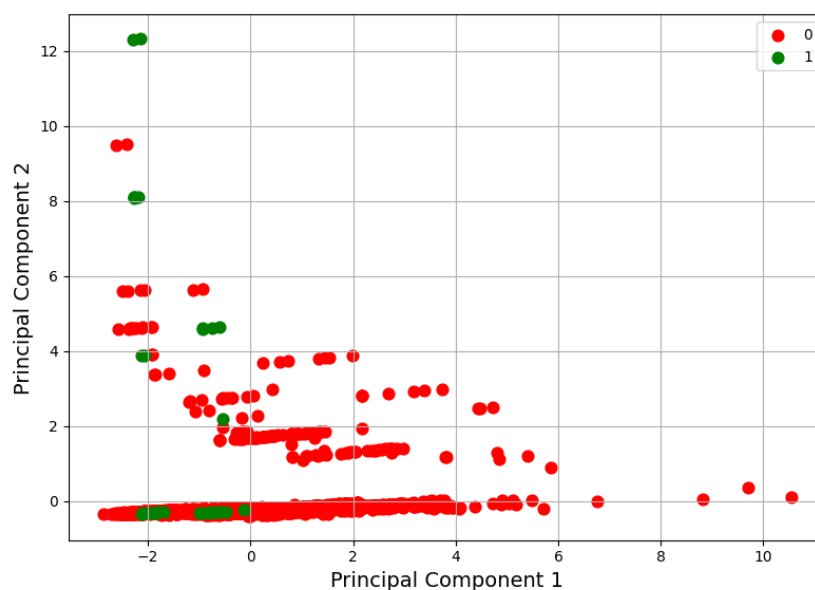


Figure 63. Principal component analysis (PCA) projection of the epitope's physicochemical properties and the prevalence of W. The PCA1 and PCA2 were plotted in a two-dimensional scatter plot with the colour of the label Function that defines the function of the epitope (allergen (red) or immunosuppressive (green)).

The two different groups of epitopes overlap, which is an indicator of a similarity between their physicochemical properties. This outcome can be interpreted in light of research efforts toward other explanations. The study of a particular amino acid identified as a critical amino acid in the protein sequence is one of the explanations.

They are identified as amino acid residues that constitute the potential allergenic peptides and have a critical role in their allergenicity (X. Wang et al., 2024). They have the ability to affect the trans-epithelial transport absorption efficiency and, consequently, the allergenic potential (Luan et al., 2024).

The critical amino acids are defined as those that cause a loss of IgE binding when substituted with an alanine residue or an amino acid with different physicochemical

properties, which can influence the potential sensitisation (Xiaoya et al., 2024). Reese et al. (2006) showed that by using mapping sequential epitopes, 12 positions deemed critical amino acids were identified and mutated, resulting in a Pen a 1 mutant with significantly reduced allergenic potency (90–98%) (Reese et al., 2005).

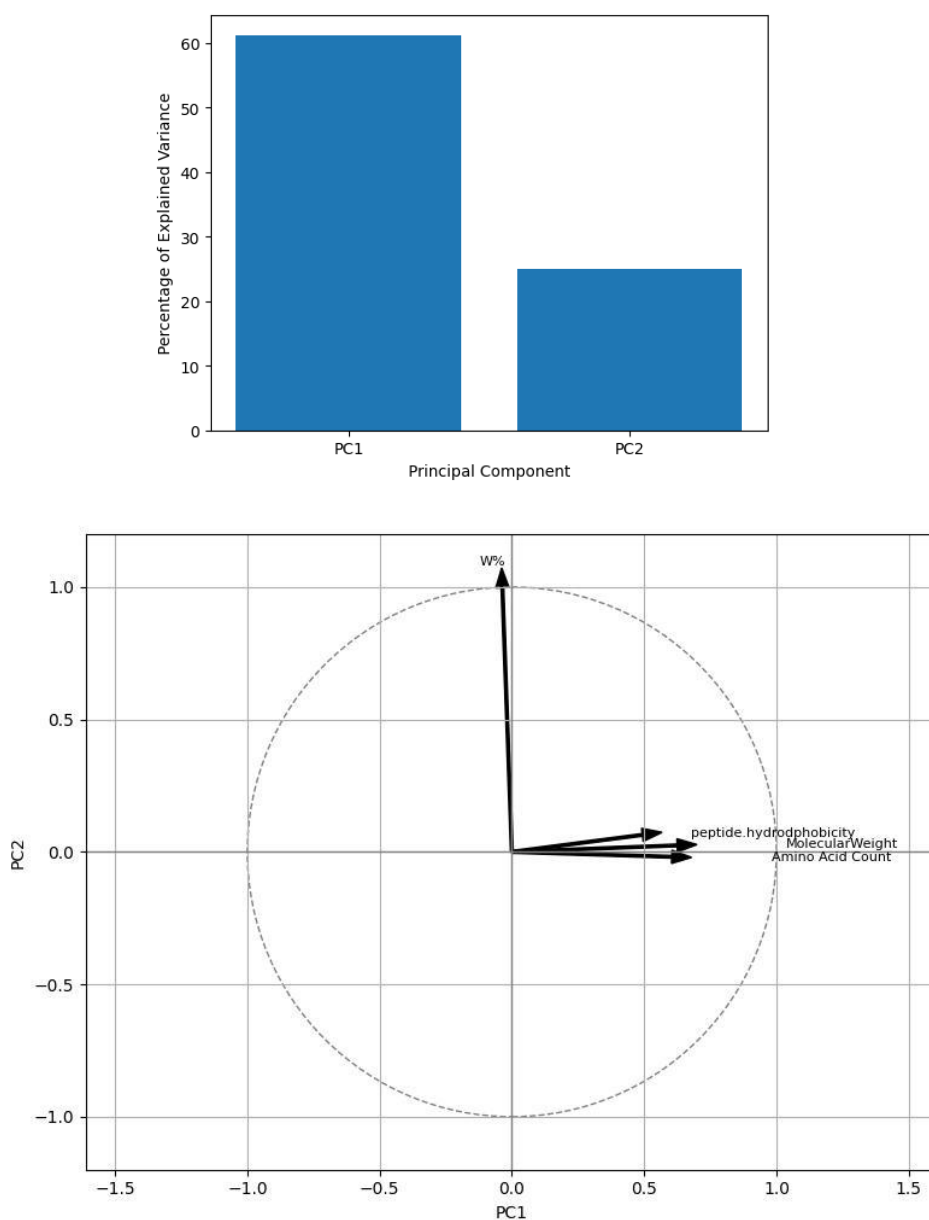


Figure 64. Diagram of the first two principal components (PCA)

By detecting the pattern in the amino acid sequences of the bioactive peptides, a group of potential epitopes was 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.

This test measures the cosine of the angle formed between two vectors, determining if the two vectors are pointing in roughly the same direction (J. Han et al., 2012).

Using only W as a parameter, the similarity (Figure 65) reveals 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 of 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.

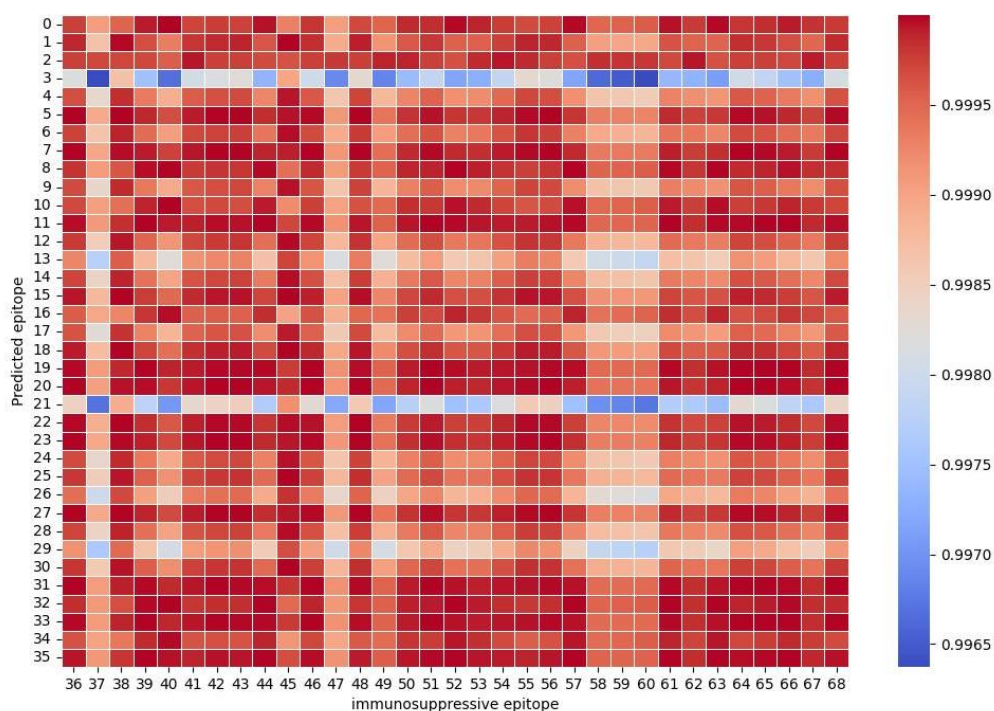


Figure 65. Heatmap of cosine similarities between the predicted epitope and the immunosuppressive epitopes. The difference in the intensity of the color indicates the degree of similarity. Each number represents the sequence of the epitope of each group.

5. CONCLUSIONS, RECOMMENDATIONS

To address the rising demand for plant protein to cope with the growing world population, safe and eco-friendly methods should be adopted. In this context, advancements in protein fractionation techniques should prioritize both process efficiency and environmental sustainability. This study accomplishes these goals by employing mild conditions (full-fat soybean flour, moderate temperatures) while extracting protein using a conventional method and a binary mixture of DES-water as a green technology. In the same context, with the growing availability of data on public databases, whether generated by machines in industrial settings or produced in laboratories, using models to predict production yield or protein purity under different conditions can save time and reduce costs. This thesis has examined the efficiency of data use to predict the extraction efficiency and to improve the understanding of the aspects that characterize the peptide functionality.

A special focus was also addressed to the nutritional aspect of the fractionation. Using the *Drosophila* model, the study examined the potential effect of an exclusive soybean protein diet on gene silencing and the toxicity of single amino acid supplements. While these findings offer a general overview of the nutritional risk, it is essential to confirm the persistence of this pattern during fly maturation and across future generations. Furthermore, investigating the genetic foundation of eye color differences can provide a more profound understanding of evolutionary adaptations related to diet.

While legume proteins are associated with allergenicity, fractionation can enable the generation of bioactive peptides exhibiting biofunctional properties. This study highlights the significance of the sequence length and molecular weight in the characterization of both peptides. Additionally, since peptides are known to stimulate the adaptive immune response via their immunogenic epitopes, identifying motif patterns within them offers valuable guidance for epitope detection and product design.

Protein fractionation may provide benefits relevant to both nutrition and industry. However, one of the primary limitations lies in the risk of protein denaturation and the generation of processed products. These products can undergo further technological modifications to yield ultra-processed food whose intake has been strongly linked to a high risk of diabetes, hypertension, dyslipidemia, and obesity (Vitale et al., 2024).

6. NEW SCIENTIFIC RESULTS

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%).

7. PRACTICAL 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.

4. 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.

8. SUMMARY

To meet the growing demand for protein while safeguarding the environment, extensive research has been devoted to various aspects of protein extraction. While efforts have focused on the origin of protein (bio-products and waste, plant, insects), other investigations have largely centered on optimizing and designing environmentally sustainable methods. In this study, a different approach was addressed by optimizing extraction methods by leveraging existing data. By adopting the conventional method (alkaline isoelectric precipitation), various algorithms were employed to model the purity of the protein isolate. The selected model was well aligned with the extracted data and was validated by the experiment held using a mild extraction technique and a full factorial design. This study underscores the potential of integrating data collection from prior experiments to optimize the extraction process.

On the other hand, with the aim of optimizing extraction efficiency through the use of DES, this study demonstrated the inefficiency of the addition of water to choline chloride: glycerol in the improvement of the extraction.

To evaluate the nutritional impact of the protein isolate, the computational method supported the understanding of the relationship between the amino acid sequences and the peptide's functionality, including allergenicity and bioactivity. Analysing a database of 263 sequences of plant (cereal and legume) proteins with allergic and bioactive properties revealed the importance of size and molecular weight. While mapping the sequences of bioactive peptides, the identical parts of the sequence were identified and compared to the immunosuppressive epitope. However, these findings are preliminary, and validation of the functional properties of the identified parts is paramount.

Furthermore, the assessment of the nutritional effect of the emerging dietary trends, such as high protein regimens and targeted single amino acid supplementation, was carried out using soybean protein, lysine, and the *Drosophila* model. The development, morphological features such as organ size, and phenotypic aspects like eye color were evaluated. A diet restricted to soybean protein resulted in a smaller larval brain and overall pupal size compared to the standard conditions. This also raised the potential of epigenetic mutation, which manifested in the intensification of the white color of the eyes.

On the other hand, while lysine in low concentrations appeared necessary for *Drosophila* development but insufficient from a qualitative perspective to complete the development

of larvae into pupae, higher concentrations demonstrated a toxic effect. This research highlights the need to consider nutrient quality and quantity when designing specific dietary strategies.

9. 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>
- Abu Risha, M., Rick, E.-M., Plum, M., & Jappe, U. (2024). Legume Allergens Pea, Chickpea, Lentil, Lupine and Beyond. *Current Allergy and Asthma Reports*, 24(9), 527–548. <https://doi.org/10.1007/s11882-024-01165-7>
- Aftabuddin, M., & Kundu, S. (2007). Hydrophobic, Hydrophilic, and Charged Amino Acid Networks within Protein. *Biophysical Journal*, 93(1), 225–231. <https://doi.org/10.1529/biophysj.106.098004>
- Ahmad, M., Chaudhary, S. U., Afzal, A. J., & Tariq, M. (2015). Starvation-Induced Dietary Behaviour in *Drosophila melanogaster* Larvae and Adults. *Scientific Reports*, 5(1), 14285. <https://doi.org/10.1038/srep14285>
- Ahmadifard, N., Murueta, J. H. C., Abedian-Kenari, A., Motamedzadegan, A., & Jamali, H. (2016). Comparison the effect of three commercial enzymes for enzymatic hydrolysis of two substrates (rice bran protein concentrate and soy-been protein) with SDS-PAGE. *Journal of Food Science and Technology*, 53(2), 1279–1284. <https://doi.org/10.1007/s13197-015-2087-6>
- Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., & Walter, P. (2002). *Drosophila and the Molecular Genetics of Pattern Formation: Genesis of the Body Plan*. In *Molecular Biology of the Cell*. 4th edition. Garland Science. <https://www.ncbi.nlm.nih.gov/books/NBK26906/>
- Aleya, A., Mihok, E., Pecsénye, B., Jolji, M., Kertész, A., Bársony, P., Vigh, 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>
- Ali, M. C., Yang, Q., Fine, A. A., Jin, W., Zhang, Z., Xing, H., & Ren, Q. (2015). Efficient removal of both basic and non-basic nitrogen compounds from fuels by deep eutectic solvents. *Green Chemistry*, 18(1), 157–164. <https://doi.org/10.1039/C5GC01823D>
- Ampofo, J., & Ngadi, M. (2022). Ultrasound-assisted processing: Science, technology and challenges for the plant-based protein industry. *Ultrasonics Sonochemistry*, 84, 105955. <https://doi.org/10.1016/j.ultsonch.2022.105955>
- Antony, J. (2023). *Design of Experiments for Engineers and Scientists* (3rd ed.). Elsevier. [https://ebookcentral.proquest.com/lib/\[SITE_ID\]/detail.action?docID=7255396](https://ebookcentral.proquest.com/lib/[SITE_ID]/detail.action?docID=7255396)
- Anyiam, P. N., Phongthai, S., Kingwascharapong, P., Pongsetkul, J., Zhang, W., Jung, Y. H., & Rawdkuen, S. (2025). A comparative review on novel-assisted extraction techniques for retrieving protein from some potential plant sources. *NFS Journal*, 40, 100239. <https://doi.org/10.1016/j.nfs.2025.100239>
- Arantes, J. da S., Arantes, M. da S., Fröhlich, H. B., Siret, L., & Bonnard, R. (2021). A novel unsupervised method for anomaly detection in time series based on statistical features for industrial predictive maintenance. *International Journal of Data Science and Analytics*, 12(4), 383–404. <https://doi.org/10.1007/s41060-021-00283-z>

- Arslan, E., Schulz, J., & Rai, K. (2021). Machine Learning in Epigenomics: Insights into Cancer Biology and Medicine. *Biochimica et Biophysica Acta. Reviews on Cancer*, 1876(2), 188588. <https://doi.org/10.1016/j.bbcan.2021.188588>
- Aydar, A. Y. (2018). Utilization of Response Surface Methodology in Optimization of Extraction of Plant Materials. In *Statistical Approaches With Emphasis on Design of Experiments Applied to Chemical Processes*. IntechOpen. <https://doi.org/10.5772/intechopen.73690>
- 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>
- Behmer, S. T. (2008). Nutrition in Insects. In J. L. Capinera (Ed.), *Encyclopedia of Entomology* (pp. 2646–2654). Springer Netherlands. https://doi.org/10.1007/978-1-4020-6359-6_2277
- Belwal, T., Devkota, H. P., Ramola, S., Andola, H. C., & Bhatt, I. D. (2020). Chapter Ten—Optimization of extraction methodologies and purification technologies to recover phytonutrients from food. In S. M. Nabavi, I. Suntar, D. Barreca, & H. Khan (Eds.), *Phytonutrients in Food* (pp. 217–235). Woodhead Publishing. <https://doi.org/10.1016/B978-0-12-815354-3.00007-1>
- Bezerra Da Silva, C. S., Park, K. R., Blood, R. A., & Walton, V. M. (2019). Intraspecific Competition Affects the Pupation Behavior of Spotted-Wing *Drosophila* (*Drosophila suzukii*). *Scientific Reports*, 9(1), 7775. <https://doi.org/10.1038/s41598-019-44248-6>
- Boateng, I. D. (2023). Evaluating the status quo of deep eutectic solvent in food chemistry. Potentials and limitations. *Food Chemistry*, 406, 135079. <https://doi.org/10.1016/j.foodchem.2022.135079>
- Boatta, F., van Hal, J., Beukeboom, L. W., & Ellers, J. (2024). High-fat and high-sugar diets induce rapid adaptations of fat storage in the house fly *Musca domestica* L. *Journal of Evolutionary Biology*, 37(11), 1368–1377. <https://doi.org/10.1093/jeb/voae122>
- Bortolotti, A., How Wong, Y., S. Korsholm, S., B. Bahring, N. H., Bobone, S., Tayyab, S., Weert, M. van de, & Stella, L. (2016). On the purported “backbone fluorescence” in protein three-dimensional fluorescence spectra. *RSC Advances*, 6(114), 112870–112876. <https://doi.org/10.1039/C6RA23426G>
- Boye, J., Zare, F., & Pletch, A. (2010). Pulse proteins: Processing, characterization, functional properties and applications in food and feed. *Food Research International*, 43(2), 414–431. <https://doi.org/10.1016/j.foodres.2009.09.003>
- Britton, J. S., & Edgar, B. A. (1998). Environmental control of the cell cycle in *Drosophila*: Nutrition activates mitotic and endoreplicative cells by distinct mechanisms. *Development*, 125(11), 2149–2158. <https://doi.org/10.1242/dev.125.11.2149>
- Cai, Y., Wang, Y., He, Y., Ren, K., Liu, Z., Zhao, L., & Wei, T. (2025). Utilizing alternative in vivo animal models for food safety and toxicity: A focus on thermal process contaminant acrylamide. *Food Chemistry*, 465, 142135. <https://doi.org/10.1016/j.foodchem.2024.142135>
- Cano-Lamadrid, M., Martínez-Zamora, L., Mozafari, L., Bueso, M. C., Kessler, M., & Artés-Hernández, F. (2023). Response Surface Methodology to Optimize the Extraction of Carotenoids from Horticultural By-Products—A Systematic Review. *Foods*, 12(24), Article 24. <https://doi.org/10.3390/foods12244456>
- Cao, P.-H., Zhang, C.-X., Ma, Y.-X., Yu, Y.-M., Liu, H.-M., Wang, X.-D., & Zheng, Y.-Z. (2023). Extraction of protein from sesame meal: Impact of deep eutectic

- solvents on protein structure and functionality. *LWT*, 187, 115366. <https://doi.org/10.1016/j.lwt.2023.115366>
- Castro, H., Costa, F., Ferreira, L., Ávila, P., Putnik, G. D., & Cruz-Cunha, M. (2022). Data Science for Industry 4.0: A Literature Review on Open Design Approach. *Procedia Computer Science*, 204, 877–884. <https://doi.org/10.1016/j.procs.2022.08.106>
- Catsimpoolas, N., Kenney, J. A., Meyer, E. W., & Szuhaj, B. F. (1971). Molecular weight and amino acid composition of glycinin subunits. *Journal of the Science of Food and Agriculture*, 22(9), 448–450. <https://doi.org/10.1002/jsfa.2740220905>
- Chakrabarti, A., & Ghosh, J. K. (2011). AIC, BIC and Recent Advances in Model Selection. In *Philosophy of Statistics* (Vol. 1–Book, Section, pp. 583–605). Elsevier B.V. <https://doi.org/10.1016/B978-0-444-51862-0.50018-6>
- 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>
- Chao, D., & Aluko, R. E. (2018). Modification of the structural, emulsifying, and foaming properties of an isolated pea protein by thermal pretreatment. *CYTA: Journal of Food*, 16(1), 357–366. <https://doi.org/10.1080/19476337.2017.1406536>
- Chaveroux, C., Lambert-Langlais, S., Cherasse, Y., Averous, J., Parry, L., Carraro, V., Jousse, C., Maurin, A.-C., Bruhat, A., & Fafournoux, P. (2010). Molecular mechanisms involved in the adaptation to amino acid limitation in mammals. *Biochimie*, 92(7), 736–745. <https://doi.org/10.1016/j.biochi.2010.02.020>
- Chen, B., McClements, D. J., Gray, D. A., & Decker, E. A. (2012). Physical and oxidative stability of pre-emulsified oil bodies extracted from soybeans. *Food Chemistry*, 132(3), 1514–1520. <https://doi.org/10.1016/j.foodchem.2011.11.144>
- Chen, C.-C., Wang, C.-Y., & Huang, Y.-H. (2021). Reversible absorption of nitrogen dioxide by choline chloride-based deep eutectic solvents and their aqueous mixtures. *Chemical Engineering Journal*, 405, 126760. <https://doi.org/10.1016/j.cej.2020.126760>
- Chen, N., Zhao, M., Chassenieux, C., & Nicolai, T. (2016). Data on the characterization of native soy globulin by SDS-Page, light scattering and titration. *Data in Brief*, 9, 749–752. <https://doi.org/10.1016/j.dib.2016.10.016>
- Chen, Q., Chaihu, L., Yao, X., Cao, X., Bi, W., Lin, J., & Chen, D. D. Y. (2021). 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>
- Chen, S., & Zheng, W. (2025). RRMSE-enhanced weighted voting regressor for improved ensemble regression. *PLOS ONE*, 20(3), e0319515. <https://doi.org/10.1371/journal.pone.0319515>
- Chen, Y., Hong, R., Rao, D., Han, Y., Chen, G., & Dong, X. (2023). Insect-Resistance of *Aquilaria sinensis* (Lour.) Leaves is Associated with Volatile Compounds. *Russian Journal of Plant Physiology*, 70(4), 60. <https://doi.org/10.1134/S1021443723700164>
- Cheng, J., Tang, J.-C., Pan, M.-X., Chen, S.-F., Zhao, D., Zhang, Y., Liao, H.-B., Zhuang, Y., Lei, R.-X., Wang, S., Liu, A.-C., Chen, J., Zhang, Z.-H., Li, H.-T., Wan, Q., & Chen, Q.-X. (2020). L-lysine confers neuroprotection by suppressing inflammatory response via microRNA-575/PTEN signaling after mouse

- intracerebral hemorrhage injury. *Experimental Neurology*, 327(C), 113214. <https://doi.org/10.1016/j.expneurol.2020.113214>
- Cheng, L. Y., Bailey, A. P., Leever, S. J., Ragan, T. J., Driscoll, P. C., & Gould, A. P. (2011). Anaplastic Lymphoma Kinase Spares Organ Growth during Nutrient Restriction in *Drosophila*. *Cell*, 146(3), 435–447. <https://doi.org/10.1016/j.cell.2011.06.040>
- Chew, L. Y., & Toh, J. T., & Ismail, A. (2018). *Application of Proteases for the Production of Bioactive Peptides* (A. Press, Ed.; pp. 247–261). Elsevier Science & Technology. <https://doi.org/10.1016/B978-0-12-813280-7.00015-3>
- 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>
- Cobb, J. S., Seale, M. A., & Janorkar, A. V. (2021). Evaluation of machine learning algorithms to predict the hydrodynamic radii and transition temperatures of chemo-biologically synthesized copolymers. *Computers in Biology and Medicine*, 128(Journal Article), 104134. <https://doi.org/10.1016/j.combiomed.2020.104134>
- Cobham, A. E., Neumann, B., & Mirth, C. K. (2022). Maintaining robust size across environmental conditions through plastic brain growth dynamics. *Open Biology*, 12(9), 220037. <https://doi.org/10.1098/rsob.220037>
- Colgrave, M. L., Dominik, S., Tobin, A. B., Stockmann, R., Simon, C., Howitt, C. A., Belobrajdic, D. P., Paull, C., & Vanhercke, T. (2021). Perspectives on Future Protein Production. *Journal of Agricultural and Food Chemistry*, 69(50), 15076–15083. <https://doi.org/10.1021/acs.jafc.1c05989>
- Colombani, J., Bianchini, L., Layalle, S., Pondeville, E., Dauphin-Villemant, C., Antoniewski, C., Carré, C., Noselli, S., & Léopold, P. (2005). Antagonistic Actions of Ecdysone and Insulins Determine Final Size in *Drosophila*. *Science*, 310(5748), 667–670. <https://doi.org/10.1126/science.1119432>
- Cong, Y. J., & Li, L. F. (2012a). Identification of the critical amino acid residues of immunoglobulin E and immunoglobulin G epitopes in β -lactoglobulin by alanine scanning analysis. *Journal of Dairy Science*, 95(11), 6307–6312. <https://doi.org/10.3168/jds.2012-5543>
- Cong, Y. J., & Li, L. F. (2012b). Identification of the critical amino acid residues of immunoglobulin E and immunoglobulin G epitopes in β -lactoglobulin by alanine scanning analysis. *Journal of Dairy Science*, 95(11), 6307–6312. <https://doi.org/10.3168/jds.2012-5543>
- Croset, V., Schleyer, M., Arguello, J. R., Gerber, B., & Benton, R. (2016). A molecular and neuronal basis for amino acid sensing in the *Drosophila* larva. *Scientific Reports*, 6(1), 34871. <https://doi.org/10.1038/srep34871>
- Dai, Y., Witkamp, G.-J., Verpoorte, R., & Choi, Y. H. (2015). Tailoring properties of natural deep eutectic solvents with water to facilitate their applications. *Food Chemistry*, 187, 14–19. <https://doi.org/10.1016/j.foodchem.2015.03.123>
- Dall’antonia, F., Pavkov-Keller, T., Zangger, K., & Keller, W. (2014). Structure of allergens and structure based epitope predictions. *Methods (San Diego, Calif.)*, 66(1), 3–21. <https://doi.org/10.1016/j.ymeth.2013.07.024>
- Delgadillo, J., & Atzil-Slonim, D. (2023). Artificial intelligence, machine learning and mental health. In H. S. Friedman & C. H. Markey (Eds.), *Encyclopedia of Mental*

- Health (Third Edition)* (Vol. 1–Book, Section, pp. 132–142). Academic Press. <https://doi.org/10.1016/B978-0-323-91497-0.00177-6>
- Demirgul, F., Simsek, O., & Sagdic, O. (2022). Amino acid, mineral, vitamin B contents and bioactivities of extracts of yeasts isolated from sourdough. *Food Bioscience*, 50, 102040. <https://doi.org/10.1016/j.fbio.2022.102040>
- Dhaliwal, D. S., & Williams, M. M. (2024). Sweet corn yield prediction using machine learning models and field-level data. *Precision Agriculture*, 25(1), 51–64. <https://doi.org/10.1007/s11119-023-10057-1>
- Dilawari, R., Kaur, N., Priyadarshi, N., Prakash, I., Patra, A., Mehta, S., Singh, B., Jain, P., & Islam, M. A. (2022). *Soybean: A Key Player for Global Food Security* (pp. 1–46). Springer International Publishing AG. https://doi.org/10.1007/978-3-031-12232-3_1
- Doğan, C., Güney, G., Güzel, K. K., Can, A., Hegedus, D. D., & Toprak, U. (2021). What You Eat Matters: Nutrient Inputs Alter the Metabolism and Neuropeptide Expression in Egyptian Cotton Leaf Worm, *Spodoptera littoralis* (Lepidoptera: Noctuidae). *Frontiers in Physiology*, 12. <https://doi.org/10.3389/fphys.2021.773688>
- 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>
- dos Santos Rodrigues, F., Delgado, G. G., Santana da Costa, T., & Tasic, L. (2023). Applications of fluorescence spectroscopy in protein conformational changes and intermolecular contacts. *BBA Advances*, 3(Journal Article). <https://doi.org/10.1016/j.bbadv.2023.100091>
- dos Santos Rodrigues, F. H., Delgado, G. G., Santana da Costa, T., & Tasic, L. (2023). Applications of fluorescence spectroscopy in protein conformational changes and intermolecular contacts. *BBA Advances*, 3, 100091. <https://doi.org/10.1016/j.bbadv.2023.100091>
- Dumitraşcu, L., Borda, D., & Aprodu, I. (2023). Alternative Processing Options for Improving the Proteins Functionality by Maillard Conjugation. *Foods*, 12(19), 3588. <https://doi.org/10.3390/foods12193588>
- Duque-Estrada, P., & Petersen, I. L. (2023). The sustainability paradox of processing plant proteins. *Npj Science of Food*, 7(1), 38. <https://doi.org/10.1038/s41538-023-00214-1>
- Dutta, S., Ray, S., & Nagarajan, K. (2013). Glutamic acid as anticancer agent: An overview. *Saudi Pharmaceutical Journal*, 21(4), 337–343. <https://doi.org/10.1016/j.jsps.2012.12.007>
- Dziak, J. J., Coffman, D. L., Lanza, S. T., Li, R., & Jermin, L. S. (2020). Sensitivity and specificity of information criteria. *Briefings in Bioinformatics*, 21(2), 553–565. <https://doi.org/10.1093/bib/bbz016>
- Elgin, S. C. R., & Reuter, G. (2013). Position-Effect Variegation, Heterochromatin Formation, and Gene Silencing in *Drosophila*. *Cold Spring Harbor Perspectives in Biology*, 5(8), a017780. <https://doi.org/10.1101/cshperspect.a017780>
- 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>

- Erdebilli, B., & Devrim-İçtenbaş, B. (2022). Ensemble Voting Regression Based on Machine Learning for Predicting Medical Waste: A Case from Turkey. *Mathematics*, 10(14), Article 14. <https://doi.org/10.3390/math10142466>
- Evans, B. A., & Howells, A. J. (1978). Control of drosoplerin synthesis in *Drosophila melanogaster*: Mutants showing an altered pattern of GTP cyclohydrolase activity during development. *Biochemical Genetics*, 16(1–2), 13–26. <https://doi.org/10.1007/BF00484381>
- 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>
- Ewy, M. W., Patel, A., Abdelmagid, M. G., Mohamed Elfadil, O., Bonnes, S. L., Salonen, B. R., Hurt, R. T., & Mundi, M. S. (2022). Plant-Based Diet: Is It as Good as an Animal-Based Diet When It Comes to Protein? *Current Nutrition Reports*, 11(2), 337–346. <https://doi.org/10.1007/s13668-022-00401-8>
- Eze, C. R., Kwofie, E. M., Adewale, P., Lam, E., & Ngadi, M. (2022). Advances in legume protein extraction technologies: A review. *Innovative Food Science & Emerging Technologies*, 82, 103199. <https://doi.org/10.1016/j.ifset.2022.103199>
- Fadimu, G. J., Olatunde, O. O., Bandara, N., & Truong, T. (2023). *Chapter 4—Reducing allergenicity in plant-based proteins* (pp. 61–77). Elsevier Inc. <https://doi.org/10.1016/B978-0-323-89842-3.00012-9>
- FAOSTAT. (2025). <https://www.fao.org/faostat/en/#data/QCL>
- Fatima, A., Singh, P., Pandey, V. K., Singh, R., & Rustagi, S. (2024). Exploring the significance of protein concentrate: A review on sources, extraction methods, and applications. *Food Chemistry Advances*, 5, 100771. <https://doi.org/10.1016/j.focha.2024.100771>
- Feng, J., Wang, H., Luo, X., Zhang, L., & Zhou, P. (2024). Identification and molecular mechanism of the anti-inflammatory effect of sea cucumber peptides: Network pharmacology, molecular docking and animal experiments. *International Journal of Biological Macromolecules*, 134958. <https://doi.org/10.1016/j.ijbiomac.2024.134958>
- Feng, Y., Ma, X., Kong, B., Chen, Q., & Liu, Q. (2021). Ethanol induced changes in structural, morphological, and functional properties of whey proteins isolates: Influence of ethanol concentration. *Food Hydrocolloids*, 111, 106379. <https://doi.org/10.1016/j.foodhyd.2020.106379>
- Ferreira, A. S. D., Craveiro, R., Duarte, A. R., Barreiros, S., Cabrita, E. J., & Paiva, A. (2021). Effect of water on the structure and dynamics of choline chloride/glycerol eutectic systems. *Journal of Molecular Liquids*, 342, 117463. <https://doi.org/10.1016/j.molliq.2021.117463>
- Fini, A. S., & Gharehghani, A. (2024). Experimental investigation of pressure effect on the PCM performance in Li-ion battery thermal management system. *Journal of Energy Storage*, 79, 110273. <https://doi.org/10.1016/j.est.2023.110273>
- Fonseca Hernandez, D., Mojica, L., & Gonzalez de Mejia, E. (2024). Legume-derived bioactive peptides: Role in cardiovascular disease prevention and control. *Current Opinion in Food Science*, 56, 101132. <https://doi.org/10.1016/j.cofs.2024.101132>
- Fraga, M. F., Ballestar, E., Paz, M. F., Ropero, S., Setien, F., Ballestar, M. L., Heine-Suñer, D., Cigudosa, J. C., Urioste, M., Benitez, J., Boix-Chornet, M., Sanchez-Aguilera, A., Ling, C., Carlsson, E., Poulsen, P., Vaag, A., Stephan, Z., Spector, T. D., Wu, Y.-Z., ... Esteller, M. (2005). Epigenetic differences arise during the lifetime of monozygotic twins. *Proceedings of the National Academy of Sciences*

- of the United States of America, 102(30), 10604–10609. <https://doi.org/10.1073/pnas.0500398102>
- Franchet, A., Jin, Y., Newell, C. L., Girard, V., Manière, G., Grosjean, Y., Barrington, C., MacRae, J. I., Gilmore, I. S., & Gould, A. P. (2025). *An intercellular metabolic relay for brain sparing in Drosophila* (p. 2025.04.02.645514). bioRxiv. <https://doi.org/10.1101/2025.04.02.645514>
- Fukagawa, N. K. (2013). Protein and amino acid supplementation in older humans. *Amino Acids*, 44(6), 1493–1509. <https://doi.org/10.1007/s00726-013-1480-6>
- Gabriele, F., Chiarini, M., Germani, R., Tiecco, M., & Spreti, N. (2019). Effect of water addition on choline chloride/glycol deep eutectic solvents: Characterization of their structural and physicochemical properties. *Journal of Molecular Liquids*, 291, 111301. <https://doi.org/10.1016/j.molliq.2019.111301>
- Galagovsky, D., Depetris-Chauvin, A., Manière, G., Geillon, F., Berthelot-Grosjean, M., Noirot, E., Alves, G., & Grosjean, Y. (2018). Sobremesa L-type Amino Acid Transporter Expressed in Glia Is Essential for Proper Timing of Development and Brain Growth. *Cell Reports*, 24(12), 3156–3166.e4. <https://doi.org/10.1016/j.celrep.2018.08.067>
- Galeazzi, A., Prifti, K., Gallo, F., & Manenti, F. (2022). A Methodology for The Optimal Surrogate Modelling of Digital Twins Using Machine Learning. In L. Montastruc & S. Negny (Eds.), *Computer Aided Chemical Engineering* (Vol. 51, pp. 1543–1548). Elsevier. <https://doi.org/10.1016/B978-0-323-95879-0.50258-7>
- 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>
- Giansanti, M. G., Frappalo, A., & Piergentili, R. (2025). *Drosophila melanogaster: How and Why It Became a Model Organism. International Journal of Molecular Sciences*, 26(15), 7485. <https://doi.org/10.3390/ijms26157485>
- Gilani, G. S., & Lee, N. (2003). *PROTEIN | Sources of Food-grade Protein* (pp. 4873–4879). Elsevier Ltd. <https://doi.org/10.1016/B0-12-227055-X/00834-8>
- Giorgio, L. D., Marcantonio, M. A., Salgado, P. R., & Mauri, A. N. (2023). Chapter 23—*Isolation, characterization, and industrial processing of soybean proteins* (pp. 557–575). Elsevier Inc. <https://doi.org/10.1016/B978-0-323-99853-6.00018-8>
- Gokhale, A. S., & Satyanarayanajois, S. (2014). Peptides and peptidomimetics as immunomodulators. *Immunotherapy*, 6(6), 755–774. <https://doi.org/10.2217/imt.14.37>
- 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>
- Goto, K., Tamehiro, N., Yoshida, T., Hanada, H., Sakuma, T., Adachi, R., Kondo, K., & Takeuchi, I. (2023). Novel machine learning method allerStat identifies statistically significant allergen-specific patterns in protein sequences. *Journal of Biological Chemistry*, 299(6), 104733. <https://doi.org/10.1016/j.jbc.2023.104733>
- Grudniewska, A., de Melo, E. M., Chan, A., Gniłka, R., Boratyński, F., & Matharu, A. S. (2018). Enhanced Protein Extraction from Oilseed Cakes Using Glycerol–Choline Chloride Deep Eutectic Solvents: A Biorefinery Approach. *ACS Sustainable Chemistry & Engineering*, 6(11), 15791–15800. <https://doi.org/10.1021/acssuschemeng.8b04359>
- 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](https://doi.org/10.1016/S1043-6618(05)80029-5)

- Guha, S., & Majumder, K. (2019). Structural-features of food-derived bioactive peptides with anti-inflammatory activity: A brief review. *Journal of Food Biochemistry*, 43(1), e12531-n/a. <https://doi.org/10.1111/jfbc.12531>
- Gunarathne, R., Guan, X., Feng, T., Zhao, Y., & Lu, J. (2025). L-lysine dietary supplementation for childhood and adolescent growth: Promises and precautions. *Journal of Advanced Research*, 70, 571–586. <https://doi.org/10.1016/j.jare.2024.05.014>
- Guo, J., Zhao, N., Zhao, Y., Jin, H., Sun, G., Yu, J., Zhang, H., Yu, M., Yang, D., & Liang, Z. (2023). Effect of Deep Eutectic Solvents on the Activity and Stability of Cellulases and Pectinases. *ACS Omega*, 8(48), 45678–45686. <https://doi.org/10.1021/acsomega.3c06088>
- Hadidi, M., Aghababaei, F., Gonzalez-Serrano, D. J., Goksen, G., Trif, M., McClements, D. J., & Moreno, A. (2024). Plant-based proteins from agro-industrial waste and by-products: Towards a more circular economy. *International Journal of Biological Macromolecules*, 261(Pt 1), 129576. <https://doi.org/10.1016/j.ijbiomac.2024.129576>
- 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>
- Halmi, M. I. E., Shukor, M. S., Masdor, N. A., Shamaan, N. A., Sabullah, M. K., & Shukor, M. Y. (2015). Testing the Normality of Residuals on Regression Model for the Growth of *Paracoccus* sp. SKG on Acetonitrile. *Journal of Environmental Bioremediation and Toxicology*, 3(1), 15–17. <https://doi.org/10.54987/jebat.v3i1.255>
- Hammond, O. S., Bowron, D. T., & Edler, K. J. (2017). The Effect of Water upon Deep Eutectic Solvent Nanostructure: An Unusual Transition from Ionic Mixture to Aqueous Solution. *Angewandte Chemie International Edition*, 56(33), 9782–9785. <https://doi.org/10.1002/anie.201702486>
- Han, H., Yin, J., Wang, B., Huang, X., Yao, J., Zheng, J., Fan, W., Li, T., & Yin, Y. (2018). Effects of dietary lysine restriction on inflammatory responses in piglets. *Scientific Reports*, 8(1), 2451–2458. <https://doi.org/10.1038/s41598-018-20689-3>
- Han, J., Kamber, M., & Pei, J. (2012). 2—Getting to Know Your Data. In J. Han, M. Kamber, & J. Pei (Eds.), *Data Mining: Concepts and Techniques (Third Edition)* (pp. 39–82). Morgan Kaufmann. <https://doi.org/10.1016/B978-0-12-381479-1.00002-2>
- Hanafi, M. A., Anwar, F., & Saari, N. (2024). Valorization of biomass for food protein via deep eutectic solvent extraction: Understanding the extraction mechanism and impact on protein structure and properties. *Food Frontiers*, 5(3), 1265–1301. <https://doi.org/10.1002/fft2.389>
- Hansen, B. B., Spittle, S., Chen, B., Poe, D., Zhang, Y., Klein, J. M., Horton, A., Adhikari, L., Zelovich, T., Doherty, B. W., Gurkan, B., Maginn, E. J., Ragauskas, A., Dadmun, M., Zawodzinski, T. A., Baker, G. A., Tuckerman, M. E., Savinell, R. F., & Sangoro, J. R. (2021). Deep Eutectic Solvents: A Review of Fundamentals and Applications. *Chemical Reviews*, 121(3), 1232–1285. <https://doi.org/10.1021/acs.chemrev.0c00385>
- Hartmann, R., Wal, J.-M., Bernard, H., & Pentzien, A.-K. (2007). Cytotoxic and Allergenic Potential of Bioactive Proteins and Peptides. *Current Pharmaceutical Design*, 13(9), 897–920. <https://doi.org/10.2174/138161207780414232>
- 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>
- Hazra, S., Sultana, S., & Roy, P. K. (2024). *Optimization Techniques for Hybrid Power Systems* (1st ed.). IGI Global. [https://ebookcentral.proquest.com/lib/\[SITE_ID\]/detail.action?docID=31537504](https://ebookcentral.proquest.com/lib/[SITE_ID]/detail.action?docID=31537504)
- He, Y., & Tao, A. (2015). Bioinformatics Methods to Predict Allergen Epitopes. In A. Tao & E. Raz (Eds.), *Allergy Bioinformatics* (pp. 223–238). Springer Netherlands. https://doi.org/10.1007/978-94-017-7444-4_14
- Hemshekhhar, M., Faiyaz, S., Choi, K.-Y. G., Krokhin, O. V., & Mookherjee, N. (2019). Immunomodulatory Functions of the Human Cathelicidin LL-37 (aa 13-31)-Derived Peptides are Associated with Predicted α -Helical Propensity and Hydrophobic Index. *Biomolecules*, 9(9), 501. <https://doi.org/10.3390/biom9090501>
- Hewage, A., Olatunde, O. O., Nimalaratne, C., House, J. D., Aluko, R. E., & Bandara, N. (2024). Improved protein extraction technology using deep eutectic solvent system for producing high purity fava bean protein isolates at mild conditions. *Food Hydrocolloids*, 147, 109283. <https://doi.org/10.1016/j.foodhyd.2023.109283>
- Hewage, A., Olatunde, O. O., Nimalaratne, C., Malalgoda, M., Aluko, R. E., & Bandara, N. (2022). Novel Extraction technologies for developing plant protein ingredients with improved functionality. *Trends in Food Science & Technology*, 129, 492–511. <https://doi.org/10.1016/j.tifs.2022.10.016>
- Hiruma, K., & Kaneko, Y. (2013). Chapter Three—Hormonal Regulation of Insect Metamorphosis with Special Reference to Juvenile Hormone Biosynthesis. In Y.-B. Shi (Ed.), *Current Topics in Developmental Biology* (Vol. 103, pp. 73–100). Academic Press. <https://doi.org/10.1016/B978-0-12-385979-2.00003-4>
- Hope, T. M. H. (2020). Chapter 4—Linear regression. In A. Mechelli & S. Vieira (Eds.), *Machine Learning* (pp. 67–81). Academic Press. <https://doi.org/10.1016/B978-0-12-815739-8.00004-3>
- Hoyer-Fender, S. (2020). Transgenerational effect of drug-mediated inhibition of LSD1 on eye pigment expression in *Drosophila*. *BMC Ecology*, 20(1), 62. <https://doi.org/10.1186/s12898-020-00330-6>
- Hsiao, Y.-H., Yu, C.-J., Li, W.-T., & Hsieh, J.-F. (2015). Coagulation of β -conglycinin, glycinin and isoflavones induced by calcium chloride in soymilk. *Scientific Reports*, 5(1), 13018. <https://doi.org/10.1038/srep13018>
- Hu, H., Wu, J., Li-Chan, E. C. Y., Zhu, L., Zhang, F., Xu, X., Fan, G., Wang, L., Huang, X., & Pan, S. (2013). Effects of ultrasound on structural and physical properties of soy protein isolate (SPI) dispersions. *Food Hydrocolloids*, 30(2), 647–655. <https://doi.org/10.1016/j.foodhyd.2012.08.001>
- Hu, S.-Y., Gu, Q.-H., Wang, J., Wang, M., Jia, X.-Y., Cui, Z., & Zhao, M.-H. (2017). The pathogenicity of T cell epitopes on human Goodpasture antigen and its critical amino acid motif. *Journal of Cellular and Molecular Medicine*, 21(9), 2117–2128. <https://doi.org/10.1111/jcmm.13134>
- Huang, J., Zhang, M., Mujumdar, A. S., Semenov, G., & Luo, Z. (2024). Technological advances in protein extraction, structure improvement and assembly, digestibility and bioavailability of plant-based foods. *Critical Reviews in Food Science and Nutrition*, 64(31), 11556–11574. <https://doi.org/10.1080/10408398.2023.2240892>

- Huang, L., Zhang, W., Ding, X., Wu, Z., & Li, Y. (2020). Effects of dual-frequency ultrasound with different energy irradiation modes on the structural and emulsifying properties of soy protein isolate. *Food and Bioprocess Processing*, 123(Journal Article), 419–426. <https://doi.org/10.1016/j.fbp.2020.07.021>
- Huang, Z., Qu, Y., Hua, X., Wang, F., Jia, X., & Yin, L. (2023). Recent advances in soybean protein processing technologies: A review of preparation, alterations in the conformational and functional properties. *International Journal of Biological Macromolecules*, 248, 125862. <https://doi.org/10.1016/j.ijbiomac.2023.125862>
- Badar, IH., Wang, Z., Chen, Q., Liu, Q., Ma, J., Liu, H., & Kong, B. (2024). Ultrasonic enhancement of structural and emulsifying properties of heat-treated soy protein isolate nanoparticles to fabricate flaxseed-derived diglyceride-based pickering emulsions. *Food Chemistry*, 442(Journal Article), 138469. <https://doi.org/10.1016/j.foodchem.2024.138469>
- IEDB Analysis Resource. (n.d.). Retrieved December 8, 2025, from <https://tools.iedb.org/main/>
- 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/>
- Jablonka, K. M., Schwaller, P., Ortega-Guerrero, A., & Smit, B. (2024). Leveraging large language models for predictive chemistry. *Nature Machine Intelligence*, 6(2), 161–169. <https://doi.org/10.1038/s42256-023-00788-1>
- Jain, A., Jain, R., & Jain, S. (2020). Protein Analysis in Food. In A. Jain, R. Jain, & S. Jain (Eds.), *Basic Techniques in Biochemistry, Microbiology and Molecular Biology: Principles and Techniques* (pp. 167–175). Springer US. https://doi.org/10.1007/978-1-4939-9861-6_41
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An Introduction to Statistical Learning*. Springer New York. <https://doi.org/10.1007/978-1-4614-7138-7>
- Jennings, B. H. (2011). *Drosophila* – a versatile model in biology & medicine. *Materials Today*, 14(5), 190–195. [https://doi.org/10.1016/S1369-7021\(11\)70113-4](https://doi.org/10.1016/S1369-7021(11)70113-4)
- Jiang, F., Wang, S., Han, D., Wei, J., Wu, Y., & Liu, Z. (2024). Glutamic-acid grafted hyaluronic acid inhibits inflammatory factors via fibroblast and skin model tests. *Skin Research and Technology*, 30(1), e13548-n/a. <https://doi.org/10.1111/srt.13548>
- Jiang, L., Wang, J., Li, Y., Wang, Z., Liang, J., Wang, R., Chen, Y., Ma, W., Qi, B., & Zhang, M. (2014). Effects of ultrasound on the structure and physical properties of black bean protein isolates. *Food Research International*, 62(Journal Article), 595–601. <https://doi.org/10.1016/j.foodres.2014.04.022>
- 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>
- Jones, J. D. (1961). Lysine Toxicity in the Chick. *The Journal of Nutrition*, 73(2), 107–112. <https://doi.org/10.1093/jn/73.2.107>
- Joshi, A., Virdi, A. S., Kaur, R., Kumar, A., & Singh, N. (2025). Deciphering food proteins: The applications of SDS-PAGE in food science. *Food Bioscience*, 66, 106126. <https://doi.org/10.1016/j.fbio.2025.106126>
- Kadyan, A., Juneja, S., & Pandey, S. (2019). Photophysical Behavior and Fluorescence Quenching of l-Tryptophan in Choline Chloride-Based Deep Eutectic Solvents.

- The Journal of Physical Chemistry B*, 123(35), 7578–7587. <https://doi.org/10.1021/acs.jpcc.9b04659>
- Kaláb, M. (1995). Practical Aspects of Electron Microscopy in Cheese Research. In E. L. Malin & M. H. Tunick (Eds.), *Chemistry of Structure-Function Relationships in Cheese* (pp. 247–276). Springer US. https://doi.org/10.1007/978-1-4615-1913-3_15
- Kamal, N., Chowdhury, S., Madan, T., Sharma, D., Attreyi, M., Haq, W., Katti, S. B., Kumar, A., & Sarma, P. U. (2005). Tryptophan residue is essential for immunoreactivity of a diagnostically relevant peptide epitope of *A. fumigatus*. *Molecular and Cellular Biochemistry*, 275(1–2), 223–231. <https://doi.org/10.1007/s11010-005-2056-x>
- Kapp, S., Choi, J.-K., & Hong, T. (2023). Predicting industrial building energy consumption with statistical and machine-learning models informed by physical system parameters. *Renewable & Sustainable Energy Reviews*, 172(C), 113045. <https://doi.org/10.1016/j.rser.2022.113045>
- Karki, B., Lamsal, B. P., Grewell, D., Pometto, A. L. I., van Leeuwen, J., Khanal, S. K., & Jung, S. (2009). Functional Properties of Soy Protein Isolates Produced from Ultrasonicated Defatted Soy Flakes. *Journal of the American Oil Chemists' Society*, 86(10), 1021–1028. <https://doi.org/10.1007/s11746-009-1433-0>
- Keerati-u-rai, M., & Corredig, M. (2011). 4.45 - Soy Protein Functionality: Emulsion and Gels. In M. Moo-Young (Ed.), *Comprehensive Biotechnology (Second Edition)* (pp. 543–551). Academic Press. <https://doi.org/10.1016/B978-0-08-088504-9.00324-X>
- Khalil, N. Z., Kong, N., & Fricke, H. (2024). The influence of GNP on the mechanical and thermomechanical properties of epoxy adhesive: Pearson correlation matrix and heatmap application in data interpretation. *Polymer Composites*, 45(10), 8997–9018. <https://doi.org/10.1002/pc.28390>
- Khetto, A., Manikpuri, S., Sarkar, A., Das, R., Bebartta, R. P., Kumar, Y., Bist, Y., Vashishth, R., & Sehrawat, R. (2024). How pulse electric field treatment affects anti-nutritional factors and plant protein digestibility: A concise review. *Food Bioscience*, 61, 104849. <https://doi.org/10.1016/j.fbio.2024.104849>
- Kikugawa, K., & Beppu, M. (1987). Involvement of lipid oxidation products in the formation of fluorescent and cross-linked proteins. *Chemistry and Physics of Lipids*, 44(2), 277–296. [https://doi.org/10.1016/0009-3084\(87\)90054-5](https://doi.org/10.1016/0009-3084(87)90054-5)
- Kim, W., Kim, M., & Park, W. (2023). Unlocking the mystery of lysine toxicity on *Microcystis aeruginosa*. *Journal of Hazardous Materials*, 448, 130932. <https://doi.org/10.1016/j.jhazmat.2023.130932>
- Klepsatel, P., Procházka, E., & Gálíková, M. (2018). Crowding of *Drosophila* larvae affects lifespan and other life-history traits via reduced availability of dietary yeast. *Experimental Gerontology*, 110, 298–308. <https://doi.org/10.1016/j.exger.2018.06.016>
- 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>
- Koethe, M., Taubert, J., & Vervuert, I. (2022). Impact of lysine supplementation on growth and development of *Hermetia illucens* larvae. *Journal of Insects as Food and Feed*, 8(1), 35–44. <https://doi.org/10.3920/JIFF2020.0059>
- Kogay, R., & Schönbach, C. (2019). Epitope predictions. In S. Ranganathan, M. Gribskov, K. Nakai, & C. Schönbach (Eds.), *Encyclopedia of Bioinformatics and*

- Computational Biology* (Vol. 2, p. 952). Elsevier. <https://doi.org/10.1016/B978-0-12-809633-8.20248-3>
- Kong, X., Guo, M., Hua, Y., Cao, D., & Zhang, C. (2008). Enzymatic preparation of immunomodulating hydrolysates from soy proteins. *Bioresource Technology*, 99(18), 8873–8879. <https://doi.org/10.1016/j.biortech.2008.04.056>
- Kong, X., Jiang, X., Zhang, B., Yuan, J., & Ge, Z. (2022). Latent variable models in the era of industrial big data: Extension and beyond. *Annual Reviews in Control*, 54, 167–199. <https://doi.org/10.1016/j.arcontrol.2022.09.005>
- Kotu, V., & Deshpande, B. (2015). Chapter 2—Data Mining Process. In V. Kotu & B. Deshpande (Eds.), *Predictive Analytics and Data Mining* (pp. 17–36). Morgan Kaufmann. <https://doi.org/10.1016/B978-0-12-801460-8.00002-1>
- Kotu, V., & Deshpande, B. (2019). Chapter 5—Regression Methods. In V. Kotu & B. Deshpande (Eds.), *Data Science (Second Edition)* (pp. 165–197). Morgan Kaufmann. <https://doi.org/10.1016/B978-0-12-814761-0.00005-8>
- Kouzehgar, K., & Eslamian, S. (2023). Chapter 2—Application of experimental data and soft computing techniques in determining the outflow and breach characteristics in embankments and landslide dams. In S. Eslamian & F. Eslamian (Eds.), *Handbook of Hydroinformatics* (pp. 11–31). Elsevier. <https://doi.org/10.1016/B978-0-12-821962-1.00002-7>
- Koyama, T., Texada, M. J., Halberg, K. A., & Rewitz, K. (2020). Metabolism and growth adaptation to environmental conditions in *Drosophila*. *Cellular and Molecular Life Sciences*, 77(22), 4523–4551. <https://doi.org/10.1007/s00018-020-03547-2>
- Kudow, N., Miura, D., Schleyer, M., Toshima, N., Gerber, B., & Tanimura, T. (2017). Preference for and learning of amino acids in larval *Drosophila*. *Biology Open*, 6(3), 365–369. <https://doi.org/10.1242/bio.020412>
- Kumar, A., Dave, M., Pant, D. C., Laxkar, R., & Tiwari, A. K. (2013). Vinca rosea leaf extract supplementation leads to developmental delay and several phenotypic anomalies in *Drosophila melanogaster*. *Toxicological & Environmental Chemistry*, 95(4), 635–645. <https://doi.org/10.1080/02772248.2013.806511>
- Kumar, M., Tomar, M., Potkule, J., Reetu, Punia, S., Dhakane-Lad, J., Singh, S., Dhumal, S., Chandra Pradhan, P., Bhushan, B., Anitha, T., Alajil, O., Alhariri, A., Amarowicz, R., & Kennedy, J. F. (2022). Functional characterization of plant-based protein to determine its quality for food applications. *Food Hydrocolloids*, 123, 106986. <https://doi.org/10.1016/j.foodhyd.2021.106986>
- Kumar, M., Tomar, M., Potkule, J., Verma, R., Punia, S., Mahapatra, A., Belwal, T., Dahuja, A., Joshi, S., Berwal, M. K., Satankar, V., Bhoite, A. G., Amarowicz, R., Kaur, C., & Kennedy, J. F. (2021). Advances in the plant protein extraction: Mechanism and recommendations. *Food Hydrocolloids*, 115, 106595. <https://doi.org/10.1016/j.foodhyd.2021.106595>
- Kutzli, I., Griener, D., Gibis, M., Schmid, C., Dawid, C., Baier, S. K., Hofmann, T., & Weiss, J. (2020). Influence of Maillard reaction conditions on the formation and solubility of pea protein isolate-maltodextrin conjugates in electrospun fibers. *Food Hydrocolloids*, 101, 105535. <https://doi.org/10.1016/j.foodhyd.2019.105535>
- Lakowicz, J. R. (Ed.). (2000). *Topics in Fluorescence Spectroscopy: Volume 6: Protein Fluorescence*. Springer US. <https://doi.org/10.1007/b115628>
- Lakowicz, J. R. (2006). *Principles of Fluorescence Spectroscopy* (3rd ed.). Springer US. <https://doi.org/10.1007/978-0-387-46312-4>

- Lan, T., Dong, Y., Jiang, L., Zhang, Y., & Sui, X. (2024). Analytical approaches for assessing protein structure in protein-rich food: A comprehensive review. *Food Chemistry: X*, 22, 101365. <https://doi.org/10.1016/j.fochx.2024.101365>
- Lanet, E., Gould, A. P., & Maurange, C. (2013). Protection of Neuronal Diversity at the Expense of Neuronal Numbers during Nutrient Restriction in the Drosophila Visual System. *Cell Reports*, 3(3), 587–594. <https://doi.org/10.1016/j.celrep.2013.02.006>
- Latimer, G. W., Jr. (Ed.). (2023). AOAC Official Method 2001.11 Protein (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>
- Lavrynenko, O., Rodenfels, J., Carvalho, M., Dye, N. A., Lafont, R., Eaton, S., & Shevchenko, A. (2015). The ecdysteroidome of Drosophila: Influence of diet and development. *Development (Cambridge, England)*, 142(21), 3758–3768. <https://doi.org/10.1242/dev.124982>
- Leitner, W., & Turner, W. R. (2017). Measurement and Analysis of Biodiversity☆. In *Reference Module in Life Sciences* (Vol. 1–Book, Section). Elsevier. <https://doi.org/10.1016/B978-0-12-809633-8.02385-2>
- 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, H., Li, W., Li, Y., & Li, Y. (2024). Machine learning-based multi-objective optimization and physical-geometrical competitive mechanisms for 3D woven thermal protection composites. *International Journal of Heat and Mass Transfer*, 232(Journal Article), 125920. <https://doi.org/10.1016/j.ijheatmasstransfer.2024.125920>
- 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](https://doi.org/10.1016/S0308-8146(98)00250-7)
- Lillycrop, K. A., Phillips, E. S., Jackson, A. A., Hanson, M. A., & Burdge, G. C. (2005). Dietary Protein Restriction of Pregnant Rats Induces and Folic Acid Supplementation Prevents Epigenetic Modification of Hepatic Gene Expression in the Offspring¹. *The Journal of Nutrition*, 135(6), 1382–1386. <https://doi.org/10.1093/jn/135.6.1382>
- Lima-Cabello, E., Robles, P., Alche, J., & Jimenez-Lopez, J. (2016). Narrow Leafed Lupin Beta-Conglutin Proteins Epitopes Identification and Molecular Features Analysis Involved in Cross-Allergenicity to Peanut and Other Legumes. *Genomics and Computational Biology*, 2, e29. <https://doi.org/10.18547/gcb.2016.vol2.iss1.e29>
- Lin, D., & Miao, S. (2021). Chapter 10—Interactions, structures, and functional properties of plant protein–polymer complexes. In C. M. Galanakis (Ed.), *Food Structure and Functionality* (pp. 201–217). Academic Press. <https://doi.org/10.1016/B978-0-12-821453-4.00004-1>
- Lin, Y., Xu, W., & Liu, Q. (2023). Editorial: Advances in color science: From color perception to color metrics and its applications in illuminated environments. *Frontiers in Psychology*, 14(Journal Article), 1142153. <https://doi.org/10.3389/fpsyg.2023.1142153>

- Lin, Z., Jiao, G., Zhang, J., Celli, G., & Brooks, M. (2021). Optimization of protein extraction from bamboo shoots and processing wastes using deep eutectic solvents in a biorefinery approach. *Biomass Conversion and Biorefinery*, 11. <https://doi.org/10.1007/s13399-020-00614-3>
- Ling, J. K. U., & Hadinoto, K. (2022). Deep Eutectic Solvent as Green Solvent in Extraction of Biological Macromolecules: A Review. *International Journal of Molecular Sciences*, 23(6), 3381. <https://doi.org/10.3390/ijms23063381>
- Liu, C., & Sathe, S. K. (2018). Food Allergen Epitope Mapping. *Journal of Agricultural and Food Chemistry*, 66(28), 7238–7248. <https://doi.org/10.1021/acs.jafc.8b01967>
- 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>
- Liu, H., Saito, Y., Riza, D. F. A., Kondo, N., Yang, X., & Han, D. (2019). Rapid evaluation of quality deterioration and freshness of beef during low temperature storage using three-dimensional fluorescence spectroscopy. *Food Chemistry*, 287, 369–374. <https://doi.org/10.1016/j.foodchem.2019.02.119>
- Liu, M., Huan, F., Han, T.-J., Liu, S.-H., Li, M.-S., Yang, Y., Wu, Y.-H., Chen, G.-X., Cao, M.-J., & Liu, G.-M. (2021). Combination Processing Method Reduced IgE-Binding Activity of *Litopenaeus vannamei* by Modifying Lysine, Arginine, and Cysteine on Multiple Allergen Epitopes. *Journal of Agricultural and Food Chemistry*, 69(16), 4865–4873. <https://doi.org/10.1021/acs.jafc.1c00718>
- Liu, W., Chen, X., Li, H., Zhang, J., An, J., & Liu, X. (2022). Anti-Inflammatory Function of Plant-Derived Bioactive Peptides: A Review. *Foods*, 11(15), 2361. <https://doi.org/10.3390/foods11152361>
- Liu, W., Wang, T., Zhao, K., Hanigan, M. D., Lin, X., Hu, Z., Hou, Q., Wang, Y., & Wang, Z. (2024). Effects of Individual Essential Amino Acids on Growth Rates of Young Rats Fed a Low-Protein Diet. *Animals*, 14(6), Article 6. <https://doi.org/10.3390/ani14060959>
- Loreto, J. S., Ferreira, S. A., Ardisson-Araújo, D. M., & Barbosa, N. V. (2021). Human type 2 diabetes mellitus-associated transcriptional disturbances in a high-sugar diet long-term exposed *Drosophila melanogaster*. *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics*, 39, 100866. <https://doi.org/10.1016/j.cbd.2021.100866>
- Lozano-Ojalvo, D., & López-Fandiño, R. (2018). Immunomodulating peptides for food allergy prevention and treatment. *Critical Reviews in Food Science and Nutrition*, 58(10), 1629–1649. <https://doi.org/10.1080/10408398.2016.1275519>
- Luan, H., Lu, J., Shi, W., & Lu, Y. (2024). Four amino acids play an important role in the allergenicity of hemocyanin allergen. *International Journal of Biological Macromolecules*, 275, 133704. <https://doi.org/10.1016/j.ijbiomac.2024.133704>
- Lundstedt, T., Seifert, E., Abramo, L., Thelin, B., Nyström, Å., Pettersen, J., & Bergman, R. (1998). Experimental design and optimization. *Chemometrics and Intelligent Laboratory Systems*, 42(1), 3–40. [https://doi.org/10.1016/S0169-7439\(98\)00065-3](https://doi.org/10.1016/S0169-7439(98)00065-3)
- Ma, C., Laaksonen, A., Liu, C., Lu, X., & Ji, X. (2018). The peculiar effect of water on ionic liquids and deep eutectic solvents. *Chemical Society Reviews*, 47(23), 8685–8720. <https://doi.org/10.1039/C8CS00325D>
- Ma, Z., Li, L., Wu, C., Huang, Y., Teng, F., & Li, Y. (2022). Effects of combined enzymatic and ultrasonic treatments on the structure and gel properties of soybean

- protein isolate. *Food Science & Technology*, 158(Journal Article), 113123. <https://doi.org/10.1016/j.lwt.2022.113123>
- Madugula, S., Pratapagiri, S., Phridviraj, M. S. B., Rao, V. C. S., Polala, N., & Kumaraswamy, P. (2023). Big data for the comprehensive data analysis of IT organizations. *The Journal of High Technology Management Research*, 34(2), 100465. <https://doi.org/10.1016/j.hitech.2023.100465>
- Maity, A., Sarkar, S., Theeyancheri, L., & Chakrabarti, R. (2020). Choline Chloride as a Nano-Crowder Protects HP-36 from Urea-Induced Denaturation: Insights from Solvent Dynamics and Protein-Solvent Interactions. *ChemPhysChem*, 21(6), 552–567. <https://doi.org/10.1002/cphc.201901078>
- 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>
- Maleh, Y., Baddi, Y., Alazab, M., Tawalbeh, L., & Romdhani, I. (2021). *Artificial Intelligence and Blockchain for Future Cybersecurity Applications* (1st ed., Vol. 90). Springer International Publishing. <https://doi.org/10.1007/978-3-030-74575-2>
- Maleki, F., Muthukrishnan, N., Ovens, K., Reinhold, C., & Forghani, R. (2020). Machine Learning Algorithm Validation: From Essentials to Advanced Applications and Implications for Regulatory Certification and Deployment. *Neuroimaging Clinics of North America*, 30(4), 433–445. <https://doi.org/10.1016/j.nic.2020.08.004>
- Manière, G., Alves, G., Berthelot-Grosjean, M., & Grosjean, Y. (2020). Growth regulation by amino acid transporters in *Drosophila* larvae. *Cellular and Molecular Life Sciences: CMLS*, 77(21), 4289–4297. <https://doi.org/10.1007/s00018-020-03535-6>
- Mannu, A., Blangetti, M., Baldino, S., & Prandi, C. (2021). Promising Technological and Industrial Applications of Deep Eutectic Systems. *Materials*, 14(10), Article 10. <https://doi.org/10.3390/ma14102494>
- Marín-García, P. J., & Llobat, L. (2021). How Does Protein Nutrition Affect the Epigenetic Changes in Pig? A Review. *Animals*, 11(2), Article 2. <https://doi.org/10.3390/ani11020544>
- Martins, M. A. R., Pinho, S. P., & Coutinho, J. A. P. (2019). Insights into the Nature of Eutectic and Deep Eutectic Mixtures. *Journal of Solution Chemistry*, 48(7), 962–982. <https://doi.org/10.1007/s10953-018-0793-1>
- Masella, R., & Mazza, G. (2009). *Glutathione and sulfur amino acids in human health and disease*. Wiley. <https://onlinelibrary.wiley.com/doi/book/10.1002/9780470475973>
- Matzkin, L. M., Johnson, S., Paight, C., Bozinovic, G., & Markow, T. A. (2011). Dietary Protein and Sugar Differentially Affect Development and Metabolic Pools in Ecologically Diverse *Drosophila*123. *The Journal of Nutrition*, 141(6), 1127–1133. <https://doi.org/10.3945/jn.111.138438>
- Maurange, C., & Lanet, E. (2014). Building a brain under nutritional restriction: Insights on sparing and plasticity from *Drosophila* studies. *Frontiers in Physiology*, 5. <https://doi.org/10.3389/fphys.2014.00117>
- Meganaharshini, M., Sudhakar, V., Dhivya Bharathi, N., & Deepak, S. (2023). Review on recent trends in the application of protein concentrates and isolates – A food industry perspective. *Food and Humanity*, 1, 308–325. <https://doi.org/10.1016/j.foohum.2023.05.022>

- Mejia, E. G. de, & Dia, V. P. (2009). Lunasin and lunasin-like peptides inhibit inflammation through suppression of NF-kappaB pathway in the macrophage. *Peptides (New York, N.Y. : 1980)*, 30(12), 2388–2398. <https://doi.org/10.1016/j.peptides.2009.08.005>
- Meng, A., Luan, B., Zhang, W., Zheng, Y., Guo, B., & Zhang, B. (2024). Exploring changes in aggregation and gel network morphology of soybean protein isolate induced by pH, NaCl, and temperature in view of interactions. *International Journal of Biological Macromolecules*, 273(Pt 1), 132911. <https://doi.org/10.1016/j.ijbiomac.2024.132911>
- Michailidis, J., Murray, N. D., & Graves, J. A. M. (1988). A correlation between development time and variegated position effect in *Drosophila melanogaster*. *Genetics Research*, 52(2), 119–123. <https://doi.org/10.1017/S0016672300027488>
- Militello, V., Vetri, V., & Leone, M. (2003). Conformational changes involved in thermal aggregation processes of bovine serum albumin. *Biophysical Chemistry*, 105(1), 133–141. [https://doi.org/10.1016/s0301-4622\(03\)00153-4](https://doi.org/10.1016/s0301-4622(03)00153-4)
- Min, S., Riaz, A., Qyyum, M. A., Choe, H., Moon, S., & Lee, M. (2022). Application of machine learning model to optimization of the hydrogen liquefaction process. In Y. Yamashita & M. Kano (Eds.), *Computer Aided Chemical Engineering* (Vol. 49, pp. 961–966). Elsevier. <https://doi.org/10.1016/B978-0-323-85159-6.50160-3>
- 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>
- Mohammed, E. A., Naugler, C., & Far, B. H. (2015). Chapter 32—Emerging Business Intelligence Framework for a Clinical Laboratory Through Big Data Analytics. In *Emerging Trends in Computational Biology, Bioinformatics, and Systems Biology* (Vol. 1—Book, Section, pp. 577–602). Elsevier Inc. <https://doi.org/10.1016/B978-0-12-802508-6.00032-6>
- Momen, S., Alavi, F., & Aider, M. (2021). Alkali-mediated treatments for extraction and functional modification of proteins: Critical and application review. *Trends in Food Science & Technology*, 110, 778–797. <https://doi.org/10.1016/j.tifs.2021.02.052>
- Mooney, C., Haslam, N. J., Pollastri, G., & Shields, D. C. (2012). Towards the improved discovery and Design of Functional Peptides: Common features of diverse classes permit generalized prediction of bioactivity. *PloS One*, 7(10), e45012. <https://doi.org/10.1371/journal.pone.0045012>
- Mulalapele, L. T., & Xi, J. (2021). Detection and inactivation of allergens in soybeans: A brief review of recent research advances. *Grain & Oil Science and Technology*, 4(4), 191–200. <https://doi.org/10.1016/j.gaost.2021.11.001>
- Naeem, M., Malik, M. I., Umar, T., Ashraf, S., & Ahmad, A. (2022). A Comprehensive Review About Bioactive Peptides: Sources to Future Perspective. *International Journal of Peptide Research and Therapeutics*, 28(6), 155. <https://doi.org/10.1007/s10989-022-10465-3>
- 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>
- Navaf, M., Sunooj, K. V., Aaliya, B., Sudheesh, C., Akhila, P. P., Mir, S. A., Nemtanu, M. R., George, J., Lackner, M., & Mousavi Khaneghah, A. (2023). Contemporary

- insights into the extraction, functional properties, and therapeutic applications of plant proteins. *Journal of Agriculture and Food Research*, 14, 100861. <https://doi.org/10.1016/j.jafr.2023.100861>
- Nazareth, Z. M., Deak, N. A., & Johnson, L. A. (2009). Functional Properties of Soy Protein Isolates Prepared from Gas-Supported Screw-Pressed Soybean Meal. *Journal of the American Oil Chemists' Society*, 86(4), 315–321. <https://doi.org/10.1007/s11746-009-1360-0>
- Neji, C., Semwal, J., Kamani, M. H., Máthé, E., & Sipos, P. (2022). Legume Protein Extracts: The Relevance of Physical Processing in the Context of Structural, Techno-Functional and Nutritional Aspects of Food Development. *Processes*, 10(12), Article 12. <https://doi.org/10.3390/pr10122586>
- Neumann, C., Harris, D. M., & Rogers, L. M. (2002). Contribution of animal source foods in improving diet quality and function in children in the developing world. *Nutrition Research*, 22(1), 193–220. [https://doi.org/10.1016/S0271-5317\(01\)00374-8](https://doi.org/10.1016/S0271-5317(01)00374-8)
- Nill, K. (2016). Soy Beans: Properties and Analysis. In *Encyclopedia of Food and Health* (pp. 54–55). Elsevier. <https://doi.org/10.1016/B978-0-12-384947-2.00642-5>
- Nisbet, R., Miner, G., & Yale, K. (2018). Chapter 10—Numerical Prediction. In R. Nisbet, G. Miner, & K. Yale (Eds.), *Handbook of Statistical Analysis and Data Mining Applications (Second Edition)* (pp. 187–213). Academic Press. <https://doi.org/10.1016/B978-0-12-416632-5.00010-4>
- Nooshkam, M., & Madadlou, A. (2016). Microwave-assisted isomerisation of lactose to lactulose and Maillard conjugation of lactulose and lactose with whey proteins and peptides. *Food Chemistry*, 200(Journal Article), 1–9. <https://doi.org/10.1016/j.foodchem.2015.12.094>
- Ojeda-Avila, T., Woods, H. A., & Raguso, R. A. (2003). Effects of dietary variation on growth, composition, and maturation of *Manduca sexta* (Sphingidae: Lepidoptera). *Journal of Insect Physiology*, 49(4), 293–306. [https://doi.org/10.1016/s0022-1910\(03\)00003-9](https://doi.org/10.1016/s0022-1910(03)00003-9)
- Oliveira, A. C., & Homem, C. C. F. (2023). Opposing effects of ecdysone signaling regulate neuroblast proliferation to ensure coordination of brain and organism development. *Developmental Biology*, 503, 53–67. <https://doi.org/10.1016/j.ydbio.2023.08.001>
- Omar, K. A., & Sadeghi, R. (2022). Physicochemical properties of deep eutectic solvents: A review. *Journal of Molecular Liquids*, 360, 119524. <https://doi.org/10.1016/j.molliq.2022.119524>
- Ormerod, K. G., LePine, O. K., Abbineni, P. S., Bridgeman, J. M., Coorssen, J. R., Mercier, A. J., & Tattersall, G. J. (2017). *Drosophila* development, physiology, behavior, and lifespan are influenced by altered dietary composition. *Fly*, 11(3), 153–170. <https://doi.org/10.1080/19336934.2017.1304331>
- Pachouly, J., Ahirrao, S., Kotecha, K., Selvachandran, G., & Abraham, A. (2022). A systematic literature review on software defect prediction using artificial intelligence: Datasets, Data Validation Methods, Approaches, and Tools. *Engineering Applications of Artificial Intelligence*, 111, 104773. <https://doi.org/10.1016/j.engappai.2022.104773>
- Pan, X., Connacher, R. P., & O'Connor, M. B. (2021). Control of the insect metamorphic transition by ecdysteroid production and secretion. *Current Opinion in Insect Science*, 43, 11–20. <https://doi.org/10.1016/j.cois.2020.09.004>
- Panato, A., Antonini, E., Bortolotti, F., & Ninfali, P. (2017). The histology of grain caryopses for nutrient location: A comparative study of six cereals. *International*

- Journal of Food Science and Technology*, 52(5), 1238–1245. <https://doi.org/10.1111/ijfs.13390>
- Pant, J., Pant, R. P., Singh, Manoj Kumar, Singh, D. P., & Pant, H. (2021). Analysis of agricultural crop yield prediction using statistical techniques of machine learning. *Materials Today: Proceedings*, 46(Journal Article), 10922–10926. <https://doi.org/10.1016/j.matpr.2021.01.948>
- Parikh, N., Yilanli, M., & Saadabadi, A. (2025a). Tranylcypromine. In *StatPearls*. StatPearls Publishing. <http://www.ncbi.nlm.nih.gov/books/NBK459162/>
- Parikh, N., Yilanli, M., & Saadabadi, A. (2025b). Tranylcypromine. In *StatPearls*. StatPearls Publishing. <http://www.ncbi.nlm.nih.gov/books/NBK459162/>
- Passos, H., Tavares, D. J. P., Ferreira, A. M., Freire, M. G., & Coutinho, J. A. P. (2016). Are Aqueous Biphasic Systems Composed of Deep Eutectic Solvents Ternary or Quaternary Systems? *ACS Sustainable Chemistry & Engineering*, 4(5), 2881–2886. <https://doi.org/10.1021/acssuschemeng.6b00485>
- Patra, A., Arun Prasath, V., & 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>
- Penfield, M. P., & Campbell, A. M. (1990). CHAPTER 15—FATS AND THEIR LIPID CONSTITUENTS. In M. P. Penfield & A. M. Campbell (Eds.), *Experimental Food Science (Third Edition)* (pp. 331–357). Academic Press. <https://doi.org/10.1016/B978-0-12-157920-3.50019-3>
- 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>
- Peng, Y., Kersten, N., Kyriakopoulou, K., & van der Goot, A. J. (2020). Functional properties of mildly fractionated soy protein as influenced by the processing pH. *Journal of Food Engineering*, 275(Journal Article), 109875. <https://doi.org/10.1016/j.jfoodeng.2019.109875>
- Pereira, R. N., Rodrigues, R. M., Machado, L., Ferreira, S., Costa, J., Villa, C., Barreiros, M. P., Mafra, I., Teixeira, J. A., & Vicente, A. A. (2021). Influence of ohmic heating on the structural and immunoreactive properties of soybean proteins. *Food Science & Technology*, 148(Journal Article), 111710. <https://doi.org/10.1016/j.lwt.2021.111710>
- Pi, X., Sun, Y., Fu, G., Wu, Z., & Cheng, J. (2021). Effect of processing on soybean allergens and their allergenicity. *Trends in Food Science & Technology*, 118, 316–327. <https://doi.org/10.1016/j.tifs.2021.10.006>
- Prabu, E., Felix, N., Uma, A., & Praveenraj, J. (2020). Effects of dietary L-lysine supplementation on growth, body composition and muscle-growth-related gene expression with an estimation of lysine requirement of GIFT tilapia. *Aquaculture Nutrition*, 26(2), 568–578. <https://doi.org/10.1111/anu.13018>
- Prandi, B., Faccini, A., Lambertini, F., Bencivenni, M., Jorba, M., Van Droogenbroeck, 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>
- Prasad, S., Mandal, I., Singh, S., Paul, A., Mandal, B., Venkatramani, R., & Swaminathan, R. (2017). Near UV-Visible electronic absorption originating from charged amino acids in a monomeric protein. *Chemical Science*, 8(8), 5416–5433. <https://doi.org/10.1039/C7SC00880E>

- Qin, P., Wang, T., & Luo, Y. (2022). A review on plant-based proteins from soybean: Health benefits and soy product development. *Journal of Agriculture and Food Research*, 7, 100265. <https://doi.org/10.1016/j.jafr.2021.100265>
- Rabbani, N., Kim, G. Y. E., Suarez, C. J., & Chen, J. H. (2022). Applications of machine learning in routine laboratory medicine: Current state and future directions. *Clinical Biochemistry*, 103(Journal Article), 1–7. <https://doi.org/10.1016/j.clinbiochem.2022.02.011>
- Radotić, K., Stanković, M., Bartolić, D., & Natić, M. (2023). Intrinsic Fluorescence Markers for Food Characteristics, Shelf Life, and Safety Estimation: Advanced Analytical Approach. *Foods*, 12(16), Article 16. <https://doi.org/10.3390/foods12163023>
- Rahman, M. M., Byanju, B., Grewell, D., & Lamsal, B. P. (2020). High-power sonication of soy proteins: Hydroxyl radicals and their effects on protein structure. *Ultrasonics Sonochemistry*, 64(Journal Article), 105019. <https://doi.org/10.1016/j.ultsonch.2020.105019>
- Rahman, M. M., & Lamsal, B. P. (2021). Ultrasound-assisted extraction and modification of plant-based proteins: Impact on physicochemical, functional, and nutritional properties. *Comprehensive Reviews in Food Science and Food Safety*, 20(2), 1457–1480. <https://doi.org/10.1111/1541-4337.12709>
- Rakesh, K. P., Suhas, R., & Gowda, D. C. (2019). Anti-inflammatory and Antioxidant Peptide-Conjugates: Modulation of Activity by Charged and Hydrophobic Residues. *International Journal of Peptide Research and Therapeutics*, 25(1), 227–234. <https://doi.org/10.1007/s10989-017-9668-3>
- Ramlee, K. A. F. K., Daud, N. M. A. N., Sarip, M. S. M., Bakar, A. R. A., & Zainudin, M. A. M. (2024). A Three-Level Factorial Model for Maximising Protein Extraction from Rice Bran with Choline Chloride: Glycerol. *Journal of Advanced Research in Micro and Nano Engineering*, 19(1), Article 1. <https://doi.org/10.37934/armne.19.1.6377>
- Randel, N. (2025). *Effects of food availability on larval development during ontogenetic niche shift in a marine annelid* (p. 2025.04.20.649073). bioRxiv. <https://doi.org/10.1101/2025.04.20.649073>
- Reese, G., Viebranz, J., Leong-Kee, S. M., Plante, M., Lauer, I., Randow, S., Moncin, M. S.-M., Ayuso, R., Lehrer, S. B., & Vieths, S. (2005). Reduced allergenic potency of VR9-1, a mutant of the major shrimp allergen Pen a 1 (tropomyosin). *Journal of Immunology (Baltimore, Md.: 1950)*, 175(12), 8354–8364. <https://doi.org/10.4049/jimmunol.175.12.8354>
- Rehman, S. U., Nadeem, A., Javed, M., Hassan, F.-U., Luo, X., Khalid, R. B., & Liu, Q. (2020). Genomic Identification, Evolution and Sequence Analysis of the Heat-Shock Protein Gene Family in Buffalo. *Genes*, 11(11), 1388. <https://doi.org/10.3390/genes11111388>
- Ren, W., Li, Y., Yin, Y., & Blachier, F. (2013). *Structure, Metabolism and Functions of Amino Acids: An Overview* (pp. 91–108). Springer Vienna. https://doi.org/10.1007/978-3-7091-1328-8_7
- Rijavec, T., & Zupin, Z. (2015). *Textiles from new soybean protein fibers (SPF)*. <https://doi.org/10.13140/RG.2.1.5169.1041>
- Rivera-Jiménez, J., Berraquero-García, C., Pérez-Gálvez, R., García-Moreno, P. J., Espejo-Carpio, F. J., Guadix, A., & Guadix, E. M. (2022). Peptides and protein hydrolysates exhibiting anti-inflammatory activity: Sources, structural features and modulation mechanisms. *Food & Function*, 13(24), 1251–1254. <https://doi.org/10.1039/d2fo02223k>

- Roberts, P. R., Burney, J. D., Black, K. W., & Zaloga, G. P. (1999). Effect of Chain Length on Absorption of Biologically Active Peptides from the Gastrointestinal Tract. *Digestion*, 60(4), 332–337. <https://doi.org/10.1159/000007679>
- Rodrigues, I. M., Coelho, J. F. J., & Carvalho, M. G. V. S. (2012). Isolation and valorisation of vegetable proteins from oilseed plants: Methods, limitations and potential. *Journal of Food Engineering*, 109(3), 337–346. <https://doi.org/10.1016/j.jfoodeng.2011.10.027>
- Rodrigues, M. A., Martins, N. E., Balancé, L. F., Broom, L. N., Dias, A. J. S., Fernandes, A. S. D., Rodrigues, F., Sucena, É., & Mirth, C. K. (2015). *Drosophila melanogaster* larvae make nutritional choices that minimize developmental time. *Journal of Insect Physiology*, 81, 69–80. <https://doi.org/10.1016/j.jinsphys.2015.07.002>
- Rovenko, B. M., Perkhulyn, N. V., Gospodaryov, D. V., Sanz, A., Lushchak, O. V., & Lushchak, V. I. (2015). High consumption of fructose rather than glucose promotes a diet-induced obese phenotype in *Drosophila melanogaster*. *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology*, 180, 75–85. <https://doi.org/10.1016/j.cbpa.2014.11.008>
- Rubinstein, N. D., Mayrose, I., Halperin, D., Yekutieli, D., Gershoni, J. M., & Pupko, T. (2008). Computational characterization of B-cell epitopes. *Molecular Immunology*, 45(12), 3477–3489. <https://doi.org/10.1016/j.molimm.2007.10.016>
- Ruggeri, F. S., Šneideris, T., Vendruscolo, M., & Knowles, T. P. J. (2019). Atomic force microscopy for single molecule characterisation of protein aggregation. *Archives of Biochemistry and Biophysics*, 664, 134–148. <https://doi.org/10.1016/j.abb.2019.02.001>
- Saito, Y., Itakura, K., Kuramoto, M., Kaho, T., Ohtake, N., Hasegawa, H., Suzuki, T., & Kondo, N. (2021). Prediction of protein and oil contents in soybeans using fluorescence excitation emission matrix. *Food Chemistry*, 365, 130403. <https://doi.org/10.1016/j.foodchem.2021.130403>
- Sanchez-Fernandez, A., Basic, M., Xiang, J., Prevost, S., Jackson, A. J., & Dicko, C. (2022a). Hydration in Deep Eutectic Solvents Induces Non-monotonic Changes in the Conformation and Stability of Proteins. *Journal of the American Chemical Society*, 144(51), 23657–23667. <https://doi.org/10.1021/jacs.2c11190>
- Sanchez-Fernandez, A., Prevost, S., & Wahlgren, M. (2022b). Deep eutectic solvents for the preservation of concentrated proteins: The case of lysozyme in 1 : 2 choline chloride : glycerol. *Green Chemistry*, 24(11), 4437–4442. <https://doi.org/10.1039/D1GC04378A>
- Santos, L. B., Assis, R. S., Barreto, J. A., Bezerra, M. A., Novaes, C. G., & Lemos, V. A. (2022). Deep eutectic solvents in liquid-phase microextraction: Contribution to green chemistry. *TrAC Trends in Analytical Chemistry*, 146, 116478. <https://doi.org/10.1016/j.trac.2021.116478>
- Sapir, L., & Harries, D. (2020). Restructuring a Deep Eutectic Solvent by Water: The Nanostructure of Hydrated Choline Chloride/Urea. *Journal of Chemical Theory and Computation*, 16(5), 3335–3342. <https://doi.org/10.1021/acs.jctc.0c00120>
- Sarker, I. H. (2021). Machine Learning: Algorithms, Real-World Applications and Research Directions. *SN Computer Science*, 2(3), 160–160. <https://doi.org/10.1007/s42979-021-00592-x>
- Schön, A., Clarkson, B. R., Jaime, M., & Freire, E. (2017). Temperature stability of proteins: Analysis of irreversible denaturation using isothermal calorimetry. *Proteins, Structure, Function, and Bioinformatics*, 85(11), 2009–2016. <https://doi.org/10.1002/prot.25354>

- Schreiber, J. B. (2008). Core reporting practices in structural equation modeling. *Research in Social and Administrative Pharmacy*, 4(2), 83–97. <https://doi.org/10.1016/j.sapharm.2007.04.003>
- 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>
- Senthil, S. K., Murugan, P. K., Selvam, S., Chandhana, J. P., Satheesh Babu, T. G., & Kandasamy, E. (2020). Fluorescence spectroscopic analysis of heavy metal induced protein denaturation. *Materials Today : Proceedings*, 33(Journal Article), 2328–2330. <https://doi.org/10.1016/j.matpr.2020.04.693>
- Seraj, A., Mohammadi-Khanaposhtani, M., Daneshfar, R., Naseri, M., Esmaeili, M., Baghban, A., Habibzadeh, S., & Eslamian, S. (2023). Chapter 5—Cross-validation. In S. Eslamian & F. Eslamian (Eds.), *Handbook of Hydroinformatics* (Vol. 1—Book, Section, pp. 89–105). Elsevier. <https://doi.org/10.1016/B978-0-12-821285-1.00021-X>
- Shaik, M. I., & Sarbon, N. M. (2022). A Review on Purification and Characterization of Anti-proliferative Peptides Derived from Fish Protein Hydrolysate. *Food Reviews International*, 38(7), 1389–1409. <https://doi.org/10.1080/87559129.2020.1812634>
- Shawky, E., Zhu, W., & Tian, J. (2025). A review of innovative extraction technologies for protein recovery from plant-based by-products: A step toward zero-waste processing. *International Journal of Biological Macromolecules*, 315, 144301. <https://doi.org/10.1016/j.ijbiomac.2025.144301>
- Shen, G., Zheng, L., Li, S., Wu, H., Li, M., Luo, Q., Yu, G., Chen, A., & Zhang, Z. (2020). The role of soy protein degradation caused by spoilage *Bacillus amyloliquefaciens* in texture deterioration of yuba, a soy product. *LWT*, 123, 109108. <https://doi.org/10.1016/j.lwt.2020.109108>
- Shen, P., Zhao, M., & Zhou, F. (2022). Design of soy protein/peptide-based colloidal particles and their role in controlling the lipid digestion of emulsions. *Current Opinion in Food Science*, 43(Journal Article), 61–70. <https://doi.org/10.1016/j.cofs.2021.10.003>
- 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>
- Shmueli, G. (2010). To Explain or to Predict? *Statistical Science*, 25(3), 289–310. <https://doi.org/10.1214/10-STS330>
- Sieradzki, Z., Mazur, M., Król, B., & Kwiatek, K. (2021). Prevalence of Genetically Modified Soybean in Animal Feedingstuffs in Poland. *Journal of Veterinary Research*, 65(1), 93–99. <https://doi.org/10.2478/jvetres-2021-0012>
- Soglia, F., Zampiga, M., Baldi, G., Malila, Y., Thanatsang, K. V., Srimarut, Y., Tatiyaborworntham, N., Unger, O., Klamchuen, A., Laghi, L., Petracci, M., & Sirri, F. (2021). Lysine Depletion during Different Feeding Phases: Effects on Growth Performances and Meat Quality of Broiler Chickens. *Animals*, 11(6), Article 6. <https://doi.org/10.3390/ani11061499>
- Soriano-Correa, C., Barrientos-Salcedo, C., Campos-Fernández, L., Alvarado-Salazar, A., & Esquivel, R. O. (2015). Importance of asparagine on the conformational stability and chemical reactivity of selected anti-inflammatory peptides. *Chemical Physics*, 457, 180–187. <https://doi.org/10.1016/j.chemphys.2015.06.005>

- Sousa-Nunes, R., Yee, L. L., & Gould, A. P. (2011). Fat cells reactivate quiescent neuroblasts via TOR and glial Insulin relays in *Drosophila*. *Nature*, 471(7339), 508–512. <https://doi.org/10.1038/nature09867>
- Souza, G. de A. L. e, Pietro, M. E. D., & Mele, A. (2024). Eutectic solvents and low molecular weight gelators for next-generation supramolecular eutectogels: A sustainable chemistry perspective. *RSC Sustainability*, 2(2), 288–319. <https://doi.org/10.1039/D3SU00264K>
- Sprangers, T., Moradei, A., Vynckier, K., Boudrez, M., Pinotti, L., & Ottoboni, M. (2024). Amino acid requirements of yellow mealworm and black soldier fly larvae. *Journal of Insects as Food and Feed*, 1(aop), 1–12. <https://doi.org/10.1163/23524588-00001271>
- Srisuradetchai, P., & Suksrikan, K. (2024). Random kernel k-nearest neighbors regression. *Frontiers in Big Data*, 7. <https://doi.org/10.3389/fdata.2024.1402384>
- Steck, K., Walker, S. J., Itskov, P. M., Baltazar, C., Moreira, J.-M., & Ribeiro, C. (2018). Internal amino acid state modulates yeast taste neurons to support protein homeostasis in *Drosophila*. *eLife*, 7, e31625. <https://doi.org/10.7554/eLife.31625>
- Stieper, B. C., Kupershtok, M., Driscoll, M. V., & Shingleton, A. W. (2008). Imaginal discs regulate developmental timing in *Drosophila melanogaster*. *Developmental Biology*, 321(1), 18–26. <https://doi.org/10.1016/j.ydbio.2008.05.556>
- Sun, C., Dai, L., Liu, F., & Gao, Y. (2016). Simultaneous treatment of heat and high pressure homogenization of zein in ethanol–water solution: Physical, structural, thermal and morphological characteristics. *Innovative Food Science & Emerging Technologies*, 34(Journal Article), 161–170. <https://doi.org/10.1016/j.ifset.2016.01.016>
- 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>
- Sun, L., Wang, D., Huang, Z., Elfalleh, W., Qin, L., & Yu, D. (2023). Structure and flavor characteristics of Maillard reaction products derived from soybean meal hydrolysates-reducing sugars. *LWT*, 185, 115097. <https://doi.org/10.1016/j.lwt.2023.115097>
- Tando, Y., & Matsui, Y. (2023). Inheritance of environment-induced phenotypic changes through epigenetic mechanisms. *Environmental Epigenetics*, 9(1), dvad008. <https://doi.org/10.1093/eep/dvad008>
- Tang, J., Yao, D., Xia, S., Cheong, L., & Tu, M. (2024). Recent progress in plant-based proteins: From extraction and modification methods to applications in the food industry. *Food Chemistry: X*, 23, 101540. <https://doi.org/10.1016/j.fochx.2024.101540>
- Tatum, E. L., & Beadle, G. W. (1939). Effect of Diet on Eye-Color Development in *Drosophila melanogaster*. *Biological Bulletin*, 77(3), 415–422. <https://doi.org/10.2307/1537651>
- Taylor, S. L., Marsh, J. T., Koppelman, S. J., Kabourek, J. L., Johnson, P. E., & Baumert, J. L. (2021). A perspective on pea allergy and pea allergens. *Trends in Food Science & Technology*, 116, 186–198. <https://doi.org/10.1016/j.tifs.2021.07.017>
- Tenhunen, T.-M., Lewandowska, A. E., Orelma, H., Johansson, L.-S., Virtanen, T., Harlin, A., Österberg, M., Eichhorn, S. J., & Tammelin, T. (2018). Understanding the interactions of cellulose fibres and deep eutectic solvent of choline chloride and urea. *Cellulose*, 25(1), 137–150. <https://doi.org/10.1007/s10570-017-1587-0>

- 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>
- Tong, S. C., Siow, L. F., Tang, T. K., & Lee, Y. Y. (2025). Effect of Enzyme-Assisted Extraction on Structural and Functional Properties of Palm Kernel Protein. *Journal of Food Biochemistry*, 2025(1), 8869943. <https://doi.org/10.1155/jfbc/8869943>
- Tran, T.-A.-T., Kim, Y.-H., Kim, G.-E., Jung, S., Kim, I.-Y., Moon, K.-S., Kim, Y.-J., Lee, T.-K., Yun, H., Lee, J.-J., Lee, H.-J., Lee, C. W., & Jung, T.-Y. (2022). The long multi-epitope peptide vaccine combined with adjuvants improved the therapeutic effects in a glioblastoma mouse model. *Frontiers in Immunology*, 13, 1007285. <https://doi.org/10.3389/fimmu.2022.1007285>
- Tseng, Y. J., Chuang, P.-J., & Appell, M. (2023). When Machine Learning and Deep Learning Come to the Big Data in Food Chemistry. *ACS Omega*, 8(18), 15854–15864. <https://doi.org/10.1021/acsomega.2c07722>
- Tu, Z., Young, A., Murphy, C., & Liang, J. F. (2009). The pH sensitivity of histidine-containing lytic peptides. *Journal of Peptide Science*, 15(11), 790–795. <https://doi.org/10.1002/psc.1180>
- Uchegbu, E. C., & Kopelman, P. G. (2013). Obesity: Treatment. In B. Caballero (Ed.), *Encyclopedia of Human Nutrition (Third Edition)* (pp. 374–382). Academic Press. <https://doi.org/10.1016/B978-0-12-375083-9.00214-2>
- Van Alphen, B., & Van Swinderen, B. (2013). *Drosophila* strategies to study psychiatric disorders. *Brain Research Bulletin*, 92, 1–11. <https://doi.org/10.1016/j.brainresbull.2011.09.007>
- Veza, I., Spraggon, M., Fattah, I. M. R., & Idris, M. (2023). Response surface methodology (RSM) for optimizing engine performance and emissions fueled with biofuel: Review of RSM for sustainability energy transition. *Results in Engineering*, 18, 101213. <https://doi.org/10.1016/j.rineng.2023.101213>
- Vijendravarma, R. K., Narasimha, S., & Kawecki, T. J. (2013). Predatory cannibalism in *Drosophila melanogaster* larvae. *Nature Communications*, 4(1), 1789. <https://doi.org/10.1038/ncomms2744>
- Vilková, M., Płotka-Wasyłka, J., & Andruch, V. (2020). The role of water in deep eutectic solvent-base extraction. *Journal of Molecular Liquids*, 304, 112747. <https://doi.org/10.1016/j.molliq.2020.112747>
- Vitale, M., Costabile, G., Testa, R., D’Abbronzio, G., Nettore, I. C., Macchia, P. E., & Giacco, R. (2024). Ultra-Processed Foods and Human Health: A Systematic Review and Meta-Analysis of Prospective Cohort Studies. *Advances in Nutrition*, 15(1), 100121. <https://doi.org/10.1016/j.advnut.2023.09.009>
- Vivian, J. T., & Callis, P. R. (2001). Mechanisms of Tryptophan Fluorescence Shifts in Proteins. *Biophysical Journal*, 80(5), 2093–2109. [https://doi.org/10.1016/S0006-3495\(01\)76183-8](https://doi.org/10.1016/S0006-3495(01)76183-8)
- Vo, T.-S., Ryu, B., & Kim, S.-K. (2013). Purification of novel anti-inflammatory peptides from enzymatic hydrolysate of the edible microalgal *Spirulina maxima*. *Journal of Functional Foods*, 5(3), 1336–1346. <https://doi.org/10.1016/j.jff.2013.05.001>
- Vu, J. P., Luong, L., Parsons, W. F., Oh, S., Sanford, D., Gabalski, A., Lighton, J. R., Pisegna, J. R., & Germano, P. M. (2017). Long-Term Intake of a High-Protein Diet Affects Body Phenotype, Metabolism, and Plasma Hormones in Mice. *The Journal of Nutrition*, 147(12), 2243–2251. <https://doi.org/10.3945/jn.117.257873>
- Wahlström, R., Rommi, K., Willberg-Keyriläinen, P., Ercili-Cura, D., Holopainen-Mantila, U., Hiltunen, J., Mäkinen, O., Nygren, H., Mikkelsen, A., & Kuutti, L.

- (2017). High Yield Protein Extraction from Brewer's Spent Grain with Novel Carboxylate Salt—Urea Aqueous Deep Eutectic Solvents. *ChemistrySelect*, 2(29), 9355–9363. <https://doi.org/10.1002/slct.201701492>
- Wan, X., Liu, H., Sun, Y., Zhang, J., Chen, X., & Chen, N. (2017). Lunasin: A promising polypeptide for the prevention and treatment of cancer. *Oncology Letters*, 13(6), 3997–4001. <https://doi.org/10.3892/ol.2017.6017>
- Wang, C., Dong, S., Zhang, L., Zhao, Y., Huang, L., Gong, X., Wang, H., & Shang, D. (2017). Cell surface binding, uptaking and anticancer activity of L-K6, a lysine/leucine-rich peptide, on human breast cancer MCF-7 cells. *Scientific Reports*, 7(1), 8293–13. <https://doi.org/10.1038/s41598-017-08963-2>
- Wang, H., Wang, N., Chen, X., Wu, Z., Zhong, W., Yu, D., & Zhang, H. (2022). Effects of moderate electric field on the structural properties and aggregation characteristics of soybean protein isolate. *Food Hydrocolloids*, 133, 107911. <https://doi.org/10.1016/j.foodhyd.2022.107911>
- Wang, L., Wei, D.-D., Wang, G.-Q., Huang, H.-Q., & Wang, J.-J. (2023). High-Sucrose Diet Exposure on Larvae Contributes to Adult Fecundity and Insecticide Tolerance in the Oriental Fruit Fly, *Bactrocera dorsalis* (Hendel). *Insects*, 14(5), 407. <https://doi.org/10.3390/insects14050407>
- Wang, L., Yang, J., Wang, Y., Zhang, J., Gao, Y., Yuan, J., Su, A., & Ju, X. (2016). Study on Antioxidant Activity and Amino Acid Analysis of Rapeseed Protein Hydrolysates. *International Journal of Food Properties*, 19(9), 1899–1911. <https://doi.org/10.1080/10942912.2015.1085397>
- Wang, X., Amei, A., De Belle, J. S., & Roberts, S. P. (2017). Environmental effects on *Drosophila* brain development and learning. *Journal of Experimental Biology*, jeb.169375. <https://doi.org/10.1242/jeb.169375>
- Wang, X., Xie, H., Hu, Y., & Tu, Z. (2024). Investigation on potential allergenic peptides and critical amino acids of the digestive products of glycated α -lactalbumin via allergenicity evaluation and molecular dynamic simulation. *LWT*, 195, 115800. <https://doi.org/10.1016/j.lwt.2024.115800>
- Wang, X., Yu, M., Wang, Z., Luo, K., Adhikari, B., Miao, S., & Liu, S. (2022). Modulation of Soy Protein Isolate Gel Properties by a Novel “Two-Step” Gelation Process: Effects of Pre-aggregation with Different Divalent Sulfates. *Food Chemistry*, 394(Journal Article), 133515. <https://doi.org/10.1016/j.foodchem.2022.133515>
- Wang, Y., Cui, Z., & Ke, R. (2023). Chapter 3—Machine learning basics. In Y. Wang, Z. Cui, & R. Ke (Eds.), *Machine Learning for Transportation Research and Applications* (Vol. 1—Book, Section, pp. 25–40). Elsevier. <https://doi.org/10.1016/B978-0-32-396126-4.00008-4>
- 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>
- Wang, Z.-L., Wang, X.-P., Li, C.-R., Xia, Z.-Z., & Li, S.-X. (2018). Effect of Dietary Protein and Carbohydrates on Survival and Growth in Larvae of the *Henosepilachna vigintioctopunctata* (F.) (Coleoptera: Coccinellidae). *Journal of Insect Science (Online)*, 18(4), 3. <https://doi.org/10.1093/jisesa/iey067>
- Waterland, R. A., & Jirtle, R. L. (2003). Transposable elements: Targets for early nutritional effects on epigenetic gene regulation. *Molecular and Cellular Biology*, 23(15), 5293–5300. <https://doi.org/10.1128/MCB.23.15.5293-5300.2003>

- Wiederrecht, G. J., Paton, D. R., & Brown, G. M. (1981). The isolation and identification of an intermediate involved in the biosynthesis of drosopterin in *Drosophila melanogaster*. *Journal of Biological Chemistry*, 256(20), 10399–10402. [https://doi.org/10.1016/S0021-9258\(19\)68632-1](https://doi.org/10.1016/S0021-9258(19)68632-1)
- Wiederstein, M., Baumgartner, S., & Lauter, K. (2023). Soybean (*Glycine max*) allergens—A Review on an Outstanding Plant Food with Allergenic Potential. *ACS Food Science & Technology*, 3(3), 363–378. <https://doi.org/10.1021/acsfoodscitech.2c00380>
- Wu, J. (2022). *Emerging sources and applications of alternative proteins* (1st edition, Vol. 101). Academic Press.
- Wu, X., Gao, T., Xu, Z., Liu, C., Teng, F., & Li, Y. (2024). Effect of combined enzyme and ultrasound treatment on the structure and gel properties of soy protein isolate: A comparative study of alkaline protease and pepsin. *Colloids and Surfaces. A, Physicochemical and Engineering Aspects*, 687(Journal Article), 133533. <https://doi.org/10.1016/j.colsurfa.2024.133533>
- Xia, W., Pan, S., Cheng, Z., Tian, Y., & Huang, X. (2020). High-intensity ultrasound treatment on soy protein after selectively proteolyzing glycinin component: Physical, structural, and aggregation properties. *Foods*, 9(6), 839. <https://doi.org/10.3390/foods9060839>
- 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>
- Xiaoya, Z., Mengfei, Zhang, Changbao, Hu, Li, Xin, Yang, Anshu, Tong, Ping, Wu, Zhihua, & and Chen, H. (2024). Effect of critical amino acids' properties on potential allergenicity of Ara h 2 epitopes. *Food and Agricultural Immunology*, 35(1), 2373064. <https://doi.org/10.1080/09540105.2024.2373064>
- Xu, J., Liu, X.-F., & Wang, Y.-T. (2016). A detection method of vegetable oils in edible blended oil based on three-dimensional fluorescence spectroscopy technique. *Food Chemistry*, 212, 72–77. <https://doi.org/10.1016/j.foodchem.2016.05.158>
- Xu, Z., Tan, Y., Ma, X., Wu, S., Zhang, B., & Luque, R. (2023). A novel method to remove nitrogen from sewage sludge during hydrothermal carbonization via inhibiting Maillard reaction. *Journal of Environmental Chemical Engineering*, 11(2), 109342. <https://doi.org/10.1016/j.jece.2023.109342>
- Yamine, A., Gao, J., & Kwan, A. H. (2019). Tryptophan Fluorescence Quenching Assays for Measuring Protein-ligand Binding Affinities: Principles and a Practical Guide. *Bio-Protocol*, 9(11), e3253. <https://doi.org/10.21769/BioProtoc.3253>
- Yang, Q., Cai, X., Huang, M., & Wang, S. (2020). A specific peptide with immunomodulatory activity from *Pseudostellaria heterophylla* and the action mechanism. *Journal of Functional Foods*, 68(C), 103887. <https://doi.org/10.1016/j.jff.2020.103887>
- Yang, Z. R., & Yang, Z. (2014). 6.01—Artificial Neural Networks. In A. Brahme (Ed.), *Comprehensive Biomedical Physics* (Vol. 1—Book, Section, pp. 1–17). Elsevier. <https://doi.org/10.1016/B978-0-444-53632-7.01101-1>
- Yeong, S. K., Idris, Z., & Hassan, H. A. (2012). 20—Palm Oleochemicals in Non-food Applications. In O.-M. Lai, C.-P. Tan, & C. C. Akoh (Eds.), *Palm Oil* (pp. 587–624). AOCS Press. <https://doi.org/10.1016/B978-0-9818936-9-3.50023-X>
- Yim, J., Kim, S., Walcher, G., & Pfeleiderer, W. (1993). Pteridines. Part C. Structure and nonenzymatic synthesis of aurodrosopterin. *Helvetica Chimica Acta*, 76(5), 1970–1979. <https://doi.org/10.1002/hlca.19930760516>

- 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>
- Zaky, A. A., Simal-Gandara, J., Eun, J.-B., Shim, J.-H., & El-Aty, A. M. A. (2022). Bioactivities, Applications, Safety, and Health Benefits of Bioactive Peptides From Food and By-Products: A Review. *Frontiers in Nutrition*, 8(8), 815640. <https://doi.org/10.3389/fnut.2021.815640>
- 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>
- Zhang, S., Sun, L., Ju, H., Bao, Z., Zeng, X., & Lin, S. (2021). Research advances and application of pulsed electric field on proteins and peptides in food. *Food Research International*, 139, 109914. <https://doi.org/10.1016/j.foodres.2020.109914>
- Zhang, T., Guo, J., Chen, J.-F., Wang, J.-M., Wan, Z.-L., & Yang, X.-Q. (2020). Heat stability and rheological properties of concentrated soy protein/egg white protein composite microparticle dispersions. *Food Hydrocolloids*, 100(Journal Article), 105449. <https://doi.org/10.1016/j.foodhyd.2019.105449>
- Zhang, T.-F., Yang, J.-F., & Lin, D. K. J. (2011). Small Box–Behnken design. *Statistics & Probability Letters*, 81(8), 1027–1033. <https://doi.org/10.1016/j.spl.2011.02.024>
- Zhang, X., Zhang, Z., Shen, A., Zhang, T., Jiang, L., El-Seedi, H., Zhang, G., & Sui, X. (2024). Legumes as an alternative protein source in plant-based foods: Applications, challenges, and strategies. *Current Research in Food Science*, 9, 100876. <https://doi.org/10.1016/j.crf.2024.100876>
- Zhang, Y., Wolosker, M. B., Zhao, Y., Ren, H., & Lemos, B. (2020). Exposure to microplastics cause gut damage, locomotor dysfunction, epigenetic silencing, and aggravate cadmium (Cd) toxicity in *Drosophila*. *Science of The Total Environment*, 744, 140979. <https://doi.org/10.1016/j.scitotenv.2020.140979>
- Zhao, S., Pan, F., Cai, S., Yi, J., Zhou, L., & Liu, Z. (2023). Secrets behind Protein Sequences: Unveiling the Potential Reasons for Varying Allergenicity Caused by Caseins from Cows, Goats, Camels, and Mares Based on Bioinformatics Analyses. *International Journal of Molecular Sciences*, 24(3), 2481. <https://doi.org/10.3390/ijms24032481>
- Zhong, W., Zhang, T., Dong, C., Li, J., Dai, J., & Wang, C. (2022). Effect of sodium chloride on formation and structure of whey protein isolate/hyaluronic acid complex and its ability to loading curcumin. *Colloids and Surfaces A: Physicochemical and Engineering Aspects*, 632, 127828. <https://doi.org/10.1016/j.colsurfa.2021.127828>
- Zhou, F., He, S., Sun, H., Wang, Y., & Zhang, Y. (2021). Advances in epitope mapping technologies for food protein allergens: A review. *Trends in Food Science & Technology*, 107, 226–239. <https://doi.org/10.1016/j.tifs.2020.10.035>
- Zhou, X., Ren, L., Zhang, Y., Zhang, J., Li, X., Yang, A., Tong, P., Wu, Z., & Chen, H. (2023). Effect of Structural Targeted Modifications on the Potential Allergenicity of Peanut Allergen Ara h 2. *Journal of Agricultural and Food Chemistry*, 71(1), 836–845. <https://doi.org/10.1021/acs.jafc.2c06359>
- Zhu, Y. F., Yao, Y., Huang, Y., Chen, C. H., Zhang, H. Y., & Huang, Z. (2022). Machine learning applications for assessment of dynamic progressive collapse of steel

moment frames. *Structures (Oxford)*, 36(Journal Article), 927–934.
<https://doi.org/10.1016/j.istruc.2021.12.067>

10. 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:
Subject:

DEENK/2/2026.PL
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





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 Törös, 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 Törös, 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.





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



11. STATEMENTS

STATEMENT

I wrote this thesis in the framework of the University of Debrecen Food Sciences Doctoral School for the purpose of obtaining a doctoral degree (Ph.D.) at the University of Debrecen.

Debrecen, 7th of January, 2026.



signature of the candidate

STATEMENT

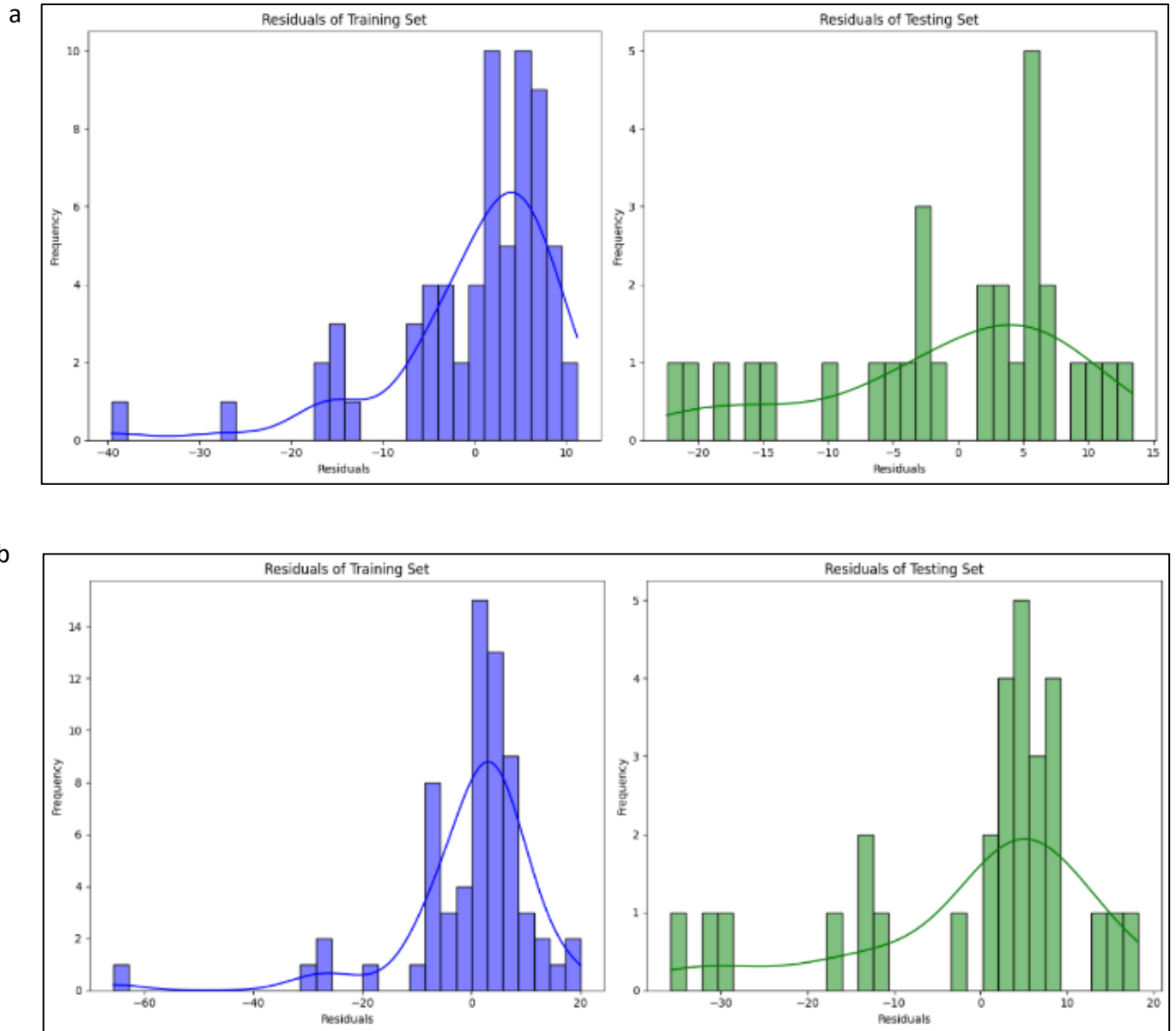
I hereby certify that the doctoral candidate Chaima Neji has carried out her work under my supervision within the framework of the above-mentioned Doctoral School between 2021-2026. The candidate has made a decisive contribution to the results of the thesis through her independent creative work, and the thesis is the candidate's independent work. I recommend that the thesis be accepted.

Debrecen, 7th of January, 2026.

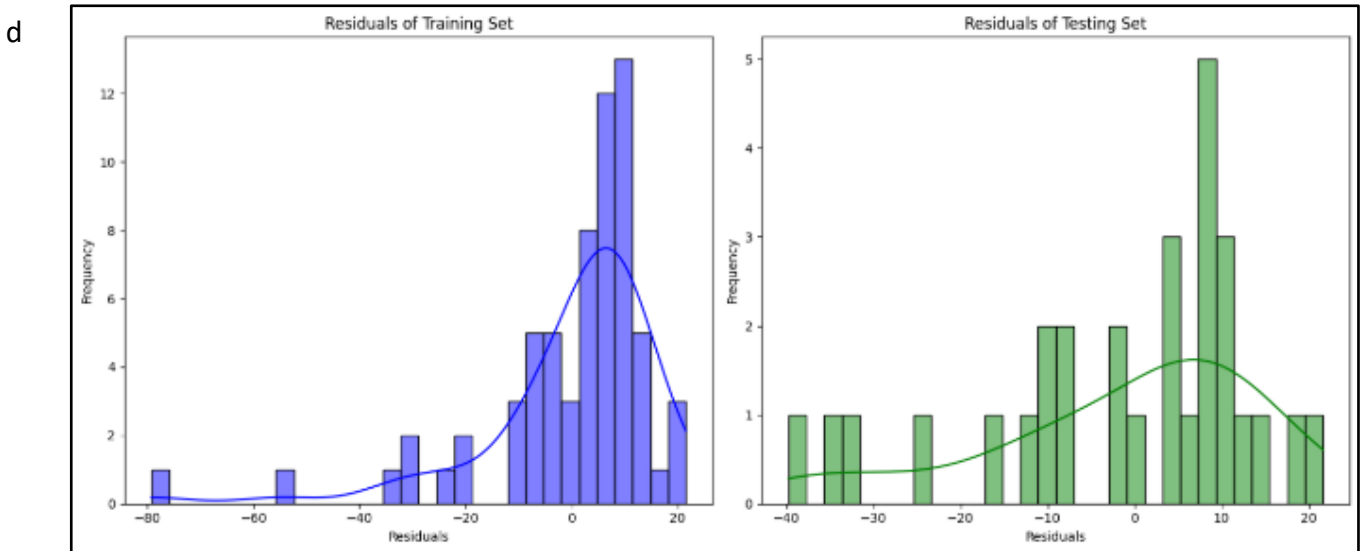
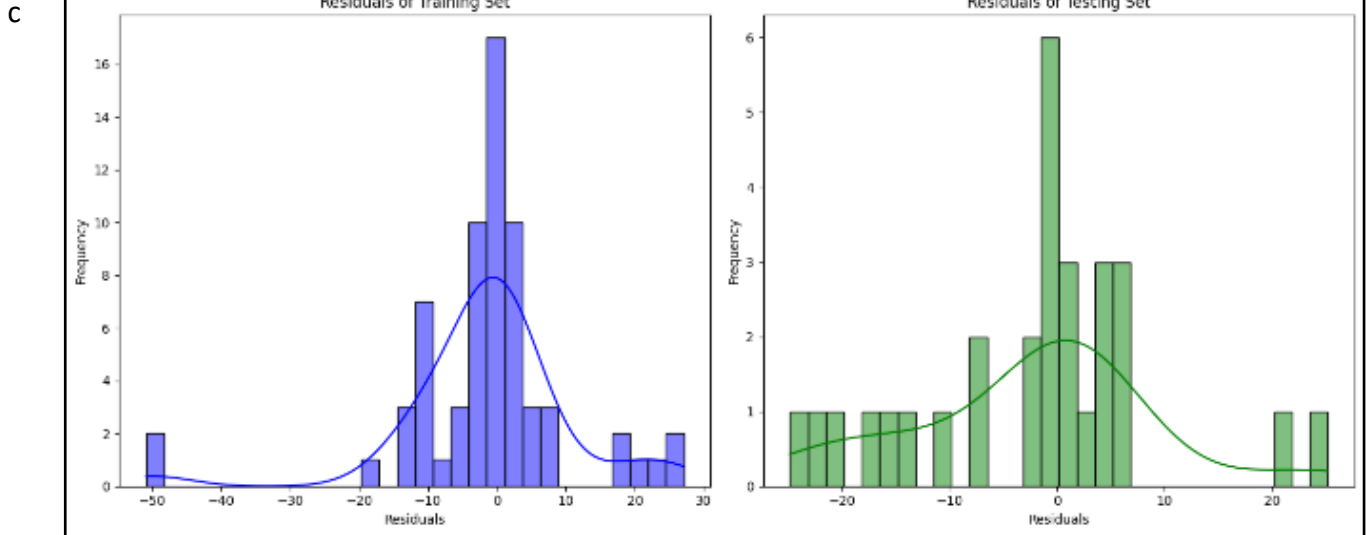


signature of the supervisor(s)

12. ANNEXES

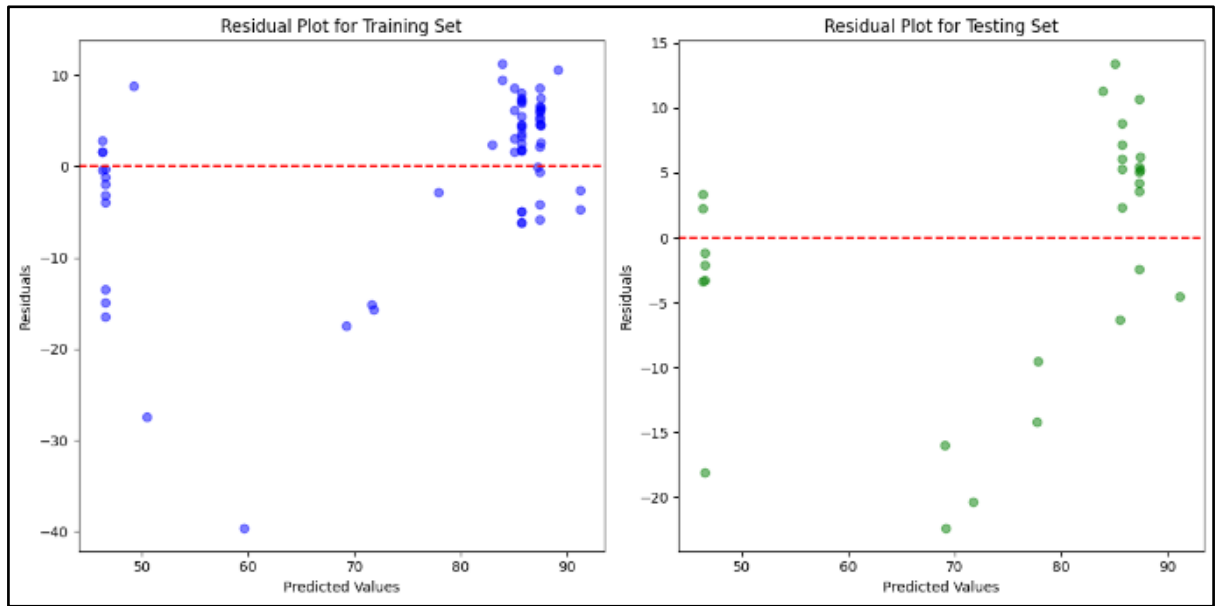


Annex 1a. Distribution of the training and testing set residuals frequency for the different adopted models: (a) Voting_reg, (b) Polynomial model

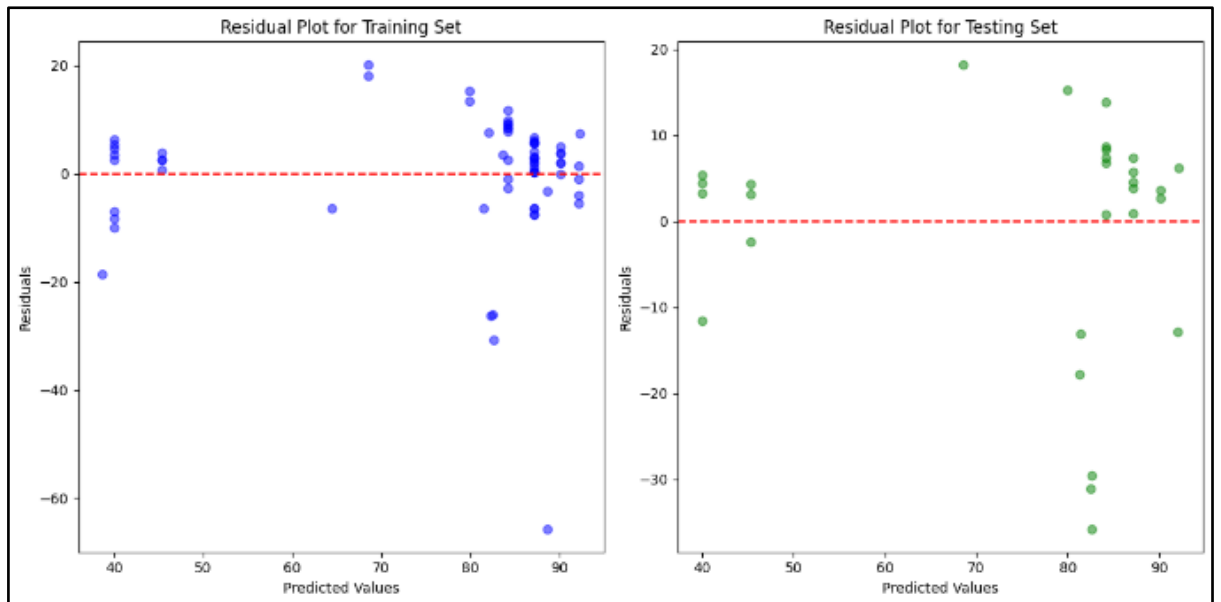


Annex 1b. Distribution of the training and testing set residuals frequency for the different adopted models: (c) Nearest Neighbors Regression model, (d) LinearRegression for model evaluation

a

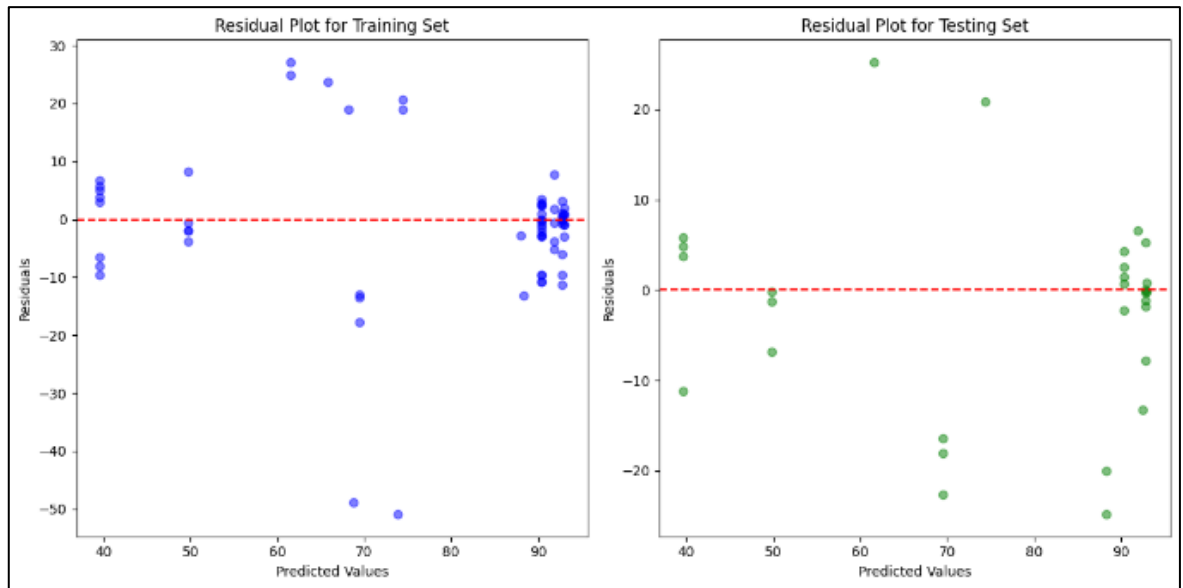


b

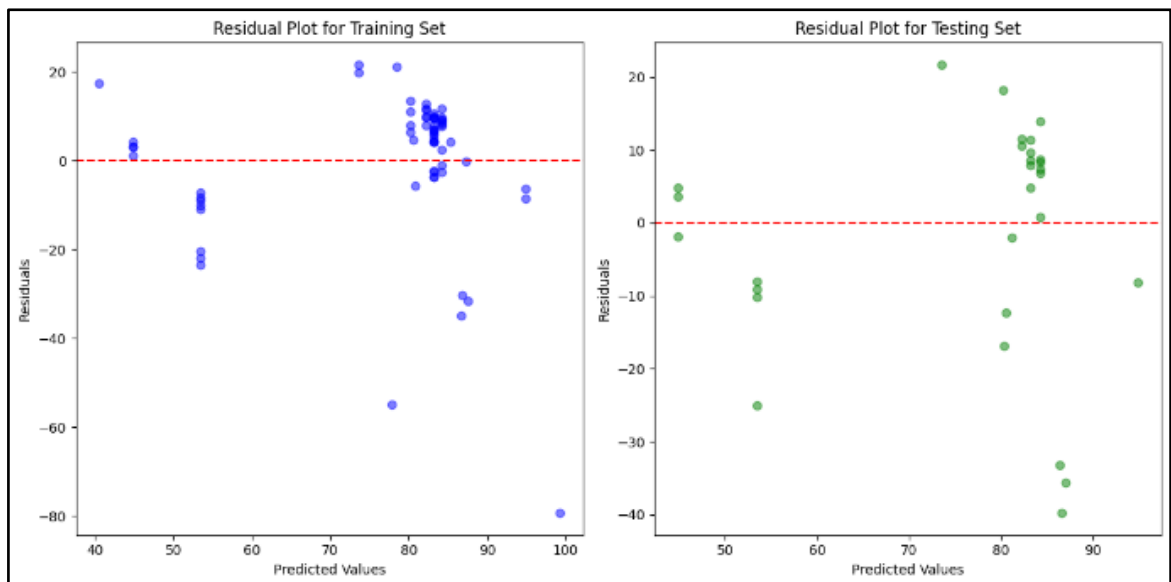


Annex 2a. Distribution of the residual frequency of training and testing set obtained by the models: (a)Voting_reg, (b) Polynomial model

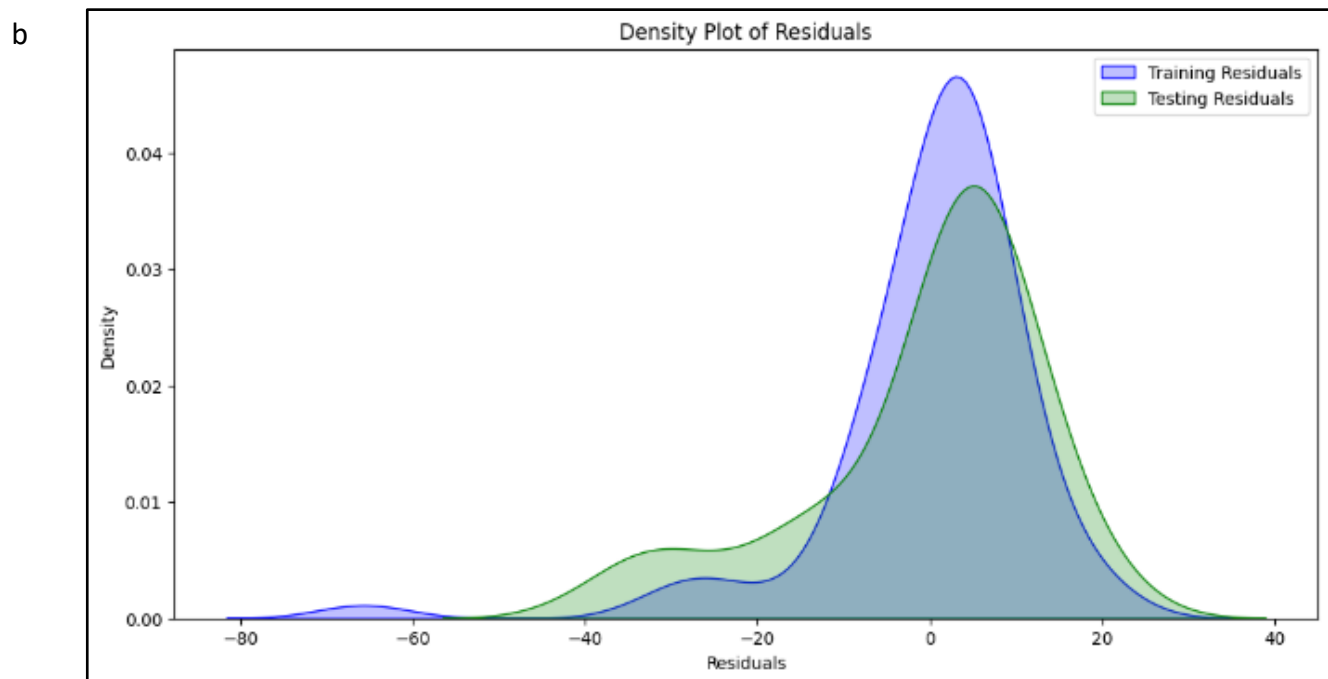
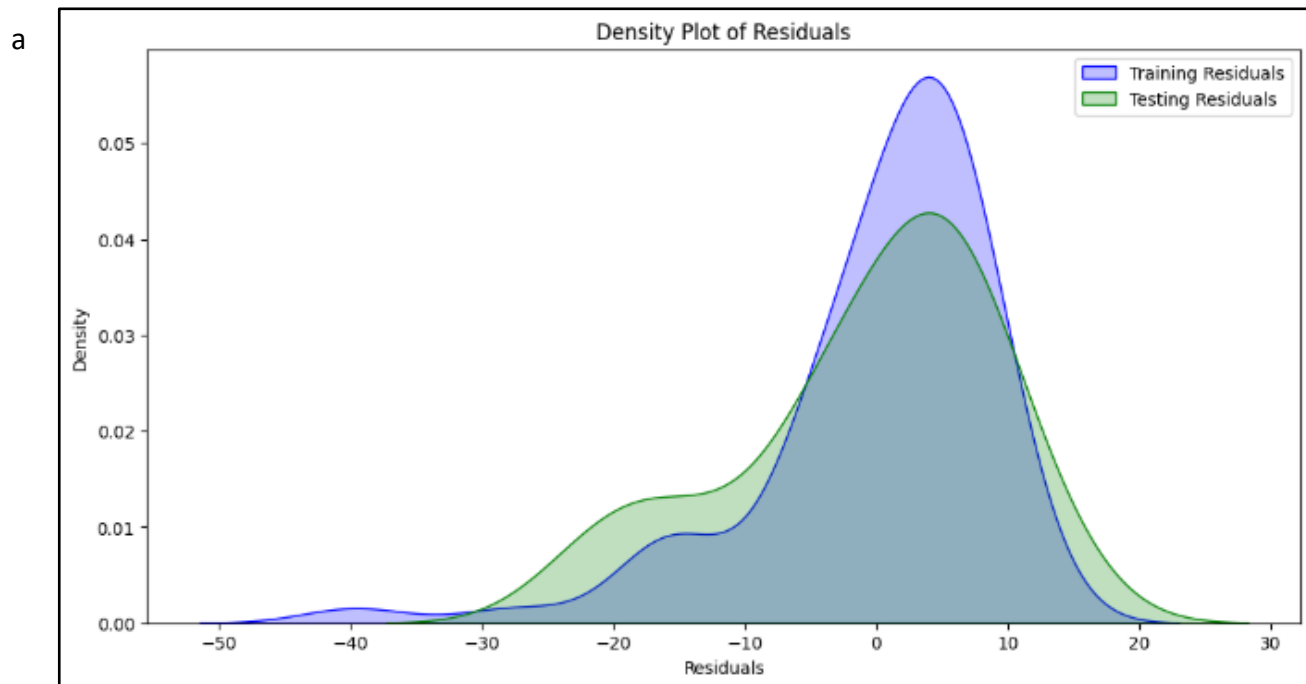
c



d

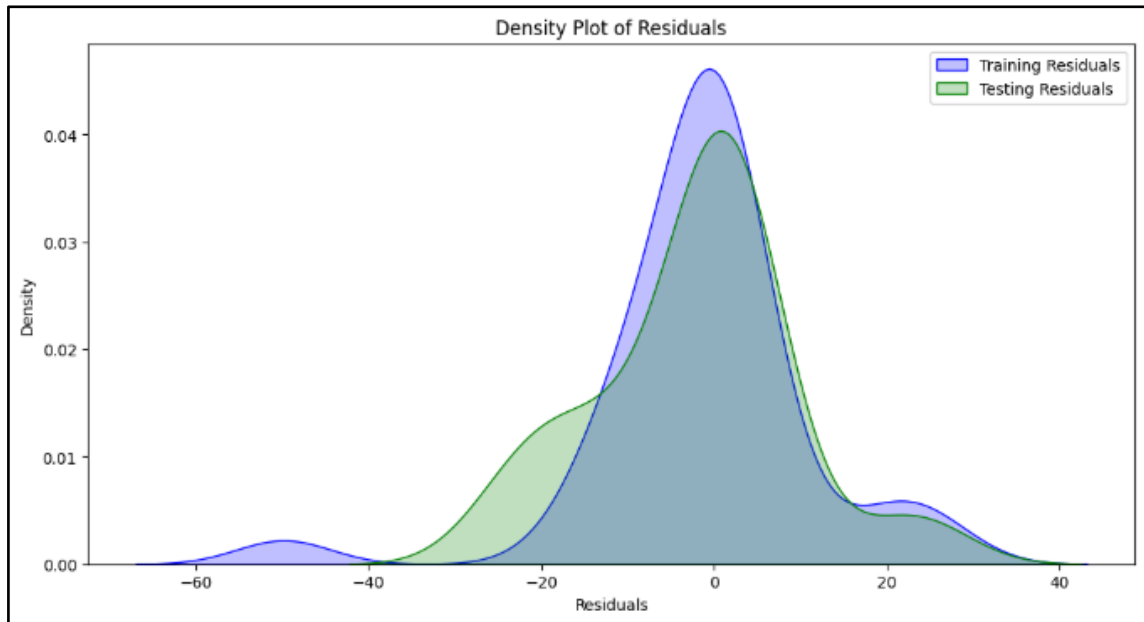


Annex 2b. Distribution of the residual frequency of training and testing set obtained by the models: (c) Nearest Neighbors Regression model, (d) LinearRegression

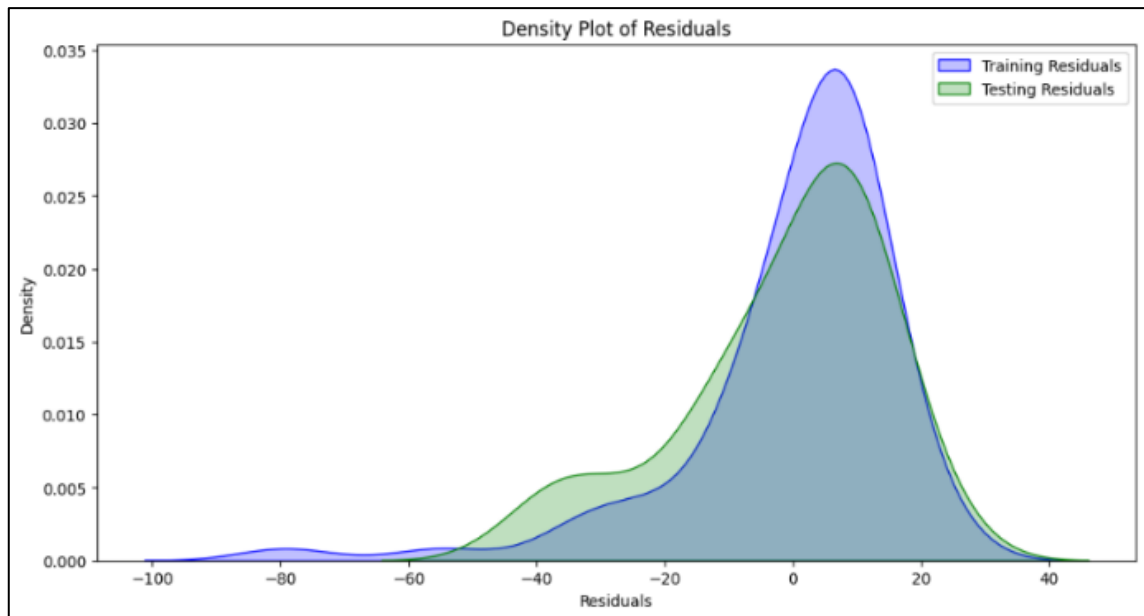


Annex 3a. Residuals density plot of the training and testing dataset obtained by the following models: (a)Voting_reg, (b): Polynomial model

c

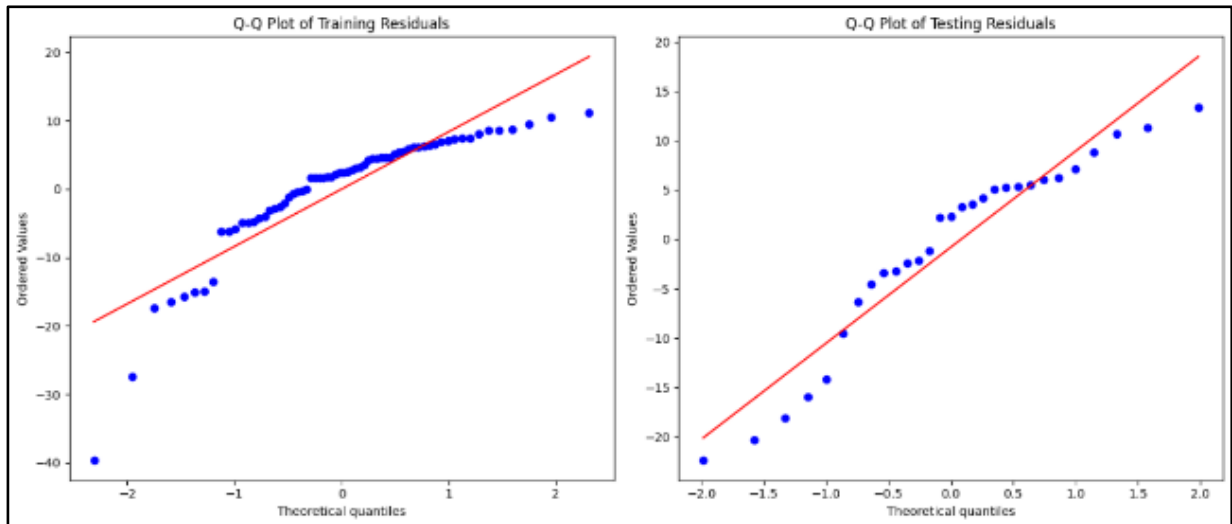


d

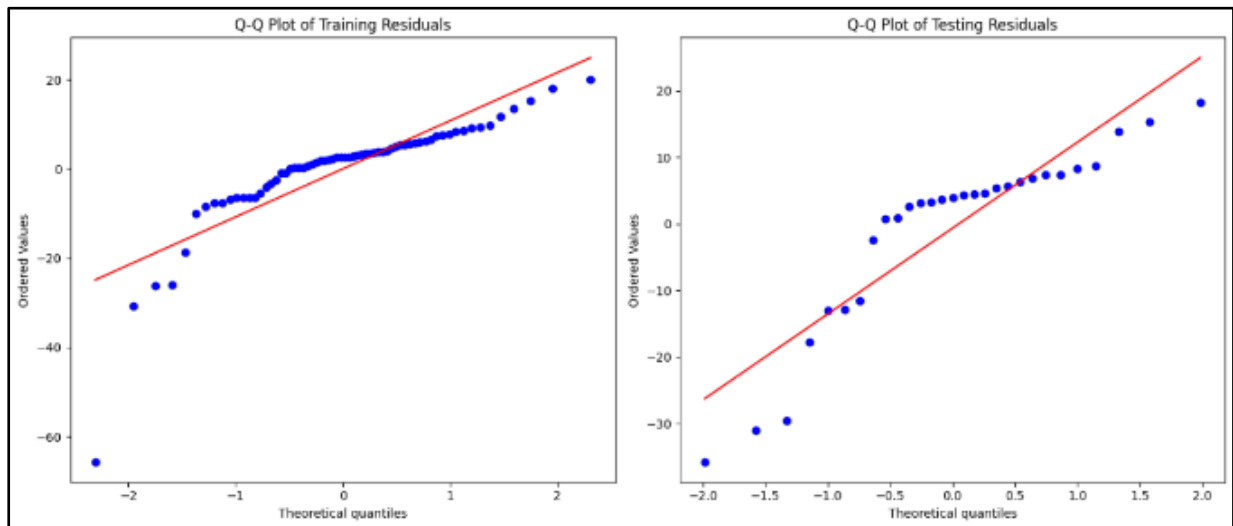


Annex 3b. Residuals density plot of the training and testing dataset obtained by the following models: (c): Nearest Neighbors Regression model, (d): LinearRegression

a

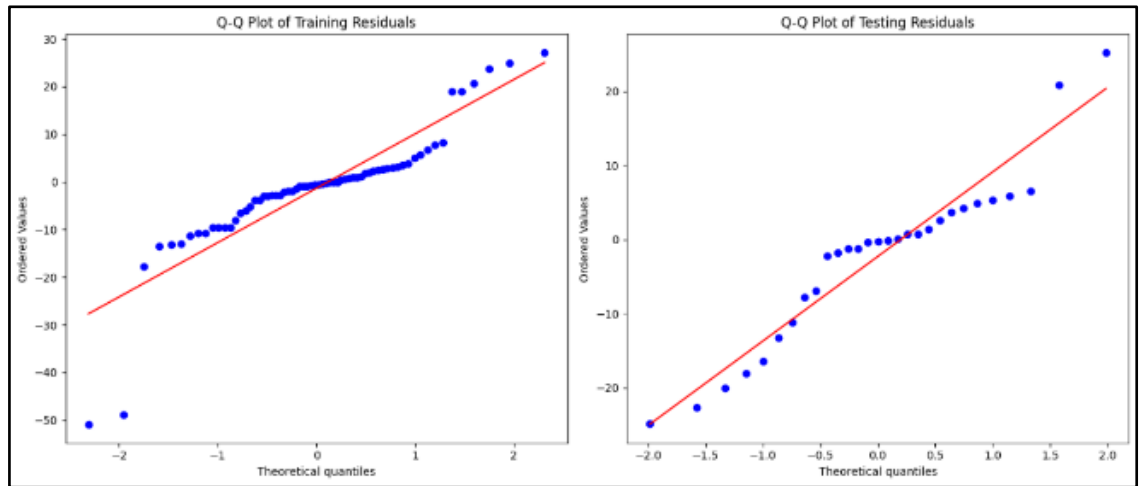


b

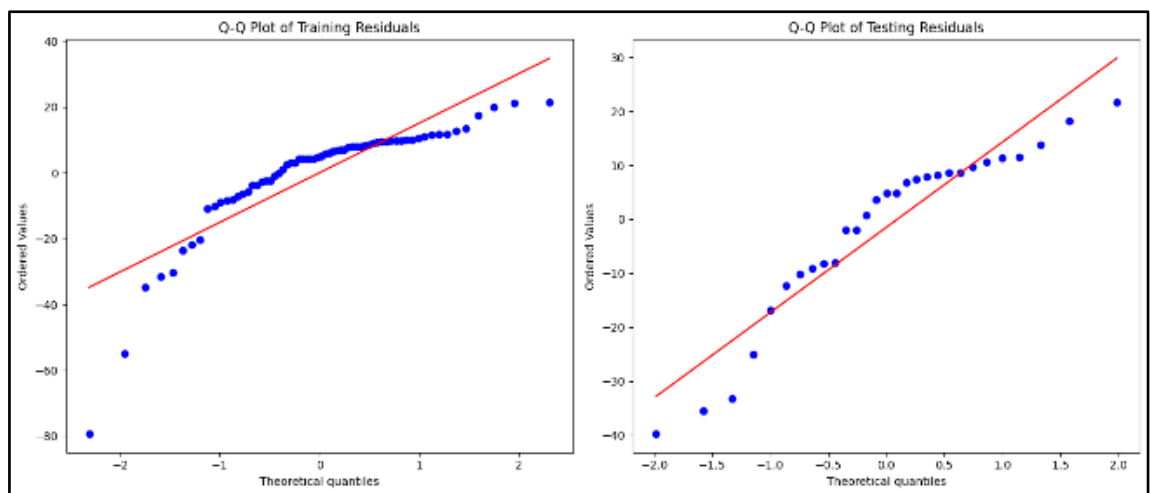


Annex 4a. Quantile-quantile (Q-Q) plot to assess the distribution of the training and the testing residuals represented for all the tested models: (a) Voting_reg, (b) Polynomial model

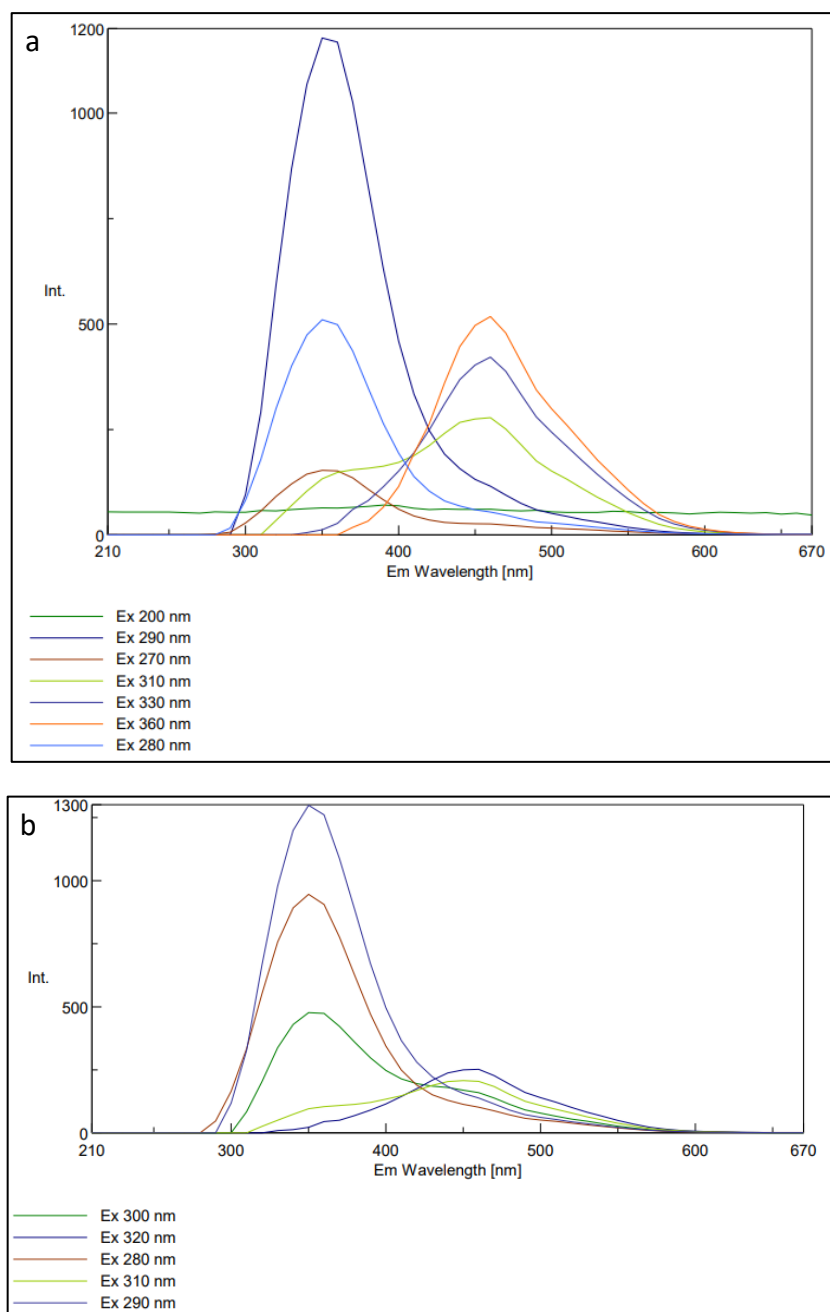
c



d

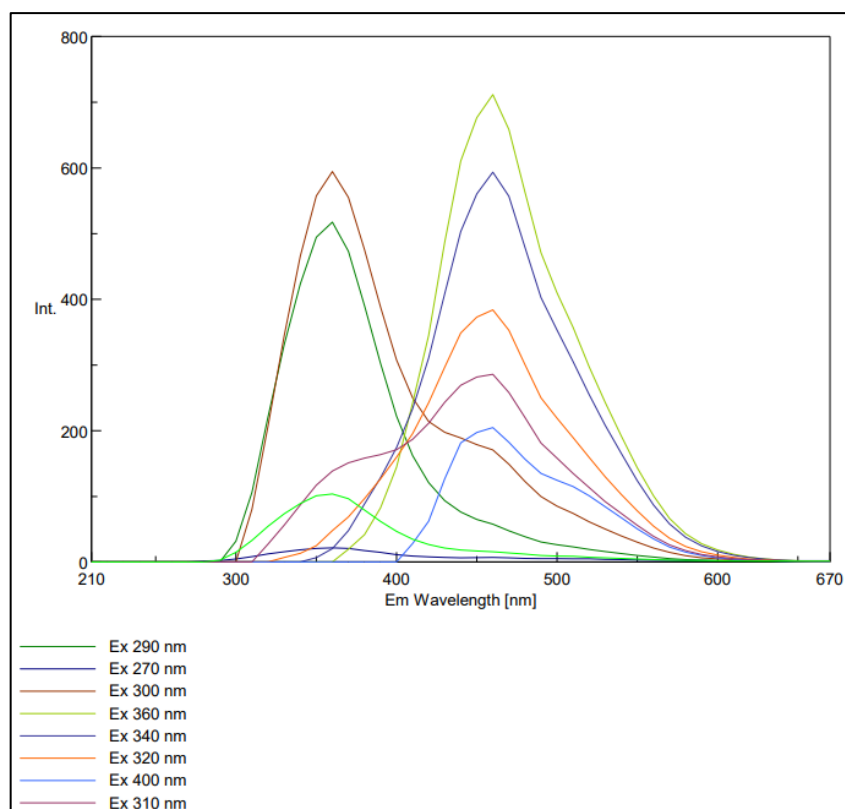


Annex 4b. Quantile-quantile (Q-Q) plot to assess the distribution of the training and the testing residuals represented for all the tested models: (c) Nearest Neighbors Regression model, (d) LinearRegression



Annex 5 a. Fluorescent emission spectra of treated and native soybean protein isolate
a: Enzymatic treatment, b: physical modifications

C



Annex 5 b. Fluorescent emission spectra of treated and native soybean protein isolate
c: native soybean protein isolate

Annex 6: Potential epitope (common sub-sequence) and the original peptide sequence (shared sequences)

Common Subsequence: ITP

Shared Sequences: ['VKGGLSIITPPEKEPR', 'GGLSIITPPEKEPRQK', 'FEITPEK NPQ', 'GGLSIITPPEKEPR']

Common Subsequence: NPQ

Shared Sequences: ['LHQNIGSSSSPDIYNPQAGRIK', 'FEITPEKNPQ', 'LHQNIGS SSSPDIYNPQAGR']

Common Subsequence: IETWNP

Shared Sequences: ['IETWNPNNKP', 'SEGGLIETWNPSNK']

Common Subsequence: GRGDDDDDDDDDD

Shared Sequences: ['SKWQHQQDSCRKQLQGVNLTPCEKHIMEKIQGRGDDDDDDDDDD', ' GRGDDDDDDDDDD']

Common Subsequence: GRG

Shared Sequences: ['SKWQHQQDSCRKQLQGVNLTPCEKHIMEKIQGRGDDDDDDDDDD', ' AEMIDLAAMLSEGRG', 'GRGDDDDDDDDDD']

Common Subsequence: GVNLTTPCEKHIMEKIQ

Shared Sequences: ['SKWQHQQDSCRKQLQGVNLTPCEKHIMEKIQGRGDDDDDDDDDD', ' GVNLTTPCEKHIMEKIQ']

Common Subsequence: SKWQHQQDSCRKQ

Shared Sequences: ['SKWQHQQDSCRKQLQGVNLTPCEKHIMEKIQGRGDDDDDDDDDD', ' SKWQHQQDSCRKQKQ']

Common Subsequence: SGN

Shared Sequences: ['LSGNK', 'FSGNRGPLVQPR']

Common Subsequence: GSS

Shared Sequences: ['LHQNIGSSSSPDIYNPQAGRIK', 'MPACGSS', 'LHQNIGSSSS PDIYNPQAGR']

Common Subsequence: RGP

Shared Sequences: ['FSGNRGPLVQPR', 'KTCENLATYYRGPCF']

Common Subsequence: PCF

Shared Sequences: ['QNWFCEAWPCF', 'KTCENLATYYRGPCF']

Common Subsequence: LQG

Shared Sequences: ['SKWQHQQDSCRKQLQGVNLTPCEKHIMEKIQGRGDDDDDDDDDD', 'H IVDKLQGRDEDEEK']

Common Subsequence: QQQ

Shared Sequences: ['EKPQQQSSRRGS', 'QQQQQQKSHGGR']

Common Subsequence: RRGS

Shared Sequences: ['EKPQQQSSRRGS', 'GRRGSESEEGDAIVK']

Common Subsequence: RGS

Shared Sequences: ['EKPQQQSSRRGS', 'RGSESEEGDAIVK']

Common Subsequence: QQS

Shared Sequences: ['EKPQQQSSRRGS', 'QQSQETDVIVK']

Common Subsequence: GGL

Shared Sequences: ['VKGGLSIITPPEKEPR', 'GGLSIITPPEKEPRQK', 'SEGGLIETWNPSNK', 'GGLSIITPPEKEPR']

Common Subsequence: SNK

Shared Sequences: ['SRNPIYSNKFGK', 'SEGGLIETWNPSNK', 'SPIYSNKFGK']

Common Subsequence: SEG

Shared Sequences: ['SEGGLIETWNPSNK', 'AEMIDLA AKMLSEGRG']

Common Subsequence: AIVK

Shared Sequences: ['GSEEESEEGDAIVK', 'GRRGSEEESEEGDAIVK', 'RGSEEESEEGDAIVK', 'NEDEEKGAIVKVK']

Common Subsequence: IVK

Shared Sequences: ['GRRGSEEESEEGDAIVK', 'QQSQETDVIVK', 'NEDEEKGAIVKVK', 'GSEEESEEGDAIVK', 'RGSEEESEEGDAIVK']

Common Subsequence: EDEEK

Shared Sequences: ['NEDEEKGAIVKVK', 'HIVDKLQGRDEDEEK']

Common Subsequence: GGLSIITPPEKEPR

Shared Sequences: ['VKGGLSIITPPEKEPR', 'GGLSIITPPEKEPRQK', 'GGLSIITPPEKEPR']

Common Subsequence: LHQNIGSSSSPDIYNPQAGR

Shared Sequences: ['LHQNIGSSSSPDIYNPQAGRIK', 'LHQNIGSSSSPDIYNPQAGR']

Common Subsequence: RGSEEESEEGDAIVK

Shared Sequences: ['GRRGSEEESEEGDAIVK', 'RGSEEESEEGDAIVK']

Common Subsequence: GSEEESEEGDAIVK

Shared Sequences: ['GRRGSEEESEEGDAIVK', 'RGSEEESEEGDAIVK', 'GSEEESEEGDAIVK']

Common Subsequence: SEE

Shared Sequences: ['RGSEEESEEGDAIVK', 'GRRGSEEESEEGDAIVK', 'IVGVSEETTGVK', 'GSEEESEEGDAIVK']

Common Subsequence: PIYSNKFGK

Shared Sequences: ['SRNPIYSNKFGK', 'SPIYSNKFGK']

Common Subsequence: EQIEELSKNAK

Shared Sequences: ['ISREQIEELSKNAK', 'EQIEELSKNAK']

Common Subsequence: SRE

Shared Sequences: ['ISREQIEELSKNAK', 'SREEETTEWEEEVAK']

Common Subsequence: EETT

Shared Sequences: ['IVGVSEETTTGVK', 'SREEETTEWEEEVAK']

Common Subsequence: VQP

Shared Sequences: ['QQPQDAVQPF', 'FSGNRGPLVQPR']

Common Subsequence: LPP

Shared Sequences: ['FLPPVT', 'LLPPFHQASSLLR']

Common Subsequence: PFNQL

Shared Sequences: ['FLPFNQL', 'PFNQLAG', 'PFNQL']

Common Subsequence: FLP

Shared Sequences: ['FLPFNQL', 'FLPPVT']

Common Subsequence: QLA

Shared Sequences: ['SQLALTNPT', 'PFNQLAG']

13. ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to the University of Debrecen for providing an exceptional academic environment and to the Stipendium Hungaricum for their financial support.

I extend my profound appreciation to my supervisor, **Prof.Dr. Péter Sipos**, for his supervision, support, and guidance throughout my research journey, and for being open and receptive to my ideas.

I am also thankful to **Dr. Endre Máthé**, the head of the Institute of Nutrition, and to **Prof. Dr.József Prokisch** for accepting me in their laboratories, and for their advice and encouragement. I would also like to extend my gratitude to **Ms. Judit Szepesi, Dr. Éva Szabolcsy-Domokosné, Mrs. Judit Koroknai, Dr. Diána Ungai, and Bence Pecsénye** for their help.

I am deeply grateful to my parents and sisters for their support and for never giving up on me. I also wish to acknowledge my friends from the Institute of Nutrition and the Nanofood laboratory, and my friends Arjun, Sidra, and Hana for their continuous support and motivation, especially during the most challenging times.