

**Purification and Characterization of Allergens
for Walnut Allergy Analysis**

BINGYU CHEN

2024

Contents

General Introduction.....	1
Chapter 1 Purification of Vicilin N-Terminal Fragments from Walnut, Cashew,	
and Pecan Seeds.....	6
Introduction	6
Materials and Methods	9
Results.....	18
Discussion	34
Chapter 2 Expression, Purification, and Characterization of Recombinant	
Vicilin N-Terminal Fragments of Walnut, Pecan, and Cashew	37
Introduction	37
Materials and Methods	40
Results.....	50
Discussion	67
Chapter 3 Expression, Purification, and Characterization of Recombinant PR-	
10 and Profilin in Walnut	69
Introduction	69
Materials and Methods	72
Results.....	79
Discussion	92

Summary	96
References.....	100
Acknowledgments.....	119

General Introduction

In recent years, with the growth of the global population and increasing attention to healthy diets, plant proteins have gained significant attention due to their environmental sustainability and nutritional value (Qin et al., 2022; Stone et al., 2022). Tree nuts, as a high-quality source of plant protein, provide not only ample protein but also dietary fiber, unsaturated fatty acids, vitamins, and minerals (O'Neil et al., 2005). Consumption of nuts has been shown to provide various health benefits, including reducing the risk of cardiovascular diseases and strokes, lowering the risk of type II diabetes and hypertension, and helping to maintain a healthy weight (Bechthold et al., 2017; Coates et al., 2018; Jiang et al., 2002; Martínez-González & Bes-Rastrollo, 2011; Mohammadifard et al., 2015; Rajaram & Sabaté, 2006). With the increasing awareness of the nutritional value of nuts, global nut production has grown significantly over the past decade. In the 2022/23 season, the total global nut production reached 5.3 million metric tons, a 160.73% increase from the 2013/14 season (INC, 2023). Japan's walnut imports have more than doubled over the past two decades, reaching 18,033 tons in 2023 (Ministry of Agriculture, Forestry and Fisheries of Japan, 2023).

However, the widespread consumption of nuts has also brought about a significant health issue, namely nut allergies. Nut allergies are a serious food allergy reaction, with prevalence varying between 1% and 3% across different regions and age groups (McWilliam et al., 2020). It is important to note that differences in age groups, geographical regions, and definitions of food allergy have led to variations in prevalence rates reported across studies. As one of the nine priority allergens, nuts may cause severe or even fatal allergic reactions (USDA, 2024). Different nuts contain different allergens,

leading to varying types of allergic reactions and individual responses. Exposure routes also affect the symptoms of allergic reactions, ranging from isolated oral symptoms to severe anaphylactic shock (Crespo et al., 2006; Dunlop & Keet, 2018). Among various nuts, walnuts, and cashews are known for their significant allergenic potential (Borres et al., 2022). Notably, a national survey on immediate-type food allergies in Japan in 2020 reported that 56.5% of tree nut allergy cases were attributed to walnuts, making walnuts the leading cause of tree nut allergies (Sugizaki et al., 2023). Walnuts and pecans belong to the same family, *Juglandaceae*, and their allergens exhibit a high degree of cross-reactivity (Andorf et al., 2017; Maloney et al., 2008; Smeekens et al., 2018). Therefore, studying walnuts and pecans together is essential for a comprehensive understanding of the characteristics of their allergenic proteins and potential cross-reactivity mechanisms.

Vicilins are an important class of seed storage proteins found in various nuts, including walnuts (Jug r 2, Jug r 6), pecans (Car i 2), and cashews (Ana o 1) (Dubiel et al., 2018; S. Teuber, Jarvis, et al., 1999; Wang et al., 2002; Y. Zhang et al., 2016). Vicilins undergo post-translational modifications during seed maturation, forming water-soluble N-terminal fragments. These fragments typically contain a motif composed of cysteine residues (C-X-X-X-C-X_[10–12 residues]-C-X-X-X-C) and form an α -helix-turn- α -helix structure, affecting the antimicrobial activity and antigenicity of the peptide (Maruyama, 2021; Payne et al., 2020). Although the antigenicity of vicilins has been widely studied, existing research indicates cross-reactivity between walnuts and pecans, cashews and pistachios (Andorf et al., 2017; Maloney et al., 2008; Smeekens et al., 2018). However, the structure and antigenicity of the N-terminal fragments, as well as their cross-reactivity between different nuts, still

require further investigation. Therefore, this study will focus on analyzing the N-terminal fragments of vicilins in walnuts, pecans, and cashews by purifying and characterizing these fragments to gain a deeper understanding of their allergenic mechanisms and potential cross-reactivity.

Pollen-Food Allergy Syndrome (PFAS) is an IgE-dependent allergic reaction that typically occurs when individuals allergic to pollen consume certain plant-based foods, resulting in symptoms such as mild oral and throat irritation and itching, although systemic symptoms may also occur in some cases (Ma et al., 2003; Poncet et al., 2020; Waserman & Watson, 2011). The food allergens mainly associated with this cross-reactivity include PR-10 (Pathogenesis-Related Protein 10), profilin, TLP (Thaumatococcus-like Protein), and others (Krikeerati et al., 2023). Jug r 5 belongs to the PR-10 family, with its scaffold composed of a highly bent seven-stranded anti-parallel β sheet and three α helices, forming a hydrophobic cavity (Fernandes et al., 2013). Jug r 7 is a profilin, a highly conserved protein with a highly conserved sequence and structure, with high identity even among distantly related organisms (Kapingidza et al., 2019). In Japan, cases of walnut allergy have shown a significant increase in recent years. The proportion of allergic reactions caused by accidental ingestion of walnuts has markedly risen among individuals aged three and above, particularly school-aged children (Sugizaki et al., 2023). This trend may be associated with the potential link between walnut allergens, including Jug r 5 and Jug r 7, and PFAS. However, there is currently a lack of systematic research on the specific relationship between walnut allergens and PFAS. Therefore, studying the structure and antigenicity of Jug r 5 and Jug r 7 is essential to understanding the sensitization mechanism of walnut allergy and the occurrence of PFAS.

To better understand and study nut allergies, it is necessary to categorize the allergens in nuts. Based on their biochemical and immunological properties, food allergens in tree nuts can be broadly classified. Some allergens, which are often more stable, can sensitize individuals as they pass through the gastrointestinal tract and typically dominate in tree nut allergies. For example, storage proteins in walnut, such as Jug r 1, Jug r 2, and Jug r 4, are considered major allergens in primary walnut sensitization (Wangorsch et al., 2017). Other allergens in tree nuts may initially sensitize individuals through inhalation and cause cross-reactions when certain foods are consumed. These allergens are often associated with milder clinical reactions, although they are more common (Willison et al., 2014; Lorenz et al., 2015; Sicherer & Sampson, 2006). It should be noted that although LTPs are not primary sensitizers in PFAS, their three-dimensional structure is stabilized by disulfide bonds, which confer high structural stability compared to the relatively unstable profilin and PR-10 (Borges et al., 2007). This stability enables food-derived LTPs to resist proteolytic digestion and thermal processing, potentially triggering severe allergic reactions (Enrique, 2002).

The allergic patterns of different types are different. Misidentification of major allergens may lead to misdiagnosis, causing patients to avoid unrelated foods to prevent allergic symptoms or delay treatment (Carlson & Coop, 2019; Ma et al., 2003). Therefore, accurately identifying and classifying major allergens in nuts is crucial for improving the accuracy and effectiveness of diagnosis, as well as providing correct dietary advice to patients.

In clinical diagnosis, although total protein extracts of foods are widely used, these preparations have issues such as natural variability, endogenous degradation, lack of standardization, and

sometimes low or absent allergen content (Andersson et al., 2007; Hoffmann-Sommergruber et al., 2008; Vieths et al., 2001). Component-resolved diagnosis (CRD) uses highly purified allergen proteins, including native and recombinant allergens. Although native allergens offer the best choice for analysis in terms of conformation and antigenicity, the characteristics of certain tree nut allergens (such as low solubility, difficulty in purification, and low abundance) limit research progress (Hoffmann-Sommergruber et al., 2008; Kattan & Wang, 2013). For example, only about 25 µg of Pru du 4 can be purified from 1 g of defatted almond powder, while approximately 2.5 to 3 mg of purified recombinant Pru du 4 can be obtained from each liter of culture medium (Tawde et al., 2006). Therefore, recombinant proteins with high purity, high concentration, and structural standardization are widely used, helping to link clinical conditions with sensitization patterns, identify individual allergens as biomarkers for the severity or persistence of food allergies, and recognize individual allergens as markers for cross-reactive allergies.

Overall, studying the allergenic components in nuts, particularly the N-terminal fragments of nut vicilins and PR-10 and profilin in walnuts, is crucial for understanding the mechanisms of nut allergies and developing better diagnostic and therapeutic methods. These studies not only enhance our understanding of nut allergies but also provide a scientific basis for future food safety and allergy prevention.

Chapter 1

Purification of Vicilin N-Terminal Fragments from Walnut, Cashew, and Pecan Seeds

Introduction

According to the International Nut and Dried Fruit Council, global nut production has significantly increased over the past decade, particularly in high-income and middle-income economies, with an average annual growth of 148,725 metric tons and 115,362 metric tons, respectively (INC, 2023). Specifically, walnut production increased from 542,367 metric tons to 1,156,765 metric tons, with a growth rate of approximately 113.28%. Cashew production increased from 601,642 metric tons to 1,095,030 metric tons, showing a growth rate of approximately 82.01%; and pecan production increased from 110,670 metric tons to 164,657 metric tons, with a growth rate of approximately 48.78%. In the global nut production for 2022/23 season, walnuts accounted for 22%, cashews for 20%, and pecans for 3% (INC, 2023). These data suggest that as awareness of dietary composition and food nutrition increases, nuts are becoming increasingly popular and play a significant role in diets.

However, increasing nut consumption has also led to a rise in nut allergy risks. Currently, the EU, USA, Canada, and Australia have classified tree nuts as potential allergenic foods, requiring labeling on food packaging to reduce the risk of allergic reactions among consumers (Gendel, 2012). Among various nuts, walnuts and cashews are particularly known for their significant allergenic potential. Pecans and walnuts belong to the same family and should be considered in cases of walnut allergies (Elizur et al., 2020; Mendes et al., 2019). According to a triennial survey by the Consumer Affairs

Agency of Japan, the number of immediate-type cases caused by walnuts significantly increased, ranking first among tree nuts with 463 cases in 2020, respectively (Sugizaki et al., 2023). In March 2023, walnuts were added to Japan's specific allergen labeling for raw materials, highlighting the growing importance of walnuts as an allergenic food (Consumer Affairs Agency, Government of Japan, 2022).

Vicilin is a storage protein found in both monocotyledonous and dicotyledonous plants, belonging to the cupin superfamily (Che et al., 2019). It has been widely identified as an allergen in walnuts (Jug r 2, Jug r 6), pecans (Car i 2), and cashews (Ana o 1) (Dubiel et al., 2018; S. Teuber, Jarvis, et al., 1999; Wang et al., 2002; Y. Zhang et al., 2016). The core region of mature vicilin in plants is highly conserved and composed of repeating domains of β -sheets and α -helices (Rasheed et al., 2020). In some plant species, vicilin contain a hydrophilic domain at the N-terminus, known as the extension region (Maruyama, 2021). During maturation, the N-terminal sequence of vicilin undergoes post-translational modifications in the seed, where the hydrophilic N-terminal domain is cleaved and accumulates (Shewry et al., 1995). Studies have found that some N-terminal fragments can bind to IgE, as observed in peanuts, almonds, macadamia nuts, and sunflowers (Aalberse et al., 2020; Che et al., 2019; Dubiel et al., 2018; Kabasser et al., 2022). Compared to the core region, the sequence identity of the extension region is relatively low among different plant species, but they share a conserved CX₃CX₁₀₋₁₂CX₃C motif and have two disulfide bonds (J. Zhang et al., 2019). To date, the specific enzyme responsible for cleaving the vicilin N-terminal fragment in walnuts has not been identified. However, in pumpkin seeds, Vacuolar Processing Enzyme (VPE) has been shown to play

a key role in processing the vicilin precursor PV100 into mature vicilin and C2 peptides containing two CX₃C motifs (Yamada et al., 1999).

This chapter aims to purify and characterize the native N-terminal fragments of vicilin from walnuts, pecans, and cashews by focusing on native proteins. First, the native N-terminal fragments from walnuts, pecans, and cashews were extracted using different solutions to account for the varying solubility of proteins from different sources. After initial extraction by gradient ammonium sulfate precipitation, the N-terminal fragments of vicilin were further purified using high-performance liquid chromatography (HPLC). The amino acid sequences of the proteins were confirmed after purification.

Studying these native proteins helps reveal the characteristics of allergens in nuts, particularly the structure and antigenicity of the vicilin N-terminal fragments. These studies enable a deeper understanding of the specific mechanisms by which these fragments trigger allergic reactions. This work not only expands our fundamental scientific knowledge of nut allergies but also provides critical evidence for the accurate diagnosis and effective prevention of allergic reactions. These findings are expected to promote the development of hypoallergenic foods and make significant contributions to public health.

Materials and Methods

Materials

The walnuts and pecans used in this study were sourced from California, USA, while the cashews were sourced from India. All nuts were raw and shelled to ensure consistent quality and processing throughout the study. These nuts were processed by Minoya Co., Ltd. (Hyogo, Japan).

Skim milk was purchased from Megmilk Snow Brand Co., Ltd. (Hokkaido, Japan). The protein marker was obtained from Biodynamics Laboratory, Inc. (Tokyo, Japan). All other reagents were purchased from Nacalai Tesque, Inc. (Kyoto, Japan) and Fujifilm Wako Pure Chemical Corporation (Osaka, Japan).

Preparation of Defatted Nut Powder

The walnuts were soaked in water at 4 °C overnight to remove the brown walnut pericarp and then dried in the draft chamber at room temperature overnight. Neither pecans nor cashews underwent this treatment.

One hundred grams of ground nuts were mixed with 700 mL of hexane and stirred at room temperature for 2 hours, discarding the fat-containing hexane. This process was repeated three times until the hexane became clear. The recovered defatted nut powder was air-dried overnight in a draft chamber and stored at 4 °C until further analysis and processing.

Ammonium Sulfate Precipitation

Preliminary purification was performed by ammonium sulfate fractionation.

Extraction of pecan proteins and cashew proteins: The defatted pecan powders and cashew powders were initially mixed with phosphate-buffered saline (PBS) in a 1:10 (w/v) ratio. The composition of the PBS was as follows: NaCl: 137 mM, KCl: 2.7 mM, Na₂HPO₄: 10 mM, KH₂PO₄: 1.8 mM. The mixtures were stirred gently at room temperature for 2 hours to facilitate the complete dissolution of proteins. Subsequently, the mixtures were centrifuged at 5,000 rpm for 30 minutes to remove undissolved particulates, collecting the clear supernatant for subsequent processing.

Under cooling conditions at 4 °C, finely grounded ammonium sulfate was gradually added to the supernatant to achieve varying levels of saturation (0%, 20%, 40%, 60%, 70%, 75%, and 80% saturation). Prior to ammonium sulfate fractionation, insoluble components were removed by centrifugation and labeled as "ppt 0%" and "sup 0%". After reaching each saturation level, the mixture was stirred continuously for 30 minutes to ensure complete dissolution of the ammonium sulfate and adequate precipitation of proteins. Following the stirring, samples were centrifuged at 5,000 rpm for 30 minutes to effectively separate the precipitated proteins from the remaining supernatant. Protein samples precipitated at different saturation levels were collected and stored at 4 °C in preparation for further purification steps.

Notably, instead of PBS, the extraction of walnut proteins was performed using a high-salt buffer, modified from the method described by Downs et al. (Downs et al., 2014). The extraction buffer was composed of 25 mM Tris-HCl, pH 7.5, containing 1.5 M NaCl and 3.5% (w/v)

polyvinylpolypyrrolidone. All other procedures were identical to those used for cashews and pecans, ensuring consistent experimental conditions and allowing for effective comparability of the data.

Anion Exchange Chromatography for Walnut Protein Purification

HiTrap Q XL column (Cytiva) was utilized to further purify crude walnut protein, which consisted of 70% saturation and 80% saturation ammonium sulfate-precipitated walnut proteins dissolved in 20 mM Tris-HCl buffer (pH 7.5). The solution was placed in a dialysis bag (Biodesign) that had been pre-treated with 20 mM Tris-HCl buffer (pH 7.5) with a molecular weight cutoff of 3.5 kDa. The bag was sealed with dialysis clips and dialyzed against 20 mM Tris-HCl buffer (pH 7.5) at 4 °C, with the buffer being changed every 8 hours for 24 hours.

The purification procedure of the dialyzed crude protein solution involved the following steps: Initially, the column was equilibrated with 25 mL of 20 mM Tris-HCl buffer (pH 7.5). The sample was then loaded onto the column, and the flow-through (Through) was collected. Subsequently, the column was washed with an additional 25 mL of 20 mM Tris-HCl buffer at pH 7.5 (Wash). Elution was performed using a stepwise gradient of NaCl in 20 mM Tris-HCl buffer (pH 7.5), with the following concentrations and volumes: 0.1 M NaCl (8 mL, Elution 1), 0.2 M NaCl (8 mL, Elution 2), 0.3 M NaCl (8 mL, Elution 3), 0.4 M NaCl (8 mL, Elution 4), 0.5 M NaCl (8 mL, Elution 5), 0.6 M NaCl (8 mL, Elution 6), 0.7 M NaCl (8 mL, Elution 7), 0.8 M NaCl (8 mL, Elution 8), 0.9 M NaCl (8 mL, Elution 9), and 1.0 M NaCl (8 mL, Elution 10).

HPLC

Anion Exchange Chromatography

Initially, the pecan protein precipitates (70% and 80% saturation ammonium sulfate precipitation) and cashew protein precipitates (75% and 80% saturation ammonium sulfate precipitation) were dissolved in 20 mM Tris-HCl buffer (pH 7.5) and dialyzed following the same method as used for walnut protein precipitates prior to loading onto the HiTrap Q XL column. The dialyzed crude protein solution was filtered through a 0.20 μ m syringe filter and concentrated to less than 20 mL using a concentrator (Vivaspin, molecular weight cut-off: 5 kDa) at 6000 g and 4 °C.

Next, the concentrated samples were purified using a Q Sepharose Fast Flow XK16 column (GE Healthcare) for anion exchange chromatography. The column was first equilibrated with 10 column volumes of 20 mM Tris-HCl buffer (pH 7.5). After sample loading, the column was washed with 5 column volumes of 20 mM Tris-HCl buffer (pH 7.5) to remove unbound substances. The bound proteins were then eluted using a linear gradient of 20 mM Tris-HCl and 1 M NaCl. The gradient started at 0 mM NaCl and increased to 800 mM NaCl over 80 minutes. The flow rate was maintained at 2 mL/min, with fractions collected every 4 mL for further analysis. The entire procedure was carried out at a constant temperature of 4 °C.

Hydrophobic Interaction Chromatography

The Elution 2 and 3 from the walnut protein purified via HiTrap Q XL column were further purified using a Butyl Toyopearl M-650 column (Tosoh Bioscience) for hydrophobic interaction

chromatography. The column was pre-equilibrated with 10 column volumes of 1.7 M ammonium sulfate, 25 mM Tris-HCl buffer (pH 7.5). After sample loading, the column was washed with 5 column volumes of 1.7 M ammonium sulfate buffer to remove unbound substances. The bound proteins were eluted with a linear gradient of ammonium sulfate, reducing the concentration from 1.7 M to 0 M over 100 minutes at a flow rate of 2 mL/min. Eluates were collected every 2 minutes.

Similarly, cashew protein fractions 19-22 obtained from Q Sepharose Fast Flow XK16 column purification were subjected to the same hydrophobic interaction chromatography process using the Butyl Toyopearl M-650 column. The entire experimental procedure was carried out at a constant temperature of 4 °C.

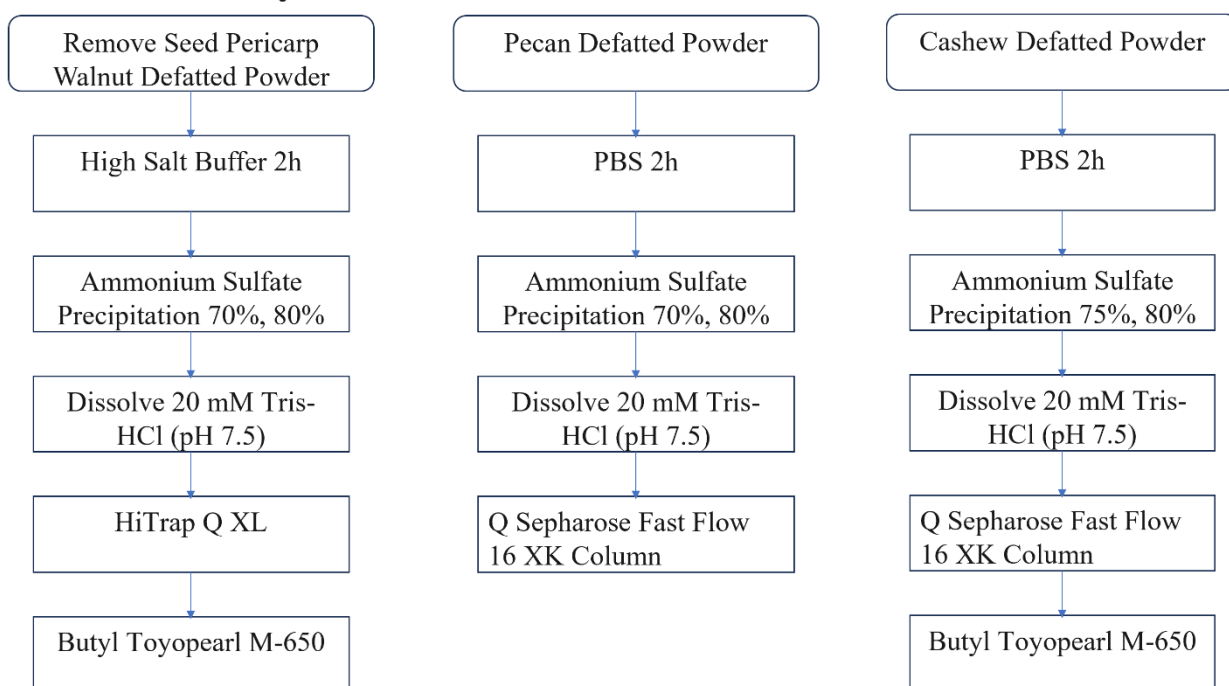


Figure 1. Protein Purification Workflow for Walnut, Pecan, and Cashew Defatted Powders

Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis (SDS-PAGE)

Sample preparation involved the following steps: First, a 2x SDS-PAGE sample buffer (20 mM Tris-HCl, pH 7.2, containing 2% v/v 2-mercaptoethanol, 10 g/L SDS, and 80 mg/L bromophenol blue) was prepared (Matsuoka et al., 2022). Liquid samples were mixed with the 2x SDS-PAGE sample buffer at a 1:1 ratio. Solid samples, on the other hand, were mixed with 1x SDS-PAGE sample buffer at a 1:1 ratio, with the 1x buffer prepared by diluting the 2x buffer with water. The mixtures were then boiled for 5 minutes followed by cooling and centrifugation to ensure consistent protein concentrations across the samples. Protein separation was performed using SDS-PAGE with a protein marker at 220 V for 40 minutes, followed by analysis of the total protein pattern through Coomassie Brilliant Blue (CBB) staining.

Additionally, when protein concentration in liquid samples was low, trichloroacetic acid (TCA) precipitation was used to concentrate the protein by a factor of 10. The specific procedure was as follows: the liquid sample was mixed with TCA to a final concentration of 25%, vortexed thoroughly, and incubated on ice for at least one hour. After incubation, the sample was centrifuged at 13,000 rpm for 20 minutes at 4 °C to ensure complete protein precipitation. The supernatant was discarded, and the pellet was washed with 750 µL of ice-cold acetone, vortexed, and centrifuged again at 13,000 rpm for 5 minutes at 4 °C to remove residual TCA. This washing step was repeated twice. After discarding the supernatant, the pellet was air-dried at room temperature to yield the solid sample. The solid sample was then mixed with 1x SDS-PAGE sample buffer at a 1:1 ratio, boiled, cooled, and subjected to SDS-PAGE analysis as described above.

Western Blot

Proteins were transferred to a nitrocellulose membrane using the semi-dry transfer method following SDS-PAGE. In this part, the gel and nitrocellulose membrane were sandwiched between layers of filter paper and foam pads to ensure even contact and facilitate efficient protein transfer. The filter papers were soaked in three different buffer solutions, applied sequentially from bottom to top. The bottom filter paper was soaked in Buffer C (25 mM Tris, pH 7.6, 20% methanol, 40 mM 6-aminohexanoic acid), and the gel was placed on top of this layer. The pre-treated nitrocellulose membrane was then carefully placed on top of the gel, ensuring the correct orientation for protein migration. The upper filter paper was soaked in Buffer B (25 mM Tris, pH 10.4, 20% methanol), and the topmost filter paper was soaked in Buffer A (0.3 M Tris, pH 10.4, 20% methanol). The higher concentration of Tris in Buffer A enhanced protein transfer efficiency, while the lower concentration in Buffer B helped prevent excessive binding of proteins to the membrane. Air bubbles were carefully removed to ensure the layers were well-aligned and all components were in close contact for optimal transfer. The transfer was performed at 40 mA for 90 minutes.

Following the transfer, the membrane was first incubated in PBSTM (20 mM sodium phosphate, pH 7.6, 150 mM NaCl, 50 g/L skim milk, and 0.05% v/v Tween 20) at room temperature for 20 minutes to block non-specific binding. Subsequently, the membrane was incubated at room temperature for 60 minutes with PBSTM containing the primary antibody. Rabbit antisera specifically raised against the recombinant Jug r 2 N-terminal fragment and Ana o 1 N-terminal fragment were employed for immunodetection in this study. These antisera were generated by immunizing rabbits

with the purified recombinant proteins expressed in the *Pichia pastoris* system, as described in Chapter 2. The membrane was then washed three times with PBSTM, each wash lasting 10 minutes. Following this, the membrane was incubated with the secondary antibody anti-rabbit IgG (Fc) (Promega), diluted 1:1000 in PBSTM, for 60 minutes at room temperature. The membrane was then washed twice with PBSTM, followed by a final wash with TBS-T (10 mM Tris-HCl, pH 7.5, 150 mM NaCl, 0.05% v/v Tween 20). Finally, the membrane was incubated in a buffer containing alkaline phosphatase substrate (10 mM Tris-HCl, pH 7.5, 0.33 mg/mL nitro blue tetrazolium chloride and 0.17 mg/mL 5-bromo-4-chloro-3-indolyl phosphate) at room temperature until sufficient signal development was achieved.

Mass Spectrometry (MS) Analysis

The stained gel bands in the SDS-PAGE were cut into 1 mm square pieces and sequentially treated with decolorizing solution 1 (acetonitrile/125 mM NH_4HCO_3 , v/v=3:1), decolorizing solution 2 (acetonitrile), and washing solution 3 (25 mM NH_4HCO_3) for 15-30 minutes each. Then, the gel pieces were incubated with 10 mM DTT in 25 mM NH_4HCO_3 at 55 °C to 60 °C for 1 hour to achieve reduction. Next, 10 mM iodoacetamide in 25 mM NH_4HCO_3 was added to the gel and incubated at room temperature in darkness for 45 minutes to achieve alkylation. Subsequently, trypsin solution was added to the gel pieces, and the gel was incubated overnight to allow enzymatic digestion. The digested peptides were collected, dehydrated using acetonitrile, and the recovered peptides were concentrated and analyzed by MS. Separation was specifically performed using an EASY-nLC II liquid chromatography system (Thermo Scientific) on a C18 column (0.075 x 150 mm, GL Sciences) with

an acetonitrile gradient containing 0.1% formic acid as the mobile phase. MS data were obtained using an LTQ Orbitrap Velos mass spectrometer (Thermo Scientific) and analyzed through Proteome Discoverer version 1.3 (Thermo Scientific).

Results

Purification of Native Walnut Vicilin N-Terminal Fragment

A rabbit antibody prepared against the recombinant Jug r 2 N-terminal fragment was utilized as the primary antibody for further confirmations. When walnut defatted powder was dissolved in PBS and subjected to ammonium sulfate gradient precipitation, it was found that even with extended extraction time and a tenfold increase in TCA concentration, the target protein could not be detected by either CBB staining or Western blot. This limitation, attributed to the low solubility of walnut proteins in PBS, has also been reported in other studies (Downs et al., 2014; Kong et al., 2019). Walnut proteins, particularly those involved in allergenic responses, tend to have lower solubility in general buffers like PBS, which hampers their detection and purification under standard conditions. Therefore, based on the research by Downs et al., a high-salt buffer was used (Downs et al., 2014). Under this condition, the solubility of walnut proteins increased. Although visibility was still limited in CBB staining, the target protein could be seen in Western blot detection around 17 kDa at ammonium sulfate saturation of 60%, 70%, and 80% (Figure 2).

During the selection of ammonium sulfate precipitation conditions, high molecular bands, particularly those at 45 kDa, were observed to be prominent and abundant below 70% ammonium sulfate saturation. These proteins could potentially correspond to storage proteins, among which some may include N-terminal fragments that were not fully cleaved and remained attached to the mature vicilin's N-terminal region. Their presence likely interferes with the resolution and purity of the target vicilin fragment. Consequently, proteins precipitated at 70% and 80% ammonium sulfate saturation

were collected, as these conditions minimized the interference from non-target proteins. Following dialysis to remove excess ammonium sulfate, the purified protein sample was subjected to further purification steps.

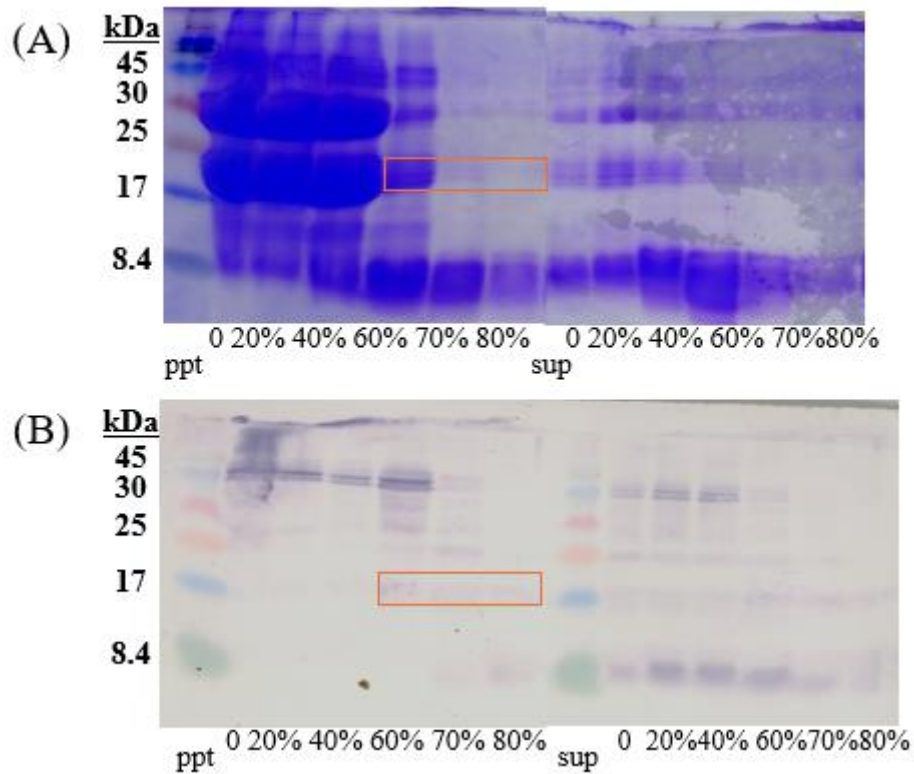


Figure 2. Walnut Protein Ammonium Sulfate Precipitation

(A) CBB Analysis of Walnut Protein Fractions Precipitated by Ammonium Sulfate

(B) Western Blot Analysis of Walnut Protein Fractions Precipitated by Ammonium Sulfate

The boxes in both panels highlight the region corresponding to the target protein around 17 kDa.

Primary antibody: Rabbit anti-recombinant Jug r 2 N-terminal fragment serum antibody.

"Ppt 0%" and "sup 0%" refer to the insoluble and soluble fractions removed by centrifugation before the ammonium sulfate fractionation process.

Dialysis not only removes residual ammonium sulfate but also equilibrates the protein sample to the buffer conditions. This is crucial for maintaining protein stability during further purification steps.

After dialysis to remove ammonium sulfate from the precipitate, further purification of walnut vicilin N-terminal fragment was performed using a HiTrap Q XL column with elution at different NaCl concentrations. Unlike other tree nuts, walnuts have a darker pericarp color, and even after manual removal, some residues may remain, potentially affecting HPLC results. This step effectively prevents pigments from walnut from precipitating into the HPLC column. The use of high-salt conditions in

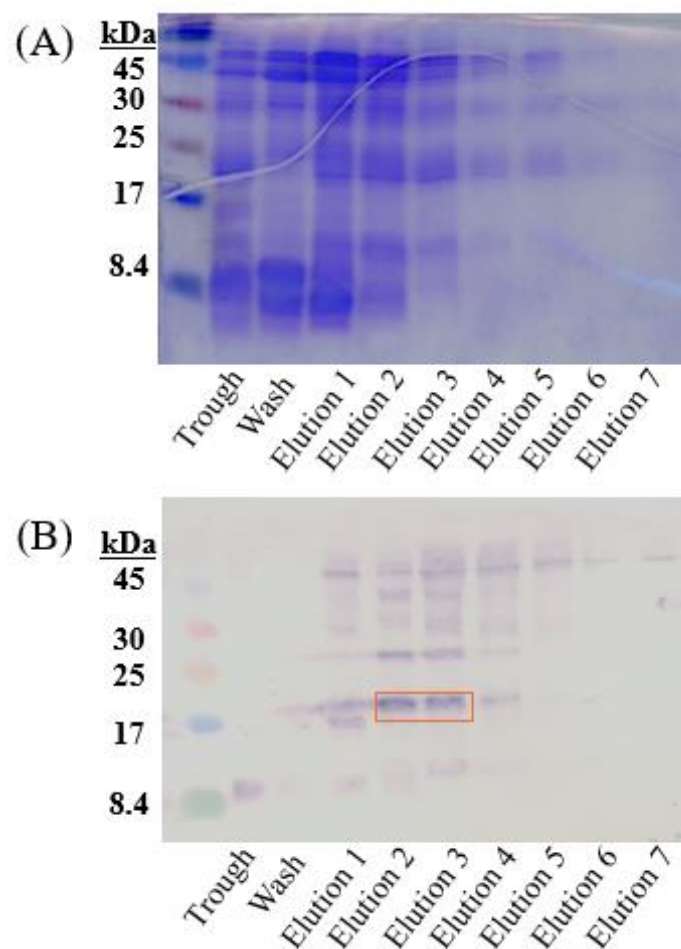


Figure 3. Walnut Protein Purified by HiTrap Q XL Column

(A) CBB Analysis of Walnut Protein Fractions Purified by HiTrap Q XL Column

(B) Western Blot Analysis of Walnut Protein Fractions Purified by HiTrap Q XL Column

The box in panel (B) highlight the region corresponding to the target protein around 17 kDa.

Primary antibody: Rabbit anti-recombinant Jug r 2 N-terminal fragment serum antibody.

combination with anion exchange chromatography is particularly effective for separating vicilin fragments due to their charge properties (Careri & Mangia, 2003). High-salt buffers help in disrupting weak interactions between the protein and other components allowing for more selective binding (Ahmed, 2017).

The Western blot results showed that the target protein was completely bound to the column, with no coloration in the Through and Wash fractions. In Elution 2 and Elution 3, the walnut vicilin N-terminal fragment was more abundant compared to other elution fractions (Figure 3), but the purity was still low. CBB staining of Elution 2 and Elution 3 indicated the presence of a large amount of non-target proteins, suggesting that the elution conditions, while successful in isolating the target protein, still co-eluted with other walnut proteins. To address this, further purification of the protein was performed using hydrophobic interaction chromatography, a method well-suited for separating proteins based on their hydrophobicity.

Further separation of the sample was achieved using a Butyl Toyopearl M-650 column, exploiting the different binding affinities between the hydrophobic regions of the protein surface and the fixed hydrophobic groups (butyl) on the chromatography column. Western blot results showed that at 52mL (fraction 14), the walnut vicilin N-terminal fragments appeared above 17 kDa, indicating successful separation. As the volume increased, the concentration also increased, peaking at 76 mL (fraction 20) before starting to decrease (Figure 4). The target protein bands at fractions 14-18 were faint and had low concentrations, thus not collected. Despite higher concentrations of the target protein at fractions 20-21, the presence of non-target proteins around 25 kDa could cause interference, and

therefore, these fractions were not collected either. Finally, based on CBB results, fraction 19 from 72-76 mL were collected as the final purified native walnut vicilin N-terminal fragment. This final fraction showed a high degree of purity, with minimal contamination from other walnut proteins.

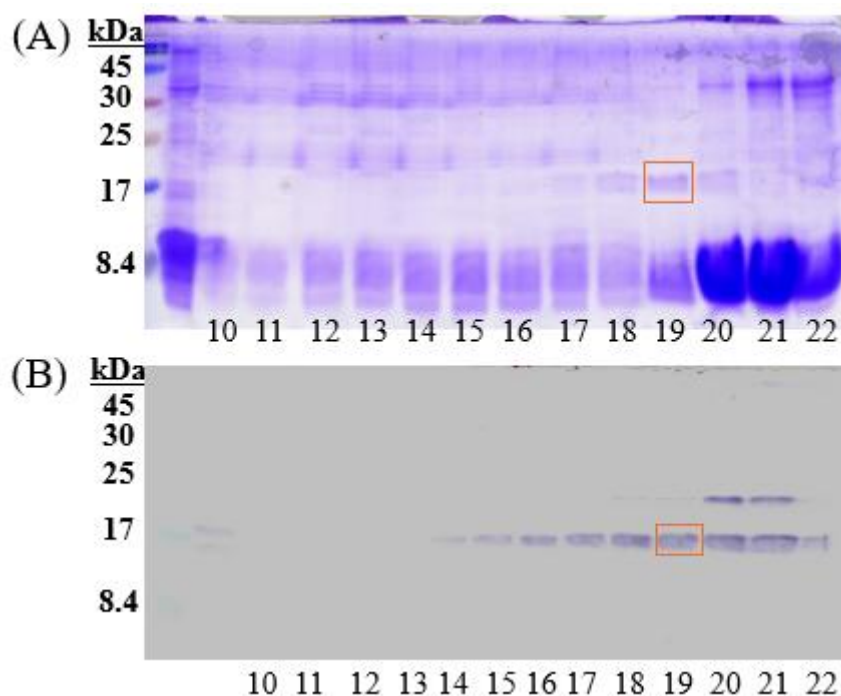


Figure 4. Walnut Protein Purified by Butyl Toyopearl M-650 Column

(A) CBB Analysis of Walnut Protein Fractions Purified by Butyl Toyopearl M-650 Column

(B) Western Blot Analysis of Walnut Protein Fractions Purified by Butyl Toyopearl M-650 Column

The boxes indicate the band excised for subsequent amino acid sequence analysis.

In both panels (A) and (B), the samples were collected at different elution volumes, corresponding to fraction numbers 10 to 22.

Primary antibody: Rabbit anti-recombinant Jug r 2 N-terminal fragment serum antibody.

Purification of Native Pecan Vicilin N-Terminal Fragment

Pecan defatted powder was dissolved in PBS and stirred to extract proteins. The supernatant was collected, and ammonium sulfate was added for gradient precipitation. Results shown in Figure 5 were

obtained by CBB staining and Western blot, using recombinant Jug r 2 N-terminal fragment antibody as the primary antibody. Color development began 5 min after adding the substrate solution and completed after 30 min. The color development results indicated that the pecan vicilin N-terminal fragment began to precipitate when the ammonium sulfate reached 70% saturation, displaying a molecular weight of approximately 17 kDa.

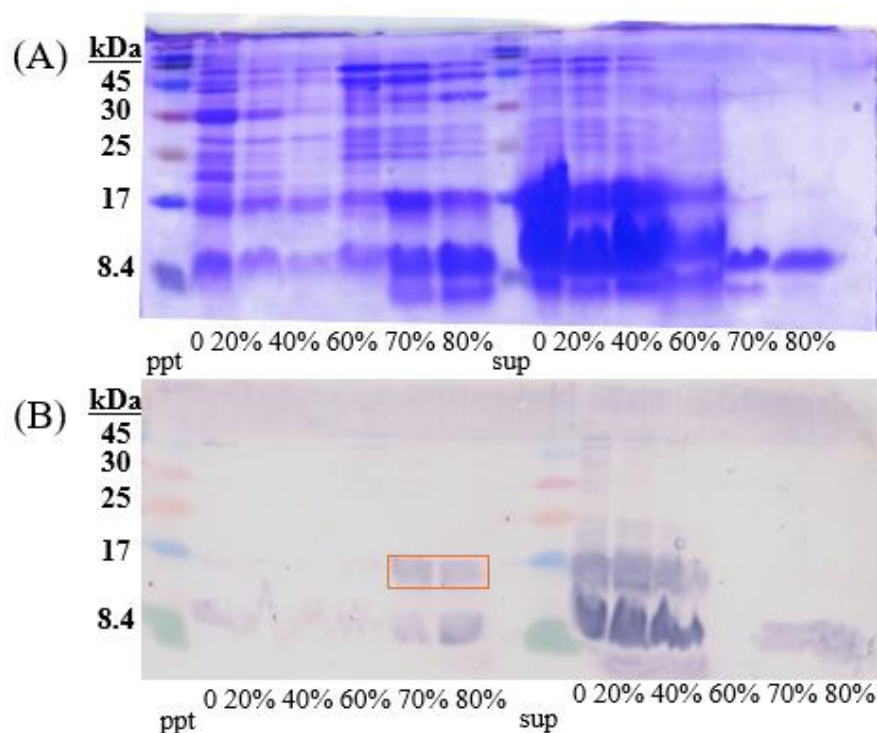


Figure 5. Pecan Protein Ammonium Sulfate Precipitation

(A) CBB Analysis of Pecan Protein Fractions Precipitated by Ammonium Sulfate

(B) Western Blot Analysis of Pecan Protein Fractions Precipitated by Ammonium Sulfate

The box in panel (B) highlight the region corresponding to the target protein around 17 kDa.

Primary antibody: Rabbit anti-recombinant Jug r 2 N-terminal fragment serum antibody.

"Ppt 0%" and "sup 0%" refer to the insoluble and soluble fractions removed by centrifugation before the ammonium sulfate fractionation process.

The precipitate formed at 70% and 80% ammonium sulfate saturation was dissolved in 20 mM Tris-HCl buffer (pH 7.5) and dialyzed. Subsequently, further purification was performed using an anion exchange column. Unlike walnut and cashew, the target pecan vicilin N-terminal fragment could be isolated using only the Q Sepharose Fast Flow XK16 column, separated at 36-40 mL (fraction 10, Figure 6). The band was excised for subsequent amino acid sequence analysis. This simplified purification process indicates that the pecan vicilin N-terminal fragment possesses favorable solubility

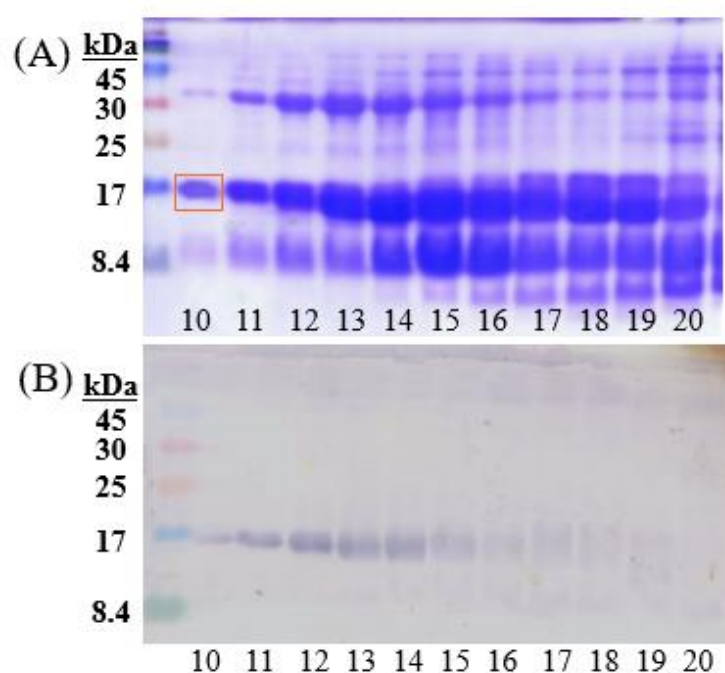


Figure 6. Pecan Protein Purified by Q Sepharose Fast Flow XK16 Column

(A) CBB Analysis of Pecan Protein Fractions Purified by Q Sepharose Fast Flow XK16 Column

(B) Western Blot Analysis of Pecan Protein Fractions Purified by Q Sepharose Fast Flow XK16 Column

The box indicates the band excised for subsequent amino acid sequence analysis.

In both panels (A) and (B), the samples were collected at different elution volumes, corresponding to fraction numbers 10 to 20.

Primary antibody: Rabbit anti-recombinant Jug r 2 N-terminal fragment serum antibody.

and binding characteristics under the conditions used, enabling efficient isolation with fewer steps. These properties may distinguish it from other vicilin proteins, potentially affecting its allergenic potential and behavior in food matrices.

Purification of Native Cashew Vicilin N-Terminal Fragment

The rabbit antibody prepared against the recombinant Ana o 1 N-terminal fragment was used as the primary antibody for confirmation experiments. Cashew defatted powder was dissolved in PBS and stirred to extract proteins. The supernatant was collected, and ammonium sulfate was added for gradient precipitation, as shown in Figure 7. Western blot results indicated that when the ammonium sulfate saturation was above 70%, specifically at 70% saturation, 75% saturation, and 80% saturation,

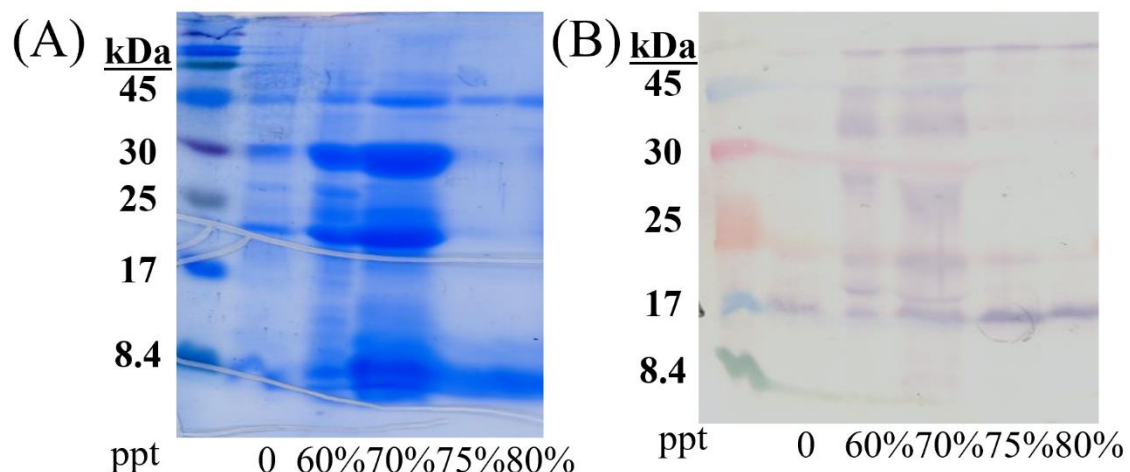


Figure 7. Cashew Protein Ammonium Sulfate Precipitation

(A) CBB Analysis of Cashew Protein Fractions Precipitated by Ammonium Sulfate

(B) Western Blot Analysis of Cashew Protein Fractions Precipitated by Ammonium Sulfate

Primary antibody: Rabbit anti-recombinant Ana o 1 N-terminal fragment serum antibody.

"Ppt 0%" and "sup 0%" refer to the insoluble and soluble fractions removed by centrifugation before the ammonium sulfate fractionation process.

the target band appeared around 17 kDa. However, at 70% ammonium sulfate saturation, there were many high molecular weight bands with high intensities, whereas at saturation above 75%, the other bands were relatively weaker. CBB staining revealed that the target protein band around 17 kDa was faint and of low concentration. Hence, protein precipitates obtained at ammonium sulfate saturation levels above 75% were used for the next step of separation. The use of higher concentrations of ammonium sulfate, followed by protein concentration steps, helps to increase the yield of the target protein, particularly for low-abundance proteins like the vicilin N-terminal fragment in cashew nuts. The strong Western blot signal after concentration confirms the presence of the target protein, despite its low initial abundance.

The protein precipitates at 75% and 80% ammonium sulfate saturation were dissolved in 20 mM Tris-HCl buffer (pH 7.5) and dialyzed. Further purification was performed using the anion exchange column, Q Sepharose Fast Flow XK16 column. The results of CBB staining and Western blot analysis are shown in Figure 8. At 60 mL (fraction 16), the target band around 17 kDa appeared in the Western blot but was not visible in CBB staining due to its low concentration. As the NaCl concentration in the mobile phase increased, more cashew vicilin N-terminal fragments were eluted, reaching the highest levels between 72-88 mL (fractions 19-22), with clear bands around the 17 kDa marker visible in CBB staining. However, due to the presence of numerous non-target proteins at low molecular weights, fractions 19-22 were collected for further purification. The consistent elution of the target protein in these fractions suggests that the vicilin N-terminal fragment has a strong affinity for the anion exchange medium under these conditions, facilitating its separation from other components.

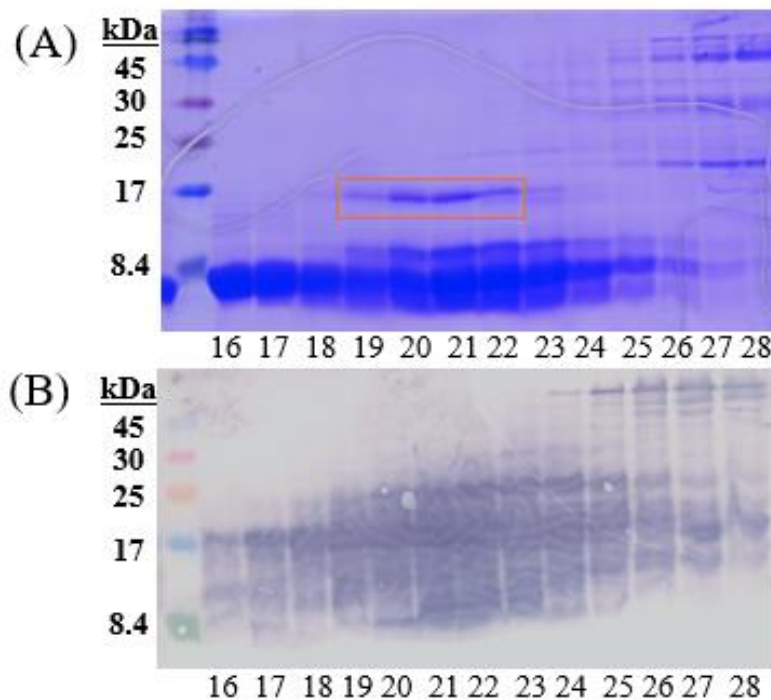


Figure 8. Cashew Protein Purified by Q Sepharose Fast Flow XK16 Column

(A) CBB Analysis of Cashew Protein Fractions Purified by Q Sepharose Fast Flow XK16 Column

(B) Western Blot Analysis of Cashew Protein Fractions Purified by Q Sepharose Fast Flow XK16 Column

The box in panel (A) highlight the region corresponding to the target protein around 17 kDa.

In both panels (A) and (B), the samples were collected at different elution volumes, corresponding to fraction numbers 16 to 28.

Primary antibody: Rabbit anti-recombinant Ana o 1 N-terminal fragment serum antibody.

The fractions 19-22 obtained from the Q Sepharose Fast Flow XK16 column were applied to the Butyl Toyopearl M-650 column for further separation and purification based on differences in hydrophobic interactions. At 100 mL (fraction 26), a band around 17 kDa appeared in Western blot (Figure 9), but CBB results showed that the band was faint and of low concentration, so it was not collected. When the volume was between 128-144 mL (fractions 33-36), the colors of the bands were

the darkest, and the concentration was the highest. This portion was collected as the purified native cashew vicilin N-terminal fragments. The final step using hydrophobic interaction chromatography ensures that the cashew vicilin N-terminal fragment is isolated with a high degree of purity. This method is particularly effective for removing remaining contaminants that have similar charge properties but differ in hydrophobicity, ensuring that the final product is suitable for further studies.

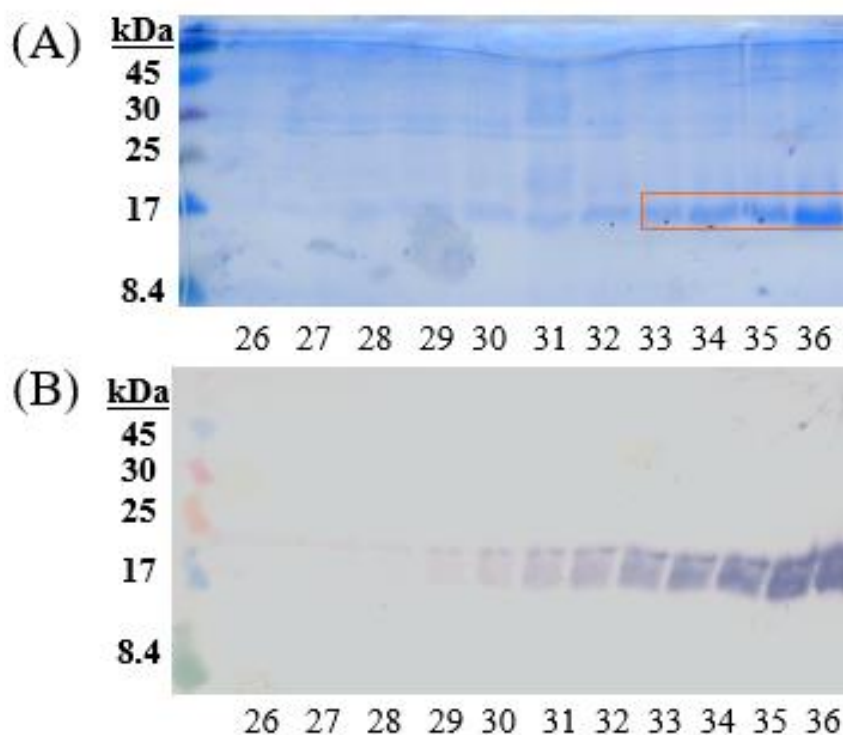


Figure 9. Cashew Protein Purified by Butyl Toyopearl M-650 Column

(A) CBB Analysis of Cashew Protein Fractions Purified by Butyl Toyopearl M-650 Column
 (B) Western Blot Analysis of Cashew Protein Fractions Purified by Butyl Toyopearl M-650 Column
 The box in panel (A) highlight the region corresponding to the target protein around 17 kDa.
 In both panels (A) and (B), the samples were collected at different elution volumes, corresponding to fraction numbers 26 to 36.
 Primary antibody: Rabbit anti-recombinant Ana o 1 N-terminal fragment serum antibody.

Identification of Native Walnut Vicilin N-Terminal Fragment

Following successfully isolating and purifying the N-terminal fragment of walnut vicilin, the approximately 17 kDa protein band was excised from a 15% acrylamide gel and subjected to trypsin digestion. The resulting peptides were then analyzed by MS. The MS analysis identified specific peptides within the N-terminal fragment of the walnut vicilin protein, including **GRDDDDDEENPR** and **QCQEYCR**, which are highlighted in bold in Figure 10. This led to the successful identification of the protein vicilin Jug r 2.0101, with UniProt accession number Q9SEW4. This identification confirmed the presence of the N-terminal fragment and validated that the isolated protein corresponds to the expected vicilin N-terminal fragment.

	10		20		30		40
R	GRDDDDDEEN	P	RDPREQYRQ	CQEY	CRRQGQ	G	QRQQQQCQI
	50		60		70		80
R	CEERLEEDQ	R	SQEERERRR	G	RDVDDQNPR	D	PEQRYEQCQ

Figure 10. Sequence of Walnut Vicilin N-Terminal Region Identified by MS

The regions in bold correspond to the peptides identified by MS that match the sequence of vicilin Jug r 2.0101 (Q9SEW4).

The relatively low sequence coverage observed for the walnut vicilin N-terminal fragment is likely attributed to the low concentration of the fragment within the 17 kDa protein band. This issue highlights the challenges often encountered in detecting low-abundance proteins, where even slight losses during purification can significantly impact the outcome of MS analysis (Liebler & Zimmerman, 2013). In the initial MS analysis, vicilin was not successfully identified, suggesting that its

concentration in the band might have been below the detection threshold. The low concentration was a key factor contributing to the limited sequence coverage. After increasing the concentration, the fragment was successfully detected and matched to the corresponding sequence. This highlights the importance of optimizing sample preparation protocols to enhance protein detection, especially for low-abundance targets.

Given the challenges associated with detecting low-abundance proteins and the limitations of the experimental conditions, the observed sequence coverage and identification results are consistent with expectations. Nevertheless, further refinement of the purification process, potentially including the use of more sensitive MS techniques or enrichment steps, could improve sequence coverage and provide a more comprehensive understanding of the protein's structure and function.

Identification of Native Pecan Vicilin N-Terminal Fragment

Following the same procedure as used for walnut vicilin identification, MS analysis was conducted with the purified pecan samples. The MS data were compared to the pecan protein database, revealing sequences corresponding to the purified protein. According to the Mascot score, the highest-scoring protein was vicilin-like protein (A0A8T1RRY4), with a score of 798, indicating a high degree of matching with the vicilin Car i 2.0101 (Table 1, Figure 11). This high Mascot score suggests a strong and confident identification, underscoring the concentration of the pecan vicilin N-terminal fragment in the sample.

Table 1. Proteins Identified by MS from Purified Car i 2 Vicilin N-Terminal Fragment

Protein (UniProt accession)	Mascot Protein Score
Vicilin-like protein (A0A8T1RRY4)	798
Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) (A0A8T1QHG8)	533

Although the extracted and purified protein had a relatively low sequence coverage of 19% with A0A8T1RRY4, this finding is consistent with the study's focus on the N-terminal fragment, which constitutes only a small portion of the full-length vicilin protein. Considering that the full-length vicilin protein has a molecular weight of approximately 95 kDa, while the extracted fragment is about 17 kDa, the theoretical maximum sequence coverage is naturally limited. Therefore, the observed 19% coverage is within the expected range, especially given that only the N-terminal region was analyzed. Despite the relatively low coverage, the identification of key peptides from the N-terminal region provides strong evidence that the isolated 17 kDa protein corresponds to the expected vicilin fragment.

10 20 30 40
DPESH **RWEFQ** **QCQERCQHEE** **RGQRQAQQCQ** **RRCEEQLRER**
50 60 70 80
ERERERE **EIV** **DPREPRKQYE** **QCRETCEKQD** **PRQQPQCERR**

Figure 11. Sequence of Pecan Vicilin N-Terminal Region Identified by MS

The regions in bold correspond to the peptides identified by MS that match the sequence of vicilin Car i 2.0101(A0A8T1RRY4).

In addition to the vicilin N-terminal fragment, the MS analysis detected the presence of other proteins, including GAPDH (Table 1). These findings suggest that while the target N-terminal fragment was successfully isolated, other seed proteins were still present in the purified samples.

Identification of Native Cashew Vicilin N-Terminal Fragment

The identified amino acid sequences for cashew vicilin N-terminal fragment are as follows (Figure 12):

10	20	30	40
L G F A L A K I D P	E L K Q C K H Q C K	V Q R Q Y D E Q Q K	E Q C V K E C E K Y
50	60	70	80
Y K E K K G R E R E	H E E E E E E W G T	G G V D E P S T H E	P A E K H L S Q C M
90	100		
R Q C E R Q E G G Q	Q K Q L C R F R C Q		

Figure 12. Sequence of Cashew Vicilin N-Terminal Region Identified by MS

The regions in bold correspond to the peptides identified by MS that match the sequence of vicilin Ana o 1.0101(Q8L5L5).

Table 2 lists the proteins matched to the purified native vicilin N-terminal fragment from cashews using MS analysis, along with their Mascot scores. The vicilin-like protein Ana o 1.0101 (Q8L5L5) has the highest Mascot score of 422. While this score is lower than that observed for the pecan vicilin fragment, it still indicates a strong match and suggests that the vicilin-like protein is the predominant component in the purified sample. This score indicates a strong match between the protein and the MS data, suggesting that this protein is most likely present in the sample. Additionally, the presence of

11S globulin and 2S albumin, though with lower scores, suggests potential minor components or contaminants from the cashew protein extract.

Table 2. Proteins Identified by MS from Purified Ana o 1 Vicilin N-Terminal Fragment

Protein (UniProt accession)	Mascot Protein Score
Vicilin-like protein (Q8L5L5)	422
11S globulin seed storage protein (Fragment) (Q8GZP6)	93
2S albumin (A0A891LTQ0)	67
2S albumin (Q8H2B8)	39

The lower Mascot scores for these additional proteins, compared to the primary target, indicate their presence but not predominance in the sample. This aligns with the goal of the purification process, which aims to isolate the cashew vicilin N-terminal fragment with minimal contamination from other proteins.

Discussion

This study successfully purified and characterized vicilin N-terminal fragments from walnut, pecan, and cashew seeds. Despite the inherent challenges posed by their low solubility and abundance, the use of ammonium sulfate precipitation, anion exchange chromatography, and hydrophobic interaction chromatography enabled the effective isolation of these protein fragments.

The purification process revealed notable differences in the solubility and binding properties of vicilin N-terminal fragments across the three nut types. Walnut vicilin N-terminal fragments exhibited lower solubility in general buffers like PBS, necessitating the use of high-salt conditions to enhance protein recovery. The subsequent application of hydrophobic interaction chromatography was crucial for achieving higher purity by effectively separating the target protein from other co-precipitated components. The strong 45 kDa band observed in the Western blot may represent an uncleaved vicilin fragment or a partially cleaved storage protein. Notably, the presence and relative abundance of this band may vary across different walnut cultivars, potentially reflecting cultivar-specific differences in the cleavage ratio between intact vicilin and its fragments. Previous studies have shown that protein content in walnuts is influenced by cultivar variations (Tapia et al., 2013). Moreover, cultivation and storage conditions could influence the enzymatic or non-enzymatic cleavage processes, contributing to variability in protein profiles. While previous study has shown that protein content in walnuts changes during the ripening process, changes in protein composition remain unexplored (Amin et al., 2017). The antigenicity of the uncleaved vicilin and its fragments remains an important area for future investigation. Structural differences between intact vicilin and its fragments could alter IgE binding,

thereby influencing allergenic potential. Future studies should focus on comparing the allergenic properties of these components and examining the impact of environmental factors and post-harvest storage on protein cleavage patterns. These insights may provide a better understanding of the link between protein processing and allergenicity.

In contrast, pecan vicilin N-terminal fragments showed more favorable solubility and binding properties, allowing for a more straightforward purification process. The effectiveness of the Q Sepharose Fast Flow XK16 column in isolating the pecan vicilin N-terminal fragment suggests it has unique properties that facilitate easier separation, potentially reflecting differences in protein structure or composition compared to walnut and cashew proteins.

For the cashew vicilin N-terminal fragments, which were initially present at low concentrations, higher ammonium sulfate levels were employed to concentrate and purify them effectively. The need for additional concentration steps highlights the challenges associated with working with low-abundance proteins during extraction. Nevertheless, the successful purification of these fragments, confirmed through MS and subsequent analyses, demonstrates the effectiveness of the methodologies used.

Although the Q Sepharose Fast Flow XK16 column was used for both walnut and cashew N-terminal fragments, and the Butyl Toyopearl M-650 column was used for both pecan and cashew N-terminal fragments, differences in elution times were observed. These differences may be attributed to variations in sample preparation, such as the higher ammonium sulfate saturation required for cashew compared to walnut and pecan, as well as differences in physicochemical properties. On the

Q Sepharose Fast Flow XK16 column, the lower isoelectric point of the cashew vicilin N-terminal fragment resulted in a stronger negative charge under the experimental pH conditions, leading to longer retention and delayed elution compared to the pecan vicilin N-terminal fragment. On the Butyl Toyopearl M-650 column, the cashew vicilin N-terminal fragment exhibited stronger hydrophobicity at pH 7.5, resulting in greater binding affinity and delayed elution. In contrast, the walnut vicilin N-terminal fragment displayed weaker hydrophobicity, leading to earlier elution.

Overall, this study contributes valuable insights to food allergen research. The findings could inform the development of hypoallergenic food products and more accurate diagnostic tools, ultimately contributing to improved management of food allergies and improved public health outcomes.

Chapter 2

Expression, Purification, and Characterization of Recombinant Vicilin N-Terminal Fragments of Walnut, Pecan, and Cashew

Introduction

As of the publication of this thesis, eight allergens have been recognized by the WHO/IUIS Allergen Nomenclature Sub-committee in walnut (*Juglans regia*): Jug r 1 (2S albumin), Jug r 2 (vicilin), Jug r 3 (nsLTP1), Jug r 4 (11S globulin), Jug r 5 (PR-10), Jug r 6 (vicilin), Jug r 7 (profilin), and Jug r 8 (nsLTP2) (Dubielka et al., 2018; Pastorello et al., 2004; Pawel Dubielka, 2021; Serge Versteeg, 2017; S. Teuber, Dandekar, et al., 1998; S. Teuber, Jarvis, et al., 1999; Wallowitz et al., 2006; Wangorsch et al., 2017). The allergens in pecan (*Carya illinoensis*) include Car i 1 (2S albumin), Car i 2 (vicilin), and Car i 4 (11S globulin) (Sharma et al., 2011a, 2011b; Y. Zhang et al., 2016). The allergens in cashews (*Anacardium occidentale*) include Ana o 1 (vicilin), Ana o 2 (legumin), and Ana o 3 (2S albumin) (Wang et al., 2002; Wang et al., 2003; Robotham et al., 2005). Among these, research on storage proteins has been quite extensive, while studies on the N-terminal fragments of vicilins are relatively few, especially regarding their recombinant proteins and sensitization frequencies.

A study found that the sensitization distribution of walnut components in Europe is similar to that of many other plant-based foods (including other nuts), which may be related to cross-reactivity (Datema et al., 2015; Lyons et al., 2021). However, the cross-reactive properties between these three vicilin N-terminal proteins remain unclear. Therefore, it is necessary to study their IgE reactivity and

cross-reactivity at the molecular level to develop better diagnostic, therapeutic, and preventive measures.

In Chapter 1, the native vicilin N-terminal fragments from walnuts, pecans, and cashews were successfully extracted. However, their antigenicity and cross-reactivity among different nuts remain unexplored and require further research. To better understand the antigenicity of these N-terminal fragments, we conducted a study on the expression, purification, and characterization of recombinant vicilin N-terminal fragments. The *Pichia pastoris* expression system with His-tagged N-terminal proteins was constructed. *Pichia pastoris* can perform post-translational modifications of proteins, including the formation of disulfide bonds, which are crucial for the correct folding and function of complex proteins (Liu et al., 2022). Additionally, *Pichia pastoris* can grow at high densities and has a high protein expression yield, with the expressed proteins exhibiting higher stability and solubility (De Brabander et al., 2023).

The medium containing target proteins was first precipitated with 80% ammonium sulfate saturation. The resulting precipitation was then dialyzed and further purified using His-tag affinity column and gel filtration chromatography. The proteins were analyzed by SDS-PAGE and Western blot analysis. The sensitization frequency was also studied. In addition to examining their expression and purification, we analyzed the sensitization frequency of these vicilin N-terminal fragments using sera from 24 children diagnosed with walnut allergies. Conformational epitope predictions were performed to identify potential immunogenic regions. These combined approaches not only help to address the gap in vicilin N-terminal fragment research but also contribute to developing improved

diagnostic and therapeutic strategies for nut allergies, providing a scientific basis for future food safety and allergy prevention.

Materials and Methods

Materials

Materials used in this chapter included the FastGene Plasmid Mini Kit and FastGene Gel/PCR Extraction Kit (Nippon Genetics, Tokyo, Japan), 2× Gflex Buffer, Tks Gflex DNA Polymerase, Ligation Mix, 10× Ex Taq Buffer, 25 mM dNTPs, and Ex Taq DNA Polymerase (all from Takara Bio Inc., Shiga, Japan), Cut Smart, NEBuffer 3.1, and restriction enzymes EcoRI-HF, NotI-HF, and BstXI (all from New England Biolabs Inc., Ipswich, MA, USA), and Zeocin (100 mg/mL, InvivoGen Inc., San Diego, CA, USA).

All enzymes were used according to the manufacturer's instructions to ensure optimal performance in DNA manipulation, while all other reagents were purchased from Nacalai Tesque, Inc. (Kyoto, Japan) and Fujifilm Wako Pure Chemical Corporation (Osaka, Japan).

Construction of Expression Vectors

The DNA sequences encoding the vicilin N-terminal fragments from walnuts, pecans, and cashews were amplified by PCR (accession numbers: AF066055.1, XM018959147, EF689893, and AF395894). The sequences for nuts vicilin N-terminal fragments were identified from the full-length allergen sequences of Jug r 2.0101, Jug r 6.0101, Car i 2.0101, Ana o 1.0101. According to WHO/IUIS, the allergens were annotated as Jug r 2.0101 (1-158), Jug r 6.0101 (28-98), Car i 2.0101 (27-196), Ana o 1.0101 (27-150).

PCR amplification was carried out in a 100 μ L reaction volume containing 50 μ L of 2x Gflex buffer; 45 μ L of nuclease-free water; 1 μ L of template DNA; 1 μ L each of forward and reverse primers; and 2 μ L of Tks Gflex DNA polymerase. The reaction conditions were optimized to ensure high specificity and yield. The thermal cycling program included an initial denaturation step at 96 °C for 30 seconds, followed by 30 cycles of denaturation at 98 °C for 10 seconds, annealing at 55 °C for 30 seconds, and extension at 68 °C for 1 minute per kilobase of target DNA. After PCR, the amplified DNA fragments were purified using the FastGene Gel/PCR Extraction Kit and agarose gel electrophoresis was used to visualize the DNA fragments, stained with ethidium bromide for detection. The purified fragments were then treated with the Cut Smart, enzymes EcoRI and NotI at 37 °C for 3 hours, and subsequently purified.

DH5 α *E. coli* cells containing the pPICZ α plasmid were inoculated into 4 mL low-salt LB medium (10 g/L peptone, 5 g/L NaCl, 5 g/L yeast extract, pH 7.5) supplemented with 1:4000 diluted Zoecin and cultured overnight at 37 °C with shaking. After overnight incubation, the plasmid was extracted and purified. Plasmid purification was carried out using the FastGene Plasmid Mini Kit, following the manufacturer's protocol to ensure high purity and yield. It was then subjected to the same restriction enzyme treatment as the DNA amplification fragments. After purification, the sizes of the plasmid were confirmed by electrophoresis.

Table 3. Sequence of PCR Primers

	5'-3' Sequence
Jug r 2 N-Terminal	ATAGAATTTCAGAGGTAGAGATGACGACGACGAAG
Fragment	ATTGCGGCCGCCTAGTGATGGTGATGGTGATGTTCTCTACCCTGCTGCTCTTTG
Jug r 6 N-Terminal	ATAGAATTCATGCAAGACCCAGAGCTGAAGCAATG
Fragment	ATTGCGGCCGCCTAGTGATGGTGATGGTGATGCTCTTCCTCTCTTTTCGCCTCCAAGTTC
Car i 2 N-Terminal	ATAGAATTCTTGAAACTGAGGACTTGTCCAACG
Fragment	ATTGCGGCCGCCTAGTGATGGTGATGGTGATGACGACGCTCTCTTTCTCTTGTCTACG
Ana o 1 N-Terminal	ATAGAATTCAAGATCGACCCAGAGCTGAAGCAATG
Fragment	ATTGCGGCCGCCTAGTGATGGTGATGGTGATGGTCGTCTCTTCTTCAGCTTCATCC

***E. coli* Transformation**

DH5 α *E. coli* cells were initially inoculated into 4 mL of low-salt LB and incubated overnight at 37 °C with shaking. The following day, 500 μ L of the overnight culture was transferred into 20 mL of fresh low-salt LB and incubated at 37 °C with shaking until the cells reached the logarithmic phase (OD₆₀₀ > 0.5). The logarithmic phase was chosen to ensure maximum competency of the cells for transformation. The culture was then chilled on ice for 10 minutes, followed by centrifugation at 3000 rpm for 5 minutes at 4 °C to pellet the cells. The supernatant was discarded, and the cell pellet was resuspended in 4 mL of ice-cold 50 mM CaCl₂. The suspension was incubated on ice for 20 minutes, followed by centrifugation at 3000 rpm for 5 minutes at 4 °C. The supernatant was discarded, and the

cells were resuspended in 1 mL of ice-cold 50 mM CaCl₂. The cell suspension was incubated on ice for an additional 20 minutes to prepare competent cells. Calcium chloride treatment was employed to increase the permeability of the bacterial cell walls to DNA.

The ligation reaction was performed by combining 5 µL of DNA, 4 µL of the pPICZα, and 9 µL of ligation mix, then incubating the mixture at 16 °C for over 1 hour. Following the ligation, 200 µL of DH5α competent *E. coli* cells were added to 18 µL of the ligation mixture, and the new mixture was incubated on ice for 30 minutes. The cells were then heat-shocked at 42 °C for 90 seconds, and immediately placed on ice for 10 minutes. The heat-shock step was critical for facilitating the uptake of the plasmid DNA by the bacterial cells. Afterward, 900 µL of SOC medium was added, and the cells were incubated at 37 °C for over 1 hour. The culture was centrifuged at 5000 rpm for 3 minutes, the supernatant was removed, and the cells were resuspended in 100 µL of SOC medium. Finally, 100 µL of the transformed cell suspension was plated onto low-salt LB agar plate containing 1:4000 diluted Zeocin and incubated overnight at 37 °C.

Preliminary colony screening was performed to identify positive clones containing the correct DNA insert. This was achieved through colony PCR, conducted using Ex Taq polymerase in a reaction volume of 15 µL. The reaction mix included 1.5 µL of 10x Ex Taq buffer, 2 µL of dNTP mixture, 9.37 µL of nuclease-free water, 1 µL of forward and reverse primers (pPICZα forward and pPICZα reverse), and 0.13 µL of Ex Taq polymerase. The thermal cycling program consisted of an initial denaturation at 95 °C for 5 minutes, followed by 30 cycles of denaturation at 98 °C for 10 seconds, annealing at 55 °C for 30 seconds, and extension at 72 °C for 1 minute.

PCR products were analyzed by agarose gel electrophoresis to confirm the size of the inserted gene. Colonies producing PCR products of the expected size were considered positive and selected for further analysis. The positive colonies were then inoculated into 4 mL of low-salt LB medium containing Zeocin and cultured overnight at 37 °C with shaking.

Following overnight culture, the vicilin N-terminal fragment-pPICZ α plasmid extraction was performed using the FastGene Plasmid Mini Kit, and the purified plasmids were subjected to restriction enzyme digestion with BstXI at 37 °C for 3 hours to linearize the plasmid and verify the presence and correct orientation of the insert. The digested products were analyzed by agarose gel electrophoresis to confirm plasmid linearization. Additionally, plasmids from the positive clones were sequenced by Eurofins Genomics K.K. (Tokyo, Japan) to verify the accuracy of the insert sequences and their correct integration into the plasmid.

Yeast Transformation

A single colony of *Pichia* X33 was inoculated into 5 mL of YPD medium (22.2 g/L peptone, 11.1 g/L yeast extract, 20 g/L dextrose) and incubated overnight at 28 °C with shaking at 200 rpm. The following day, 100 μ L of the overnight culture was transferred into 100 mL of fresh YPD medium and incubated at 28 °C with shaking at 200 rpm until the OD₆₀₀ exceeded 1.3. The cells were harvested by centrifugation and washed twice with ice-cold deionized water. The cell pellet was resuspended in 4 mL of 1 M sorbitol, and the suspension was centrifuged at 1500 g for 5 minutes at 4 °C. The

supernatant was discarded, and the cells were resuspended in 250 μ L of 1 M sorbitol which was used to stabilize the cells during the electroporation process.

For the transformation, 80 μ L of the cell suspension was mixed with the vicilin N-terminal fragment-pPICZ α plasmid (linearized with BstXI) in a 1.5 mL tube and incubated on ice for 5 minutes. The mixture was then transferred to a pre-chilled 2 mm electroporation cuvette (CLP Cell Projects Ltd.) and subjected to electroporation using the Fungi (Pic) mode on a Bio-Rad electroporator. Electroporation was selected for its high efficiency in introducing DNA into yeast cells. Immediately after electroporation, 1 mL of 1 M sorbitol and 1 mL of YPD medium were added, and the cells were incubated at 28 °C with shaking at 200 rpm for 1.5 hours. The cells were then centrifuged at 1500 g for 5 minutes at 4 °C, the supernatant was removed, and the cell pellet was resuspended in 200 μ L of 1 M sorbitol. The entire volume was spread onto YPDS agar plate (22.2 g/L peptone, 11.1 g/L yeast extract, 20 g/L dextrose, 15 g/L agar, 1 M sorbitol) containing Zeocin and incubated at 28 °C for 3 days. Zeocin was used for selection of successful transformants, ensuring only yeast cells with the plasmid could grow. Colony PCR was performed on the resulting colonies, following the same procedure as described in the *E. coli* transformation section.

Expression of Recombinant Proteins

The single colony was inoculated into 40 mL of BMGY medium (1% yeast extract, 2% peptone, 100 mM potassium phosphate, pH 6.0, 1% glycerol, 1% YNB, 4 μ g/mL biotin) in a 300 mL baffled flask. Cultures were grown at 28 °C and 200 rpm until OD₆₀₀ exceeded 5.0. The cells were harvested

by centrifugation at 1,500 g for 5 minutes, washed twice with BMMY medium (1% yeast extract, 2% peptone, 100 mm potassium phosphate, pH 6.0, 1% methanol, 1% YNB, 4 µg/ml biotin) to remove any residual BMGY, and then resuspended in fresh BMMY medium. The suspension was transferred to a 2 L baffled flask. Baffled flasks were used to increase oxygen transfer, which is critical for optimal yeast growth and protein expression. The cultures were induced with 1% (v/v) methanol every 24 hours and grown at 28 °C and 200 rpm for 2 days. The yeast cultures were then centrifuged at 5000 rpm for 15 minutes at 4 °C, and the supernatant was collected.

Ammonium Sulfate Precipitation

Ammonium sulfate was added to the yeast culture supernatant to a final concentration of 80% saturation and stirred at 4 °C for 30 minutes. This step was performed to selectively precipitate the recombinant vicilin N-terminal fragments from the culture supernatant. The solution was centrifuged at 5000 rpm for 30 minutes at 4 °C, and the supernatant was discarded. The precipitate was dissolved in PBS and dialyzed in PBS at 4 °C with a dialysis clip, changing the buffer every 8 hours for 24 hours. Dialysis was used to remove excess ammonium sulfate and other small molecules, which could interfere with subsequent purification steps.

His-tag Affinity Purification

The dialyzed samples were purified using a HisTrap FF 5 mL column (Cytiva). His-tag affinity column was employed to specifically bind and purify the vicilin N-terminal fragments tagged with a

polyhistidine sequence. The column was first equilibrated by washing with 25 mL of pre-chilled PBS. The dialyzed samples were then loaded onto the column on ice using a syringe, maintaining a consistent flow rate. After loading, the column was washed with 25 mL of PBS containing 20 mM imidazole to remove unbound material. The bound proteins were subsequently eluted in three fractions, each with 8 mL of PBS containing 500 mM imidazole. The eluted fractions were analyzed by SDS-PAGE and Western blot to confirm the presence and purity of the target protein.

Gel Filtration Chromatography

The analyzed fractions containing the vicilin N-terminal fragments were subjected to gel filtration chromatography using a Superdex 75 pg 16/60 column (GE Healthcare) pre-equilibrated with PBS. Prior to chromatography, the samples were concentrated to less than 8 mL using a concentrator (Vivaspin, molecular weight cut-off: 5 kDa) at 6000 g and 4 °C. The concentrated samples were then filtered through a 0.20 µm syringe filter to remove any particulates. Gel filtration chromatography was used to separate proteins based on their size. The chromatography was run at a flow rate of 0.7 mL/min, and fractions were collected every 5 minutes, 3.5 mL per fraction, from minute 50 to minute 150. The fractions were identified by SDS-PAGE and Western blot analysis. These fractions were then collected and stored at 4 °C for further use.

SDS-PAGE

The experimental method used in this chapter is consistent with that described for SDS-PAGE in Chapter 1.

Western Blot

After incubation in PBSTM for 20 minutes at room temperature, the membrane was incubated for 60 minutes at room temperature with PBSTM containing mouse antisera against Anti-6xHis Mouse IgG1- κ (Nacalai Tesque), diluted 1:1000. The membrane was then washed three times, each for 10 minutes, with PBSTM. Following this, it was incubated with the secondary antibody Anti-Mouse IgG (H+L) from Promega, diluted 1:1000 in PBSTM, at room temperature for 60 minutes. The primary and secondary antibodies used throughout this chapter were consistent, ensuring the reproducibility and reliability of the results. The remaining experimental steps are identical to those in Chapter 1.

ELISA

Serum samples were collected from 24 pediatric patients who had clinically documented immediate reactions to walnuts. These patients were selected from the pediatric population receiving care at the National Hospital Organization Sagamihara Hospital in Japan. Ethical approval for participant recruitment was granted by the ethics committees of all involved hospitals. Informed consent was obtained from the parents or legal guardians of all participating children. This study was approved by the Ethical Board of Kyoto University under the approval ID: R3659.

Protein Structure Visualization

The protein structures were visualized using AlphaFold (Jumper et al., 2021).

Epitope Prediction Visualization

The epitope prediction and visualization were conducted by Epitope 3D (Moreira Da Silva et al., 2022).

Results

Purification of Recombinant Jug r 2 N-Terminal Fragment

After successful colony screening and large-scale cultivation, the culture medium was subjected to ammonium sulfate precipitation at a saturation level of 80%. This process effectively precipitated the target protein, thereby facilitating its separation from the culture medium and removal of other impurities while maintaining the protein's relative integrity. Subsequently, the protein was

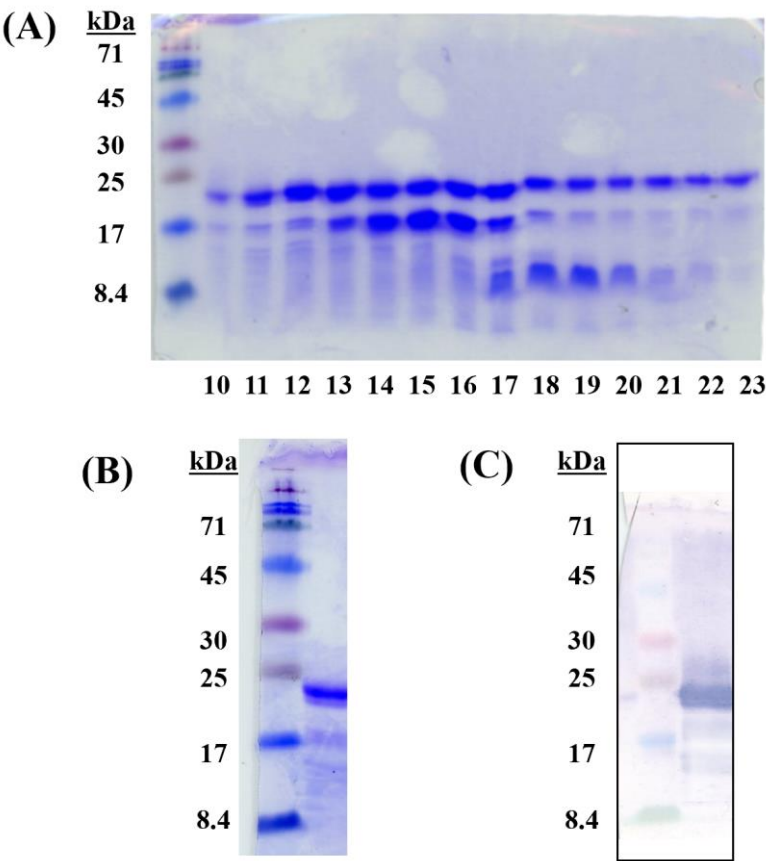


Figure 13. Recombinant Jug r 2 N-Terminal Fragment

(A) CBB Analysis of Jug r 2 N-Terminal Fragment Purified by Superdex 75 pg 16/60 Column

(B) CBB Analysis of Purified Jug r 2 N-Terminal Fragment

(C) Western Blot Analysis of Purified Jug r 2 N-Terminal Fragment

In panel (A), the samples were collected at different elution volumes, corresponding to fraction numbers 10 to 23. Primary antibody: Anti-6xHis mouse IgG1-κ.

successfully purified using His-tag affinity column and size-exclusion chromatography techniques. These steps enabled the isolation and purification of the Jug r 2 N-terminal fragment from the complex culture medium, which was subsequently confirmed and analyzed by SDS-PAGE and Western blot (Figure 13).

As shown in Figure 13A, SDS-PAGE analysis of the fractions collected from the Superdex 75 pg 16/60 column reveals distinct protein bands at different elution volumes (fractions 10-23), with a prominent band observed at 70.0 mL (fraction 11), which corresponds to the expected size of the Jug r 2 N-terminal fragment. The presence of these bands, ranging between 17-25 kDa, is consistent with the expected molecular weight of the Jug r 2 N-terminal fragment.

Western blot analysis (Figure 13C) further validated the presence of the Jug r 2 N-terminal fragment by its specific reaction with anti-6xHis antibody, demonstrating the correct expression and successful purification of the recombinant protein. The distinct bands observed on the Western blot confirm that the target protein was effectively isolated, although some minor bands might indicate incomplete purification or minor degradation.

Purification of Recombinant Jug r 6 N-Terminal Fragment

Among the allergens identified in walnuts, Jug r 2 was the first vicilin to be characterized. However, in 2018, Dubiel et al. identified an additional vicilin, Jug r 6 (Dubiel et al., 2018). Similar to Jug r 2, the construction of a plasmid containing Jug r 6 was successfully completed and subsequently introduced into yeast cells via electroporation. Following protein expression and

ammonium sulfate precipitation, the protein was successfully purified using His-tag affinity column and gel filtration chromatography. The results of these purification steps are presented in Figure 14.

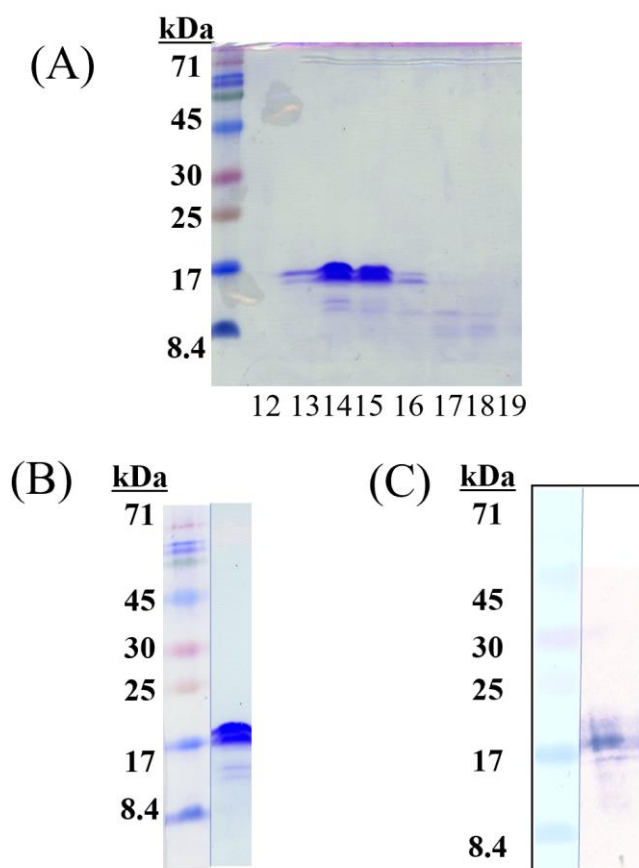


Figure 14. Recombinant Jug r 6 N-Terminal Fragment

(A) CBB Analysis of Jug r 6 N-Terminal Fragment Purified by Superdex 75 pg 16/60 Column

(B) CBB Analysis of Purified Jug r 6 N-Terminal Fragment

(C) Western Blot Analysis of Purified Jug r 6 N-Terminal Fragment

In panel (A), the samples were collected at different elution volumes, corresponding to fraction numbers 12 to 19. Primary antibody: Anti-6xHis mouse IgG1- κ .

During this step, the protein began to elute at approximately 77.0 mL (fraction 13), with the highest concentration observed at 84.0 mL (fraction 15), as shown by the clear single peak in the

chromatogram. The elution profile indicates that Jug r 6 N-terminal fragment was effectively isolated, with minimal contamination from other proteins (Figure 14A).

SDS-PAGE analysis of the collected fractions confirmed the presence of a distinct band corresponding to the expected molecular weight of the Jug r 6 N-terminal fragment, approximately 17 kDa (Figure 14B). This result suggests that the protein was successfully purified. Western blot analysis further validated the identity of the purified protein, showing a specific interaction with the anti-6xHis antibody. The development of a distinct band in the expected molecular weight range further confirms the successful expression and purification of the Jug r 6 N-terminal fragment (Figure 14C).

Purification of Recombinant Car i 2 N-Terminal Fragment

Figure 15 presents the analysis of the recombinant Car i 2 N-terminal fragment purified by Superdex 75 pg 16/60 column. SDS-PAGE analysis (Figure 15A) demonstrated multiple protein bands, with the major band appearing at approximately 25 kDa, which aligns with the expected molecular weight of the Car i 2 N-terminal fragment. This band emerged at an elution volume of 63.0 mL (fraction 9) and peaked at 73.5 mL (fraction 12), indicating successful purification. The additional smaller bands likely represent partial degradation or non-specific binding, but their low intensity indicates that the majority of the protein remained intact.

Figure 15C probed with anti-6xHis antibody confirmed the presence of Car i 2 N-terminal fragment around 25 kDa, further verifying successful expression and purification. The clarity of the

blot indicates a high level of purity, suitable for downstream applications. These results confirm the integrity and successful purification of the Car i 2 N-terminal fragment, with minimal degradation.

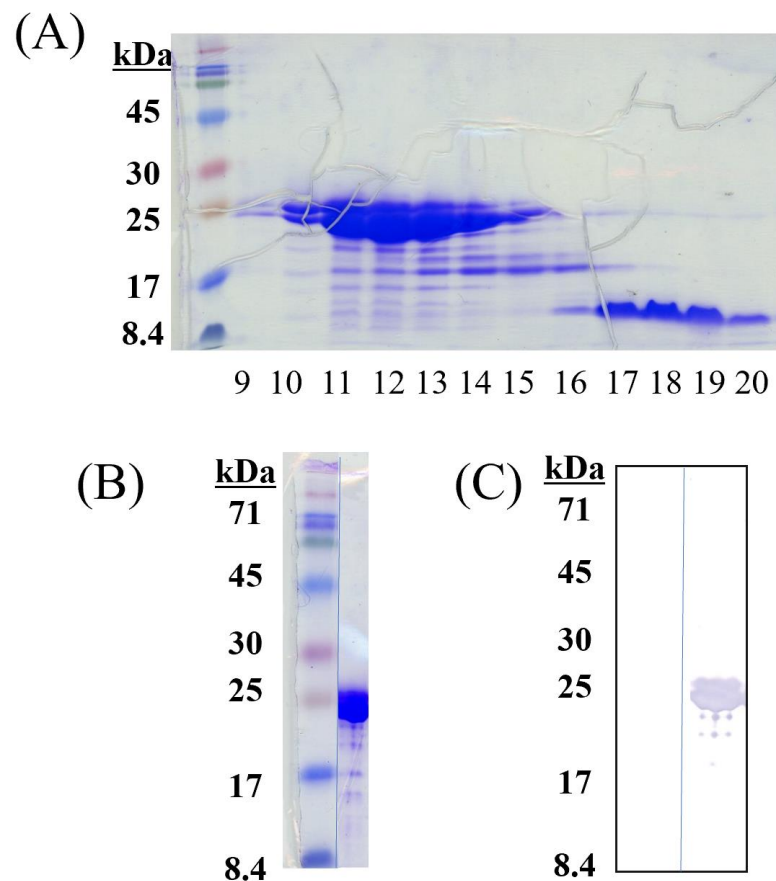


Figure 15. Recombinant Car i 2 Vicilin N-Terminal Fragment

(A) CBB Analysis of Car i 2 N-Terminal Fragment Purified by Superdex 75 pg 16/60 Column

(B) CBB Analysis of Purified Car i 2 N-Terminal Fragment

(C) Western Blot Analysis of Purified Car i 2 N-Terminal Fragment

In panel (A), the samples were collected at different elution volumes, corresponding to fraction numbers 9 to 20. Primary antibody: Anti-6xHis mouse IgG1-κ.

Purification of Recombinant Ana o 1 N-Terminal Fragment

Figure 16 presents the analysis of the recombinant Ana o 1 N-terminal fragment. SDS-PAGE analysis (Figure 16A) demonstrated a prominent band around 17 kDa, which corresponds to the

expected size of the Ana o 1 N-terminal fragment. This single, distinct band indicates successful purification, although the yield appears to be lower compared to other recombinant proteins. Notably, no additional protein bands were observed, suggesting a high degree of purity for this recombinant protein.

Figure 16C illustrates the results of the Western blot, where the purified Ana o 1 N-terminal fragment was probed with anti-6xHis antibody. The Western blot confirmed the presence of the target protein, revealing a specific band at approximately 17 kDa. This finding verifies the correct expression and successful purification of the recombinant Ana o 1 N-terminal fragment, demonstrating that the procedures used were effective in maintaining the integrity and purity of the protein.

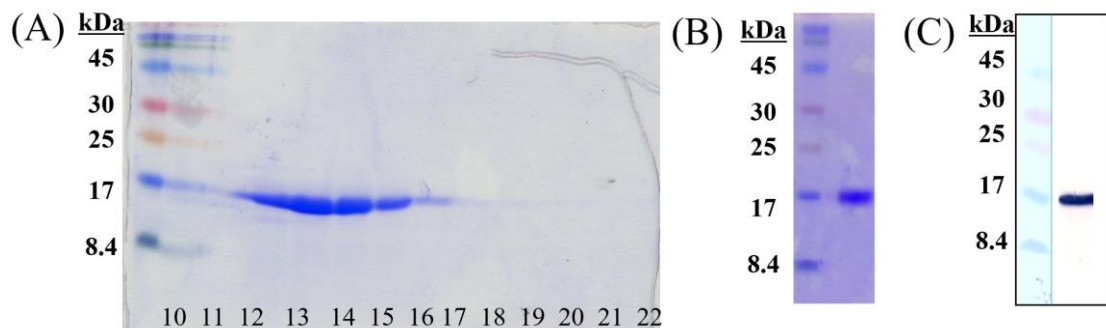


Figure 16. Recombinant Ana o 1 N-Terminal Fragment

(A) CBB Analysis of Recombinant Ana o 1 N-Terminal Fragment Purified by Superdex 75 pg 16/60 Column

(B) CBB Analysis of Purified Ana o 1 N-Terminal Fragment

(C) Western Blot Analysis of Purified Ana o 1 Vicilin N-Terminal Fragment

In panel (A), the samples were collected at different elution volumes, corresponding to fraction numbers 10 to 22. Primary antibody: Anti-6xHis mouse IgG1-κ.

Overall, these results validate the effectiveness of the expression and purification protocols used for the recombinant vicilin N-terminal fragments from Jug r 2, Jug r 6, Car i 2, and Ana o 1. The high specificity of the Western blots and the consistent molecular weights observed in SDS-PAGE confirm that the recombinant proteins were successfully isolated and are ready for further structural and functional analyses.

ELISA

The IgE reactivity of serum from 24 pediatric patients allergic to walnuts was assessed against vicilin N-terminal fragments from nuts. These fragments included the purified recombinant Jug r 2 N-terminal fragment, Jug r 6 N-terminal fragment, Car i 2 N-terminal fragment, and Ana o 1 N-terminal fragment. ELISA was employed to analyze each patient's serum sample, and the intensity of the reactions was plotted in a scatter diagram as shown in Figure 17.

Each point on the diagram represents the absorbance of a serum sample's response to the respective allergen, with a horizontal dashed line indicating the threshold for a positive IgE reaction. Among the tested fragments, both the Jug r 2 N-terminal fragment and Jug r 6 N-terminal fragment exhibited IgE reactivity exceeding the threshold in 70.8% of the samples, followed by the Car i 2 N-terminal fragment, which showed a positive reaction rate of 66.7%. The Ana o 1 N-terminal fragment showed the lowest positive reaction rate at only 12.5%, indicating a relatively lower potential for causing cross-reactivity in patients allergic to walnuts. Notably, sera from five participants (P1, P6, P8, P11 and P13) exhibited no positive response to any of the tested allergens, while sera from seven

participants (P1, P6, P8, P11, P12, P13, and P17) exhibited no positive response to walnut vicilin N-terminal fragments, suggesting that these individuals might not be sensitized to these specific proteins, and their walnut allergy may be attributed to other proteins.

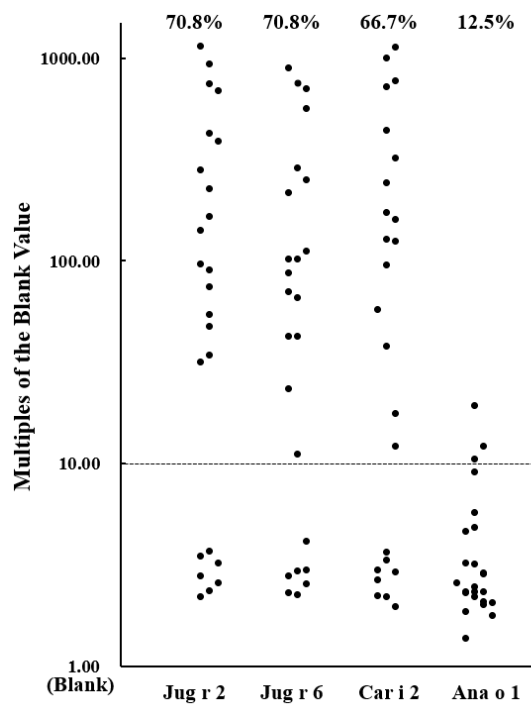


Figure 17. IgE Reactivity of Sera to Vicilin N-Terminal Fragments from Nuts

The y-axis represents the multiples of the blank value, and the horizontal dashed line represents the threshold for a positive reaction (10 times the blank value). Each dot represents the IgE reactivity level for an individual patient’s serum. Sera from patients with reactivity levels above the threshold are considered positive for specific IgE against the respective allergen.

The Venn diagram further categorizes the IgE reactivity to these vicilin N-terminal fragments (Figure 18). The Ana o 1 N-terminal fragment elicited responses in only three serum samples (P3, P7 and P22), all of which also reacted to other tested allergens, suggesting that the Ana o 1 N-terminal

fragment may share immunological characteristics with other vicilin N-terminal fragments but has a weaker capacity to trigger an IgE response.

Out of the 24 serum samples, 19 exhibited positive reactivity to at least one of the tested vicilin N-terminal fragments. Notably, 17 out of 19 serum samples (89.5%) reacted to the Jug r 2 N-terminal fragment, and all 17 of these samples also showed reactivity to the Jug r 6 N-terminal fragment, suggesting a high degree of cross-reactivity between these two walnut-derived fragments. Furthermore, 16 out of these 19 serum samples (84.2%) also reacted to the pecan-derived Car i 2 N-terminal fragment.

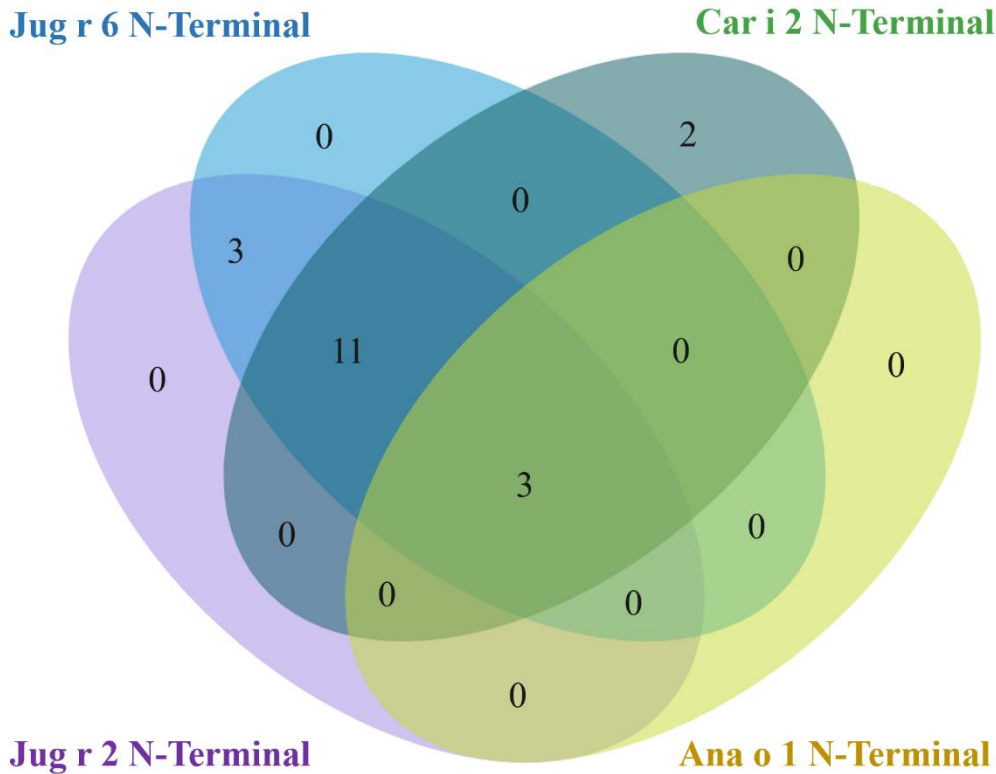


Figure 18. Classification of IgE Reactivity to Nuts Vicilin N-Terminal Fragments in Sera

The Venn diagram shows the overlap in IgE reactivity to the different vicilin N-terminal fragments (Jug r 2, Jug r 6, Car i 2, and Ana o 1) in sera from 24 patients with walnut allergy. The numbers represent the count of sera reactive to specific combinations of these fragments.

Additionally, sera from two patients (P12 and P17) uniquely reacted to the Car i 2 N-terminal fragment, without any response to other nut-derived vicilin N-terminal fragments. In contrast, sera from three patients (P15, P16, and P19) reacted exclusively to the walnut vicilin N-terminal fragments, with no response to the Car i 2 N-terminal fragment. These observations resulted in a different positivity rate for the pecan-derived Car i 2 N-terminal fragment compared to the walnut-derived fragments, suggesting immunological epitope differences between them. To further explore this unique cross-reactivity, an amino acid sequence alignment analysis of these proteins was conducted.

Table 4. Sequence Identity Among Nuts Vicilin N-Terminal Fragments

Protein	Jug r 2	Jug r 6	Car i 2	Ana o 1
Jug r 2	1.00	0.21	0.33	0.23
Jug r 6		1.00	0.27	0.66
Car i 2			1.00	0.22
Ana o 1				1.00

0.00~0.25	0.25~0.50	0.50~0.75	0.75~1.00
-----------	-----------	-----------	-----------

"Jug r 2", "Jug r 6", "Car i 2", and "Ana o 1" refer to the vicilin N-terminal fragments of Jug r 2, Jug r 6, Car i 2, and Ana o 1, respectively.

Generally, it is anticipated that proteins with a sequence identity exceeding 70% may demonstrate cross-reactivity, whereas such reactivity is uncommon when the sequence identity falls below 50% (Aalberse, 2000). In the comparative analysis of amino acid sequences from four nuts vicilin N-

terminal fragments, the results, presented in Table 4, reveal significantly lower sequence identities than this threshold. These findings suggest that sequence identity alone cannot sufficiently explain the observed cross-reactivity among these allergens.

Subsequently, the conformation of vicilin N-terminal fragments from different sources was analyzed. As shown in Table 5, sequence alignment highlights conserved cysteine residues C, underlined in the sequences, which demonstrate a high similarity in the critical CXXXCX₁₀₋₁₂CXXXC motifs despite notable differences in overall sequence identity. These conserved motifs stabilize α -

Table 5. Amino Acid Sequence Alignment of Vicilin N-Terminal Fragments

	20	40	60
Jug r 2	RGRDD-DDEE NPR-DP---R	EQYRQCQEYC RRQGQGQRQ-	QQQCQIRCEE RLEEDQRSQE
Jug r 6	MQ-----	----DP---- -ELKQCKHQC	RHQRQFDEQE KEHCQRSCDE YHIE----KK
Car i 2	LETEDLSNEL NPHHPESH	WEFQQCQERC QHEER-GQRQ	AQQCQRRCEE QLRE-----R
Ana o 1	KI-----	----DP---- -ELKQCKHQC	KVQRQYDEQQ KEQCVKECEK YYKE----KK
	80	100	120
Jug r 2	ERERRRRGRD- -----VDD	QNPRDP-EQR YEQCQQQCER	QRRGQEQTLC RRRCE----Q
Jug r 6	ARERA-----	-----	-----E
Car i 2	EREREREI- -----VD-	--PREP-RKQ YEQCRETCEK	QDPRQ-QPQC ERRCERQFQE
Ana o 1	GREEREHEEEE EEWGTGGVDE	PSTHEPAEKH LSQCMRQCER	QEGGQQKQLC RFRCQ----E
	140	160	180
Jug r 2	RRQQEERERQ RGRDR-----	-QDPQQQYHR CQRRCQIQEQ	SPERQRQCQQ RCERQYKEQQ
Jug r 6	RRRSEEGS-- -SREEG----	-----	-----YEEE
Car i 2	QQERERRERR RGRDDDDKEN	PRDPREQYRQ CEEHCRRQGQ	GQRQQQQCQS RCEERFEEEQ
Ana o 1	RYKKERGQHN YKREDD----	-----	-----EDED
Jug r 2	GRE-----	158	
Jug r 6	ELGGEREEE	71	
Car i 2	RRQEERERR	170	
Ana o 1	EDEAEDEDE	124	

"Jug r 2", "Jug r 6", "Car i 2", and "Ana o 1" refer to the vicilin N-terminal fragments of Jug r 2, Jug r 6, Car i 2, and Ana o 1, respectively. The cysteine residues within the CXXXCX₁₀₋₁₂CXXXC motif are represented as underlined C.

hairpin secondary structures through disulfide bonds formed between cysteine residues. The α -hairpin structures are particularly significant, as they may contribute to the allergenic cross-reactivity observed among tree nuts by exposing similar epitopes to the immune system (Foo et al., 2022).

Conformational epitope prediction was conducted by combining spatial conformations predicted by AlphaFold with subsequent analysis using Epitope3D (Jumper et al., 2021; Moreira Da Silva et al., 2022). AlphaFold generates multiple models with varying confidence scores (pLDDT values) and structural features. Screening based on overall pLDDT scores identified the model with the highest average pLDDT score, which was used as input for Epitope3D, a machine-learning-based tool specifically designed to predict conformational epitopes (Moreira Da Silva et al., 2022).

To ensure reliability and representativeness of the selected AlphaFold model, all generated models were compared with a focus on the spatial arrangement of key structural elements, particularly the hairpin structures. The analysis demonstrated a high degree of consistency in the number and overall configurations of hairpin structures across all models, confirming the robustness of these structural predictions. The confidence levels for the hairpin structures were consistently high (Figure 19), indicating their stability and reliability. However, the linkers connecting the hairpins exhibited relatively low confidence levels, with notable variations in their conformations across different models. This suggests that while the hairpin structures themselves are stable, the inter-hairpin linkers show significant flexibility, which may influence the protein's structural dynamics and functional adaptability. These findings provide a reliable basis for subsequent epitope predictions and functional analyses.

Notably, there is a variation in the number of hairpin structures among the different fragments. As demonstrated in Figure 19, three-dimensional structural simulations conducted using AlphaFold indicate that the Jug r 2 N-terminal fragment and the Car i 2 N-terminal fragment each contain three hairpin structures, whereas the Ana o 1 N-terminal fragment comprises two, and the Jug r 6 N-terminal fragment possesses only a single hairpin structure. These findings suggest that while the conserved CXXXCX₁₀₋₁₂CXXXC motifs and hairpin structures are critical for stability and allergenicity, their variation across tree nut vicilin N-terminal fragments may further influence the degree of cross-reactivity.

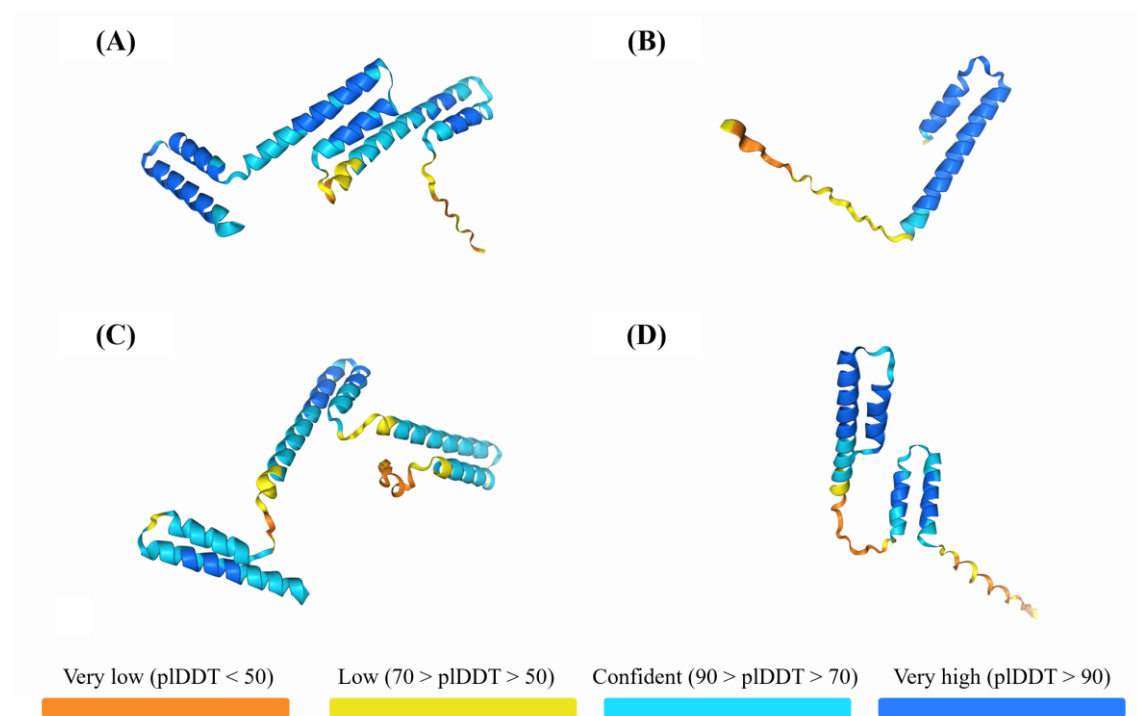


Figure 19. Three-Dimensional Models of Nuts Vicilin N-Terminal Fragments

Panels (A), (B), (C), and (D) present the three-dimensional structures of the Jug r 2 N-terminal fragment, Jug r 6 N-terminal fragment, Car i 2 N-terminal fragment, and Ana o 1 N-terminal fragment, respectively.

B-cell epitopes can be classified into linear and conformational epitopes, with approximately 90% of B-cell epitopes being conformational (Van Regenmortel, 2009). Therefore, predicting the conformational epitopes of these four N-terminal fragments is crucial to understanding the underlying causes of cross-reactivity in allergies. To investigate the potential mechanisms of cross-reactivity and explain the differences in positivity rates among four vicilin N-terminal fragments, conformational B-cell epitopes were predicted using Epitope 3D (Figure 20). This approach offered insights into potential reactive sites, with simulations revealing active regions on each N-terminal fragments, thereby aiding in the elucidation of cross-reactivity among these proteins.

The regions highlighted in red are identified as the predicted conformational B-cell epitopes, suggesting areas with potential antigenic activity within each protein structure. For ease of reference, these regions have been sequentially named based on their appearance within the sequence of each vicilin N-terminal fragments. Specifically, the epitopes have been designated as J2N-1, J2N-2, J2N-3 for Jug r 2 N-terminal fragment; J6N-1, J6N-2 for Jug r 6 N-terminal fragment; C2N-1, C2N-2, C2N-3 for Car i 2 N-terminal fragment; and A1N-1, A1N-2 for Ana o 1 N-terminal fragment. This nomenclature enables clear communication of specific loci of interest across the different protein models.

As shown in Figure 20, both of Jug r 2 and Car i 2 N-terminal fragments feature three red regions each, positioned within three distinct hairpin structures. The smaller regions J2N-2, J2N-3, C2N-1, and C2N-2 are located internally within the hairpin structures, indicating reduced accessibility to antigens. In contrast, the larger areas J2N-1 and C2N-3, composed of relatively contiguous amino

acids, are positioned on the external arches of the hairpins, enhancing their likelihood of antigenic interaction. Interestingly, a detailed comparison of the amino acid sequences in and around J2N-1 and C2N-3 reveals that despite their presence in different hairpins, they share the identical sequence CRRQGQGQRQQQCQ (Table 5), situated between two cysteine residues. This structural similarity suggests that these regions could influence their antigenic potential and contribute to cross-reactivity among these proteins.

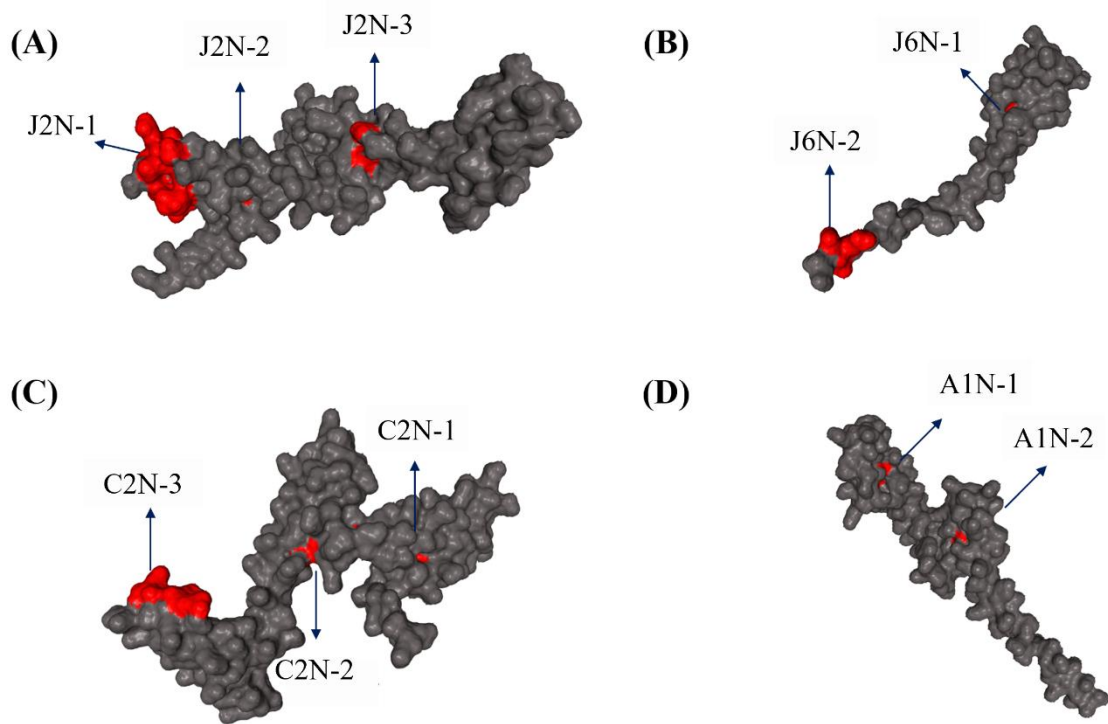


Figure 20. Conformational Epitope Predictions for Nuts Vicilin N-terminal Fragments

Regions highlighted in red indicate the predicted conformational B-cell epitopes sites.

Each panel shows the predicted epitopes for different N-terminal fragments: (A) Jug r 2 (J2N-1, J2N-2, J2N-3), (B) Jug r 6 (J6N-1, J6N-2), (C) Car i 2 (C2N-1, C2N-2, C2N-3), (D) Ana o 1 (A1N-1, A1N-2).

J6N-2's location is unique in that it is not situated within a hairpin structure but rather on a more flexible chain, lacking structural stabilization from disulfide bonds. Compared to other epitope locations within internal hairpin structures, such as J6N-1, J2N-2, J2N-3, C2N-1, and C2N-2, this positioning may increase its exposure and accessibility, potentially enhancing its ability to interact with antibodies. While J6N-1's location within a hairpin structure may limit its exposure, the distinctive positioning of J6N-2 outside of conventional hairpin structure likely contributes to its heightened reactivity in serological assays. Compared to the more complex structures of Jug r 2 N-terminal fragment and Car i 2 N-terminal fragment, the simpler structural configuration of Jug r 6 N-terminal fragment may result in a distinct pattern of immune recognition. Although flexible chains located outside the hairpin structure are also observed in other allergens, the terminal regions in these cases are relatively closer to the nearest cysteine residues that form the critical disulfide bonds. This proximity may limit their flexibility compared to the extended chain observed in Jug r 6 N-terminal fragment. The higher degree of freedom in Jug r 6's flexible chain could contribute to its unique structural characteristics.

At the same time, it should be noted that the extended J6N-2 regions in the AlphaFold predictions exhibit relatively low pLDDT scores, indicating low confidence. Moreover, both experimentally derived protein structures and AI-predicted models have inherent limitations (Acharya & Lloyd, 2005; Moore et al., 2022; Shao et al., 2022; Terwilliger et al., 2024). Proteins are inherently flexible and dynamic, with their conformational states influenced by factors such as temperature, solvent conditions, and interactions with ligands or other proteins (Fraser et al., 2011). High-resolution crystal

structures often capture dominant conformations under specific experimental conditions but may deviate when environmental conditions change (Acharya & Lloyd, 2005). Similarly, while AI-based models are highly accurate in predicting three-dimensional structures, they still face challenges in accounting for varying environments, the dynamic interplay of protein–protein interactions, and multiple conformational states (Baek et al., 2021; Jumper et al., 2021; Terwilliger et al., 2024).

Ana o 1 N-terminal fragment comprises only two hairpin structures, which likely contribute to its unique antigenic characteristics. Unlike J6N-2, which is external to hairpin structure and thereby more accessible, the epitope sites A1N-1 and A1N-2 within Ana o 1 N-terminal fragment are embedded within these hairpin structures. This placement restricts their exposure compared to epitopes like J2N-3 and C2N-1, which are external and present continuous epitope regions. Similarly, akin to epitopes J2N-1, J2N-2, C2N-2, and C2N-3, which are positioned within the internal spaces of hairpins, A1N-1 and A1N-2 have limited accessibility to the immune system, thereby potentially diminishing antibody binding interactions, contributing to the lower immunoreactivity observed. Additionally, the fact that the sera were derived from walnut-allergic patients, and that cashew and walnut belong to different genera, along with the structural arrangement of the epitopes, provides a reasonable explanation for the finding that only 12.5% of patients showed a positive reaction to the Ana o 1 N-terminal fragment (Borres et al., 2022).

Discussion

This chapter focused on the expression, purification, and allergenicity analysis of recombinant vicilin N-terminal fragments from walnut (Jug r 2, Jug r 6), pecan (Car i 2), and cashew (Ana o 1) using the *Pichia pastoris* system. This system efficiently facilitated the expression of recombinant proteins and ensured their proper post-translational modifications. The formation of stable α -hairpin structures through the CXXXXCX₁₀₋₁₂CXXXC motif was critical for the proper folding and functional expression of these complex allergenic proteins. Using His-tag affinity column followed by gel filtration chromatography ensured high purity and integrity of the proteins, essential for subsequent allergenicity testing.

These recombinant proteins were used to assess their IgE reactivity with sera from 24 walnut-allergic patients, revealing significant cross-reactivity and variability in allergenic potential among the different proteins. Specifically, Jug r 2 and Jug r 6 N-terminal fragments demonstrated IgE reactivity in 70.8% of the samples, while the Car i 2 N-terminal fragment had the second positivity rate at 66.7%. In contrast, the Ana o 1 N-terminal fragment showed a much lower positivity rate of 12.5%, indicating its lower allergenic potential. These findings suggest that the immunological differences among vicilin N-terminal fragments might be related to their structural features, may be linked to their structural features and cross-reactivity mechanisms.

To explore this cross-reactivity, further research on the amino acid sequences and structures was conducted. A sequence comparison revealed that the sequence identity among these fragments was below the commonly accepted 70% threshold for cross-reactivity (Van Regenmortel, 2009). The lower

sequence identity suggests that IgE recognition is mediated primarily through conformational rather than linear epitopes (Foo et al., 2023). Conformational epitope prediction analysis showed that all fragments shared a conserved sequence motif, and each vicilin N-terminal fragment was capable of forming at least one disulfide-bonded hairpin structure due to this motif, supporting the occurrence of cross-reactivity. This finding aligns with previous studies, which indicate that this conserved disulfide-bonded hairpin motif, common across various plant species, may facilitate cross-reactivity between species (Foo et al., 2022, 2023; Oparin et al., 2012; Payne et al., 2020; J. Zhang et al., 2019). When exploring differences in positivity rates, structural epitope predictions showed a high degree of similarity between J2N-3 and C2N-1, which may be a key factor in triggering cross-reactivity. Additionally, the unique positioning of J6N-2 outside the hairpin structure, with greater surface accessibility, may contribute to its antigenicity, warranting further experimental investigation to confirm its role as an antigenic epitope.

In conclusion, this chapter deepens our understanding of the structural and immunogenic characteristics of vicilin N-terminal fragments from different nuts and elucidates the complex interactions between protein structure and allergenic potential. The insights gained in this research pave the way for innovative allergy management and food safety approaches for tree nut allergies. Future research will continue to explore the immunogenicity of vicilin N-terminal fragments, particularly their roles in clinical allergic reactions, and how structural analysis can predict and mitigate allergic responses, providing essential scientific support for developing new diagnostic and therapeutic strategies.

Chapter 3

Expression, Purification, and Characterization of Recombinant PR-10 and Profilin in Walnut

Introduction

While the major allergen in walnuts, such as Jug r 1 (2S albumin), Jug r 2 (7S globulin), and Jug r 4 (2S globulin), are significant allergens, they may not be the sole triggers of sensitization (S. Teuber, Dandekar, et al., 1998; S. Teuber, Jarvis, et al., 1999; Wallowitz et al., 2006; Wangorsch et al., 2017). Other proteins, including profilin and PR-10, associated with PFAS, may also contribute to allergic responses in susceptible individuals (Carlson & Coop, 2019). PFAS refers to IgE-mediated allergic reactions occurring in individuals allergic to pollen when they consume certain plant-based foods (Valenta & Kraft, 1996). This reaction is triggered by the cross-reactivity of IgE antibodies against certain pollen proteins that have similar proteins present in foods (Poncet et al., 2020; Waserman & Watson, 2011). Symptoms of PFAS typically occur shortly after consumption, with 97% of patients experiencing mild symptoms such as tingling and itching of the mouth and throat, while 3% of patients may experience systemic symptoms (Ma et al., 2003).

Food allergens related to PFAS include PR-10, profilin, TLP, nsLTP (Non-specific Lipid Transfer Protein), and GRP (Gibberellin-Regulated Protein) (Krikeerati et al., 2023). Among them, PR-10 and profilin are the allergens most often associated with PFAS (Maruyama, 2021; Poncet et al., 2020). Existing research on PFAS food allergens primarily focuses on fruits (such as pineapple, kiwi, apple, peach, plum, apricot, orange, watermelon, banana, etc.), vegetables (such as carrot, celery, potato, pepper), and tree nuts (almond, hazelnut), with relatively few studies on walnuts, which are

the world's most consumed tree nut (Carlson & Coop, 2019; INC, 2023; Jeon, 2020). To address this gap, this chapter examines PR-10 (Jug r 5) and profilin (Jug r 7) in walnuts to enrich the study content.

Jug r 5 belongs to the Bet v 1/PR-10 family. Bet v 1 from birch pollen is a major antigen that causes PFAS through cross-reactivity. Its structure consists of seven antiparallel β -strands and three α -helices, surrounding a large hydrophobic cavity (Fernandes et al., 2013). Jug r 7 belongs to profilin, a small cytoplasmic molecule (12-15 kDa) with highly conserved sequences, showing greater than 75% sequence identity even in distantly related organisms (Hauser et al., 2010). Its structure primarily consists of seven β -strands and three to four α -helices (Kapingidza et al., 2019). Due to its involvement in numerous key cellular processes, profilin is present in most eukaryotic cells and is considered a pan-allergen (Costa et al., 2014). A survey conducted in Europe found that Jug r 5 and Jug r 7 had the highest overall sensitization rates among walnut allergens, with Jug r 5 showing higher sensitization rates in regions rich in birch trees (such as Central and Northern Europe), and Jug r 7 having a more uniform sensitization distribution throughout Europe, likely related to pollen exposure (Lyons et al., 2021).

In the context of peanuts, research has shown that individuals sensitized solely to PR-10 (Ara h 8), as opposed to Ara h 1, Ara h 2, or Ara h 3, may still experience allergic symptoms upon peanut ingestion (Asarnoj et al., 2010; Glaumann et al., 2015). Isolated sensitization to Ara h 8 generally suggests a degree of peanut tolerance, which may allow for the cautious introduction of peanuts at home in sensitized children (Asarnoj et al., 2012). In contrast, similar studies on walnuts are currently

lacking. Therefore, there is an urgent need for more refined laboratory diagnostics to differentiate true walnut allergens, preventing misdiagnosis and unnecessary dietary restrictions.

To further investigate these allergens, an *E. coli* recombinant expression system was constructed using pColdGST (Jug r 5) and pGEX (Jug r 7) plasmids, incorporating GST and His-tags. After overnight cultivation, cells were harvested and lysed via sonication, followed by GST and His-tag affinity purification. Jug r 5 and Jug r 7 were further purified using HPLC. Additionally, the storage protein 11S globulin, 7S globulin and 2S albumin were extracted from natural walnut seeds. The successful purification of the target proteins was confirmed through SDS-PAGE and Western blot analysis, and the allergenicities of Jug r 5, Jug r 7, and storage proteins were studied using sera from 24 pediatric patients diagnosed with walnut allergies.

This chapter aims to provide valuable insights into the immunological profiles of walnut allergens, focusing on the PR-10 and profilin families. By addressing existing knowledge gaps, this study seeks to enhance the understanding of PFAS-associated allergens and cross-reactivity patterns, ultimately contributing to more accurate diagnoses and improved management strategies for walnut allergies.

Materials and Methods

Materials

The restriction enzymes used included BamHI-HF, NotI-HF, and XbaI-HF, all sourced from New England Biolabs Inc. (Ipswich, MA, USA). HRV 3C Protease (1 U/ μ L; Takara Bio Inc., Shiga, Japan) was employed for the cleavage of fusion proteins. Additionally, reduced glutathione (Fujifilm Wako Pure Chemical Corporation, Osaka, Japan) was used for the elution of GST-tagged proteins, due to its strong affinity for the GST tag, ensuring effective protein recovery. Isopropyl β -D-1-thiogalactopyranoside (IPTG) (Nacalai Tesque, Inc., Kyoto, Japan) was utilized to induce protein expression, and carbenicillin sodium salt (Fujifilm Wako Pure Chemical Corporation, Osaka, Japan) was used as an antibiotic for selection in bacterial cultures. Other reagents were identical to those described in Chapter 2.

Expression and Purification of Recombinant Jug r 5 and Jug r 7

Construction of Expression Vectors

Based on the sequences available from the National Center for Biotechnology Information (accession numbers KX034087.1 and MG366484), gene-specific primers were synthesized for Jug r 5 and Jug r 7. The primers used for PCR amplification are summarized in Table 6.

PCR amplification was performed following the same method described in Chapter 2, with identical thermal cycling conditions and purification steps.

Table 6. Sequence of Recombinant Jug r 5 and Jug r 7 PCR Primers

	5'-3' Sequence
Jug r 5	TTTGGATCCGGTGTTTTCACCTACGAGACTGAATC AAATCTAGACTAGTTGTAAGCATCTGGATGAGCCAAC
Jug r 7	TTTGGATCCTCCTGGCAGACTTACGTTGACGAGC AAAGCGGCCCGCCTAATGGTGATGGTGATGGTGCAAAC CTTGCTCGACCAAGTAGTCAC

For both Jug r 5 and Jug r 7 constructs, the DNA fragments were inserted into the appropriate sites of expression vectors: Jug r 5 into the BamHI-XbaI site of the pColdGST vector, and Jug r 7 into the BamHI-NotI site of the pGEX vector. For each construct, a DH5 α *E. coli* colony containing the corresponding plasmid (pColdGST or pGEX) was inoculated into 4 mL of LB medium (10 g/L peptone, 10 g/L NaCl, 5 g/L yeast extract, pH 7.5) supplemented with 1:1000 diluted carbenicillin and cultured overnight at 37 °C with shaking. After incubation, plasmids were extracted and purified using the FastGene Plasmid Mini Kit, following the manufacturer's protocol. The vectors (pColdGST for Jug r 5 and pGEX for Jug r 7) and their respective DNA fragments were digested with restriction enzymes (BamHI-XbaI for Jug r 5 and BamHI-NotI for Jug r 7) at 37 °C for 3 hours to generate compatible ends for ligation. The digested products were purified by using the FastGene Gel/PCR Extraction Kit. Ligation reactions for both constructs were performed at 16 °C for over 1 hour.

Transformation of DH5 α *E. coli* cells was carried out using the heat shock method, as described in Chapter 2, with carbenicillin used for selection instead of Zeocin. After heat shock at 42 °C for 90

seconds, the transformed cells were plated onto LB agar plates containing 1:1000 diluted carbenicillin and incubated overnight at 37 °C. For each construct, plasmids from positive clones were sequenced by Eurofins Genomics K.K. (Tokyo, Japan) to confirm the accuracy of the insert sequences and their proper integration into the plasmid vectors.

***E. coli* Transformation for Jug r 5 and Jug r 7**

For Jug r 5, the confirmed DH5 α *E. coli* colony containing the Jug r 5-pColdGST plasmid was inoculated into 4 mL of LB medium supplemented with carbenicillin and cultured overnight to ensure sufficient plasmid yield. Simultaneously, BL21 cells were cultured in LB medium. The next day, 500 μ L of the BL21 culture was transferred into 20 mL of fresh LB medium. Once the cells reached logarithmic phase (OD₆₀₀ > 0.5), they were cooled on ice for transformation. The Jug r 5-pColdGST plasmid was introduced into BL21 cells using the heat shock method, with selection performed on LB agar plates containing carbenicillin.

For Jug r 7, the DH5 α *E. coli* colony harboring the Jug r 7-pGEX plasmid was similarly inoculated into 4 mL of LB medium with carbenicillin and cultured overnight, followed by plasmid extraction and purification. SHuffle T7 cells were cultured overnight in LB medium and then subcultured in fresh medium. After the cells reached the logarithmic phase, they were chilled on ice to prepare for transformation. The Jug r 7-pGEX plasmid was introduced into SHuffle T7 cells via the heat shock method, with selection on LB agar plates containing carbenicillin.

Expression of Recombinant Jug r 5 and Jug r 7

A single, transformed *E. coli* BL21 colony harboring the Jug r 5-pColdGST plasmid was inoculated into 20 mL of LB medium containing carbenicillin and cultured overnight at 37 °C with shaking. The culture was then transferred to 500 mL of LB medium containing carbenicillin and grown at 37 °C, 200 rpm until the OD600 reached 0.4-0.6. The culture was cooled on ice for 15 minutes to minimize background expression, ensuring controlled induction. Protein expression was induced with 0.2 mM IPTG at 15 °C with shaking at 100 rpm overnight. Lowering the temperature to 15 °C and reducing the shaking speed were strategic decisions to enhance the solubility of Jug r 5, thereby reducing the formation of inclusion bodies. After induction, the cells were harvested by centrifugation at 5000 rpm for 10 minutes at 4 °C.

For Jug r 7, a transformed *E. coli* SHuffle T7 colony harboring the Jug r 7-pGEX plasmid was similarly inoculated into 20 mL of LB medium containing carbenicillin and cultured overnight at 30 °C with shaking. The culture was then transferred to 500 mL of LB medium containing carbenicillin and grown at 30 °C, 200 rpm until the OD600 reached 0.4-0.6. The culture was cooled, and protein expression was induced with 0.2 mM IPTG at 18 °C with shaking at 100 rpm overnight. After induction, the cells were harvested by centrifugation at 5000 rpm for 10 minutes at 4 °C.

Cell Lysis of Recombinant Jug r 5 and Jug r 7

The harvested cells were resuspended in 80 mL of ice-cold PBS buffer and sonicated on ice for a total of 60 minutes using a sonicator (Qsonica). The sonication process was carried out with

intermittent pulses of 20 seconds on and 40 seconds off to prevent overheating, ensuring effective cell lysis. The precise control of sonication intervals was critical to maximizing cell disruption while minimizing protein denaturation. The lysate was then centrifuged at 10,000 rpm for 10 minutes at 4 °C to pellet the cell debris.

Affinity Purification of Recombinant Jug r 5 and Jug r 7

The clear supernatant, containing the soluble protein fraction, was carefully collected and applied to a GSTrap FF 5 mL column (Cytiva) pre-equilibrated with PBS buffer. After washing with 5 column volumes of PBS to remove non-specifically bound proteins, elution was performed with 24 mL of 10 mM glutathione solution (dissolved in 50 mM Tris-HCl buffer, pH 8.0). This step efficiently eluted the GST-tagged protein, ensuring its high purity. The eluted fractions containing the target protein were treated overnight at 4 °C with HRV 3C protease at a concentration of 1 mg/mL to remove the GST tag, using a rotating incubator to ensure uniform mixing during the treatment.

Following cleavage, the same His-tag affinity purification procedure outlined in Chapter 2 was used.

Gel Filtration Chromatography of Recombinant Jug r 5 and Jug r 7

The eluted fractions were subjected to gel filtration chromatography using the same procedure described in Chapter 2, with a Superdex 75 pg 16/60 column pre-equilibrated with PBS.

Purification of Storage Proteins from Walnut Seeds

Ammonium Sulfate Precipitation of Storage Proteins

The defatted walnut powder was prepared following the method outlined in Chapter 1.

For the extraction of storage proteins from walnut seeds, the high-salt buffer was utilized, following the method described in Chapter 1. The difference lies in the use of varying saturation levels: 0%, 20%, 40%, 60%, 70%, and 80%. Prior to ammonium sulfate fractionation, insoluble components were removed by centrifugation and labeled as "ppt 0%" and "sup 0%".

Gel Filtration Chromatography of Storage Proteins

The walnut protein precipitate from the 40% ammonium sulfate saturation was dissolved in high-salt buffer and placed into a pre-treated dialysis bag (Biodesign, molecular weight cut-off: 3.5 kDa). Dialysis was performed at 4 °C with high-salt buffer as the dialysis medium, with buffer changes every 8 hours for a total of 24 hours. The dialyzed sample was then concentrated to less than 8 mL using a concentrator (Vivaspin, molecular weight cut-off: 5 kDa) at 6000 g and 4 °C, followed by filtration through a 0.20 µm syringe filter to remove particulates.

Gel filtration chromatography was performed under the same conditions as described in Chapter 2, except that high-salt buffer was used as the mobile phase.

SDS-PAGE

The experimental method used in this chapter is consistent with that described for SDS-PAGE in Chapter 1.

Western Blot

For the identification and confirmation during the purification of recombinant Jug r 5 and Jug r 7, the experimental methods were consistent with those described in Chapter 2.

For the identification and confirmation during the purification of natural walnut storage proteins, the experimental methods followed those outlined in Chapter 1. However, the primary antibody used was rabbit anti-soybean A1aB1b serum antibody.

MS Analysis

The experimental method used in this chapter is consistent with that described for MS in Chapter 1.

ELISA

Serum samples were collected from 24 pediatric participants whose clinical histories indicated an immediate reaction to walnuts. The ELISA procedure was performed according to the same method outlined in Chapter 2.

Results

Purification of Recombinant Jug r 5

Jug r 5 was successfully expressed as a His-tag and GST-tagged fusion protein and underwent purification through a combination of affinity columns and gel filtration chromatography. The initial purification step utilized HiTrap and GSTrap affinity columns, which selectively captured the fusion protein by exploiting the specific interactions of the His-tag and GST with the column ligands. This process effectively separated the fusion protein from the crude extract, achieving initial purification prior to further refinement.

Subsequently, the purified protein underwent further purification via gel filtration chromatography using a Superdex 75 pg 16/60 column. This step is critical for size-based separation of the protein, ensuring that the final product is of high purity. During the gel filtration, the protein began to elute at approximately 70.0 mL (fraction 11) and reached its highest concentration at 73.5 mL (fraction 12), indicating a uniform size distribution and suggesting successful separation of Jug r 5 from other potential contaminants (Figure 21A). The peak observed around fraction 12 suggests that Jug r 5 was predominantly in a monomeric form, free from significant aggregation or degradation. Fractions 11 through 13, which exhibited the highest purity, were carefully collected for subsequent processing.

The final purified Jug r 5 was analyzed using SDS-PAGE, where it appeared as a single band at the expected molecular weight of above 17 kDa (Figure 21B). This result confirms that the protein was not only successfully purified but also maintained its integrity without significant degradation.

The observed molecular weight is consistent with the expected size reported in previous studies (Wangorsch et al., 2017), further validating the successful expression and purification of the recombinant protein.

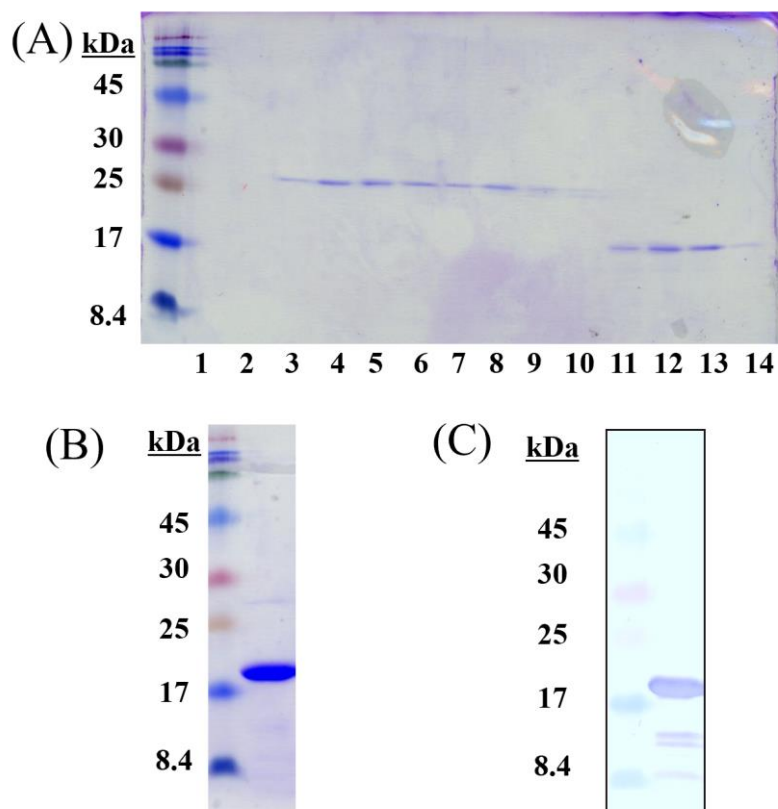


Figure 21. Recombinant Jug r 5

(A) CBB Analysis of Recombinant Jug r 5 Purified by Superdex 75 pg 16/60 Column

(B) CBB Analysis of Purified Jug r 5; (C) Western Blot Analysis of Purified Jug r 5

In panel (A), the samples were collected at different elution volumes, corresponding to fraction numbers 1 to 14. Primary antibody: Anti-6xHis mouse IgG1-κ.

Purification of Recombinant Jug r 7

Jug r 7, like Jug r 5, was expressed as a His-tagged and GST-tagged fusion protein to facilitate efficient purification via two-step affinity columns. Following initial purification on HiTrap and

GSTrap columns, the protein underwent gel filtration chromatography on a Superdex 75 pg 16/60 column for size-based separation.

Due to the lower molecular weight of Jug r 7 compared to Jug r 5, it eluted slightly later, with the highest concentrations observed at approximately 77.0 mL (fraction 13) and 80.5mL (fraction 14) as indicated by CBB staining (Figure 22A). This elution profile suggests that Jug r 7 was successfully isolated in a monodisperse form, free from larger aggregates or potential contaminants.

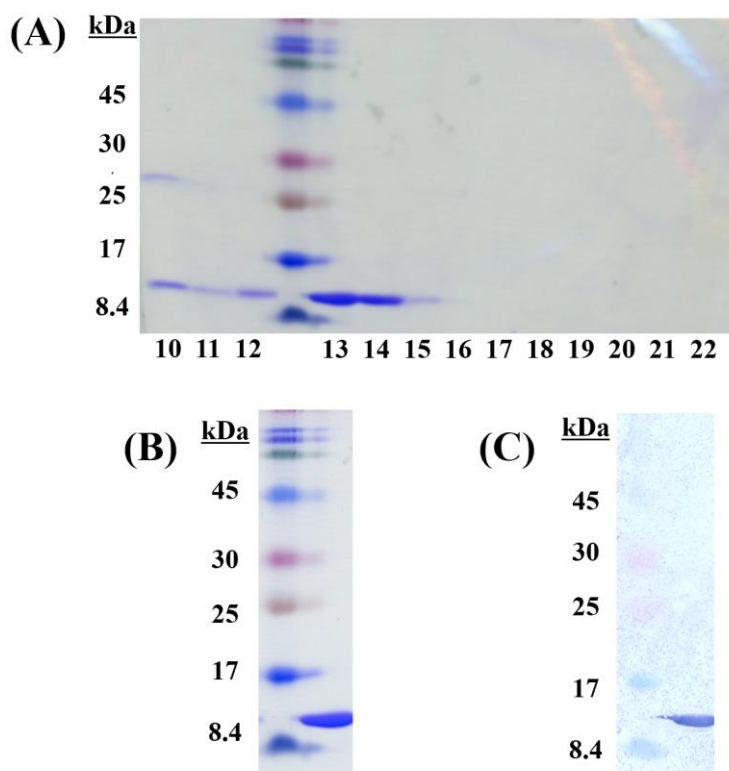


Figure 22. Recombinant Jug r 7

(A) CBB Analysis of Recombinant Jug r 7 Purified by Superdex 75 pg 16/60 Column
 (B) CBB Analysis of Purified Recombinant Jug r 7; (C) Western Blot Analysis of Purified Recombinant Jug r 7. In panel (A), the samples were collected at different elution volumes, corresponding to fraction numbers 10 to 22. Primary antibody: Anti-6xHis mouse IgG1-κ.

SDS-PAGE analysis of purified Jug r 7 showed a clear band between the 8.4 kDa and 17 kDa markers, consistent with the expected size, confirming the protein's high purity and integrity. Western blot analysis further confirmed its specific binding to the anti-6xHis antibody with a 5-minute development time (Figure 22C), underscoring the efficiency of the purification process. The observed molecular weight aligns with previous reports (Versteeg, 2017), validating the accuracy of the expression and purification methods.

Purification of Native Walnut Storage Proteins

To investigate the primary allergens in patients, walnut storage proteins 11S globulin, 7S globulin and 2S albumin were also purified. Given the high abundance of storage protein in walnut seeds, direct extraction from the seeds was more efficient, negating the need for recombinant production. For protein detection, a rabbit anti-soybean A1aB1b serum antibody was used, based on foundational research previously conducted in this laboratory (Maruyama et al., 2018). Soybean glycinin, a member of the 11S globulin family, contains five major subunits, among which A1a1b is one (Choi et al., 2004). Due to the homology between soybean 11S and walnut 11S proteins, the use of this antibody provided a theoretical basis for detecting walnut 11S globulin.

Following the procedure outlined in Chapter 1, walnut storage proteins were extracted using the high-salt buffer, followed by ammonium sulfate precipitation. Preliminary removal of undesired proteins was performed, with optimal conditions achieved at 40% ammonium sulfate saturation (Figure 23).

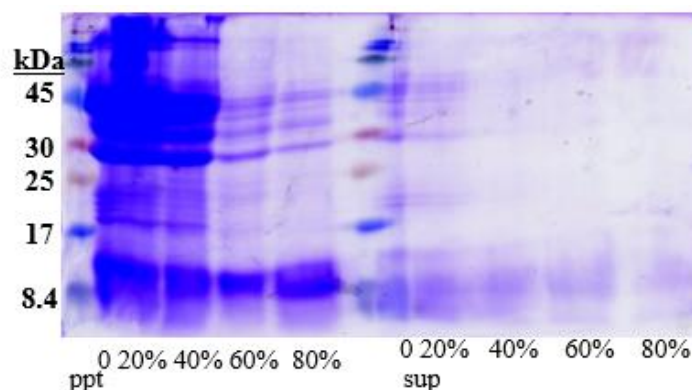


Figure 23. Walnut Protein Ammonium Sulfate Precipitation

"Ppt 0%" and "sup 0%" refer to the insoluble and soluble fractions removed by centrifugation before the ammonium sulfate fractionation process.

After dialysis to remove residual ammonium sulfate, the samples were subjected to Superdex 75 pg 16/60 column for size-based separation. Three peaks were observed, corresponding to fractions 3 to 7, 14 to 16, and 20 to 23. Subsequent analysis of these fractions was performed using CBB staining and Western blotting (Figure 24). The high-molecular-weight (HMW) fraction (fractions 4-6) demonstrated the presence of larger proteins, likely corresponding to 11S globulin and 7S globulin, as evidenced by a positive reaction with the rabbit anti-soybean A1aB1b serum antibody, which indicates the presence of storage proteins. These fractions were subsequently collected and concentrated for further testing. Similarly, the low-molecular-weight (LMW) fraction (fractions 15-16) also showed a high protein concentration and reactivity with the antibody, warranting its collection and concentration for subsequent experiments.

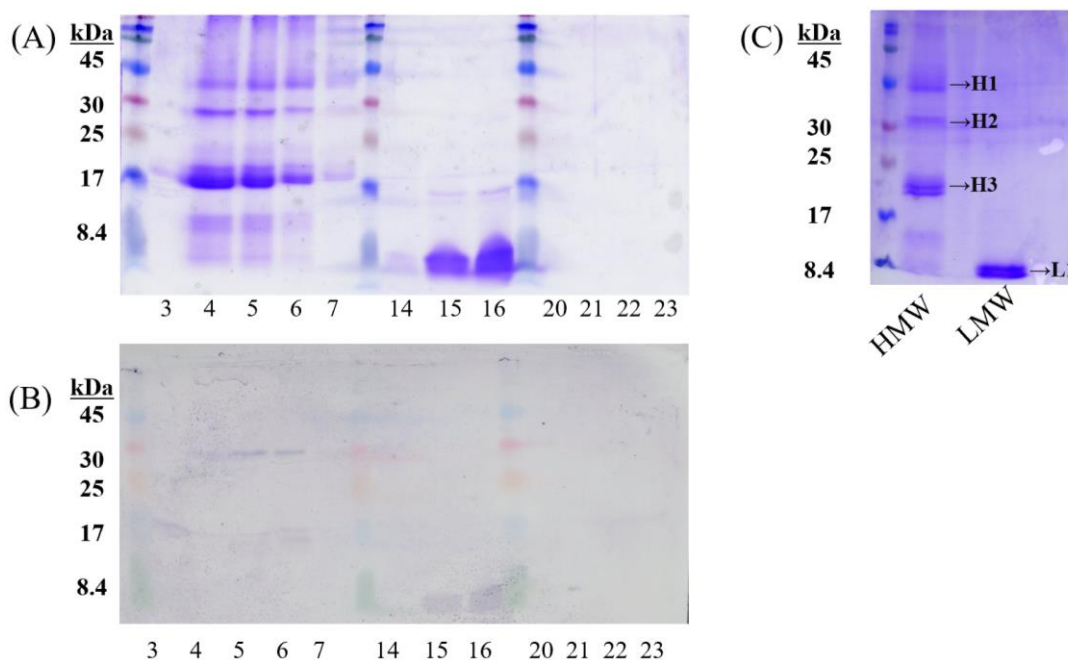


Figure 24. Walnut Protein Purified by Superdex 75 pg 16/60 Column

(A) CBB Analysis of Walnut Protein Fractions Purified by Superdex 75 pg 16/60 Column

(B) Western Blot Analysis of Walnut Protein Fractions Purified by Superdex 75 pg 16/60 Column

In both panels (A) and (B), the samples were collected at different elution volumes, corresponding to fraction numbers 3 to 7, 14 to 16, and 20 to 23.

H1, H2, H3, and L1 in panel (C) represent different samples prepared for MS analysis.

Primary antibody: Rabbit anti-soybean A1aB1b serum antibody.

Identification of Native Walnut Storage Proteins

Based on Figure 24 panel (C), HMW comprises three distinct bands, labeled as H1, H2, and H3 in descending order of molecular weight. In contrast, LMW contains a single low-molecular-weight band, designated as L1. Subsequent MS analysis was performed on the CBB-stained bands from HMW (H1, H2, H3) and LMW (L1), with the detailed identification results presented in Table 7.

Table 7. Proteins Identified by MS from HMW and LMW

Band	Protein (UniProt accession)	Mascot Protein Score
H1	Vicilin Jug r 2.0101 (Fragment) (Q9SEW4)	2450
	Legumin B-like (A0A2I4GEH1)	1918
	Vicilin Jug r 6.0101 (A0A2I4E5L6)	1773
H2	Legumin B-like (A0A2I4GEH1)	2122
	Vicilin Jug r 2.0101 (Fragment) (Q9SEW4)	1932
	11S globulin seed storage protein 2-like (A0A2I4EG83)	1766
	11S globulin (A0A833YDI2)	1755
H3	Legumin B-like (A0A2I4GEH1)	2668
	11S globulin seed storage protein 2-like (A0A2I4EG83)	1934
	11S globulin (A0A833YDI2)	1897
L1	2S sulfur-rich seed storage protein 2 (A0A2I4EX90)	483
	2S albumin seed storage protein (Fragment) (Q7Y1C2)	357

In HMW, the presence of both 11S globulin and 7S globulin were detected across the H1, H2, and H3 bands. The H1 band showed a higher sequence similarity to 7S globulin (Q9SEW4), with a sequence coverage of 56% (Figure 25), indicating that 7S globulin is one of the primary protein components in H1. In contrast, the H2 and H3 bands exhibited a greater similarity to 11S globulin (A0A2I4GEH1), with a sequence coverage as high as 79% (Figure 26). These findings indicate that

11S globulin is the predominant protein component in the HMW fraction, with 7S globulin also present in a certain proportion. This co-presence may be attributed to the natural coexistence of 11S and 7S globulins in walnut seeds and their similar physicochemical properties, which influence their behavior during separation and purification.

10	20	30	40
RGRDDDDDEEN	PRDPREQYRQ	CQEYCRREQGQ	GQRQQQQQCQI
50	60	70	80
RCEERLEEDQ	RSQEERERRR	GRDVDDQNPR	DPEQRYEQCQ
90	100	110	120
QQCERQRRGQ	EQTLCRRRCE	QRRQQEERER	QRGRDRQDPQ
130	140	150	160
QQYHRCQRRR	QIQEQSPERQ	RQCQQRCERQ	YKEQQGR ERG
170	180	190	200
PEASPRRESR	GREEEQQRHN	PYYFHSQSIR	SRHESEEGEV
210	220	230	240
KYLERFTERT	ELLRGIENYR	VVILDANPNT	SMLPHHK DAE
250	260	270	280
SVAVVTRGRA	TLTLVSQETR	ESFNLECGDV	IRVPAGATVY
290	300	310	320
VINQDSNERL	EMVKLLQPVN	NPGQFREYYA	AGAKSPDQSY
330	340	350	360
LRVFSNDILV	AALNTPRDRL	ERFFDQQEQR	EGVIIRASQE
370	380	390	400
KLRALSQHAM	SAGQRPWGRR	SSGGPISLKS	ESPSYSNQFG
410	420	430	440
QFFEACPEEH	RQLQEMDVLV	NYAEIKRGAM	MVPHYNSKAT
450	460	470	480
VVVYVVEGTG	RYEMACPHVS	SQSYEGQGRR	EQEEEEESTGR
490	500	510	520
FQKVTARLAR	GDIFVIPAGH	PIAITASQNE	NLRLLGFDIN
530	540	550	560
GENNQRDFLA	GQNNIINQLE	REAKELSFNM	PREEIEEIFE
570	580	590	593
SQMESYFVPT	ERQSRRGQGR	DHPLASILDF	AFF

Figure 25. Sequence of Walnut 7S Globulin (Q9SEW4) Identified by MS

Matched peptides shown in bold with a sequence coverage of 56%.

10	20	30	40
MAYSYLFSLS	LCFLVLFHGC	LAQLEQVTGQ	SQRQQQQR FQ
50	60	70	80
TECRINNLNA	QEPGRRVESE	AGLSEYWNRN	DEQFRCAGVD
90	100	110	120
LVRHTIQRRG	LLLPSFSNAP	RLVYVVQGRG	LHGAAIPGCP
130	140	150	160
ETFQSESSSQ	FRGEQGSQRL	SRDQHQKVRE	IREGDVVAIP
170	180	190	200
AGVAHWIYND	GESQLIVMIL	YDTSNQANQL	DENARRFYLA
210	220	230	240
GNPHQQQGG	RQRRRESWPG	SRSRSRSQSP	EERERSQQSG
250	260	270	280
SNIFSGFDEE	FLADSFNIDN	ELAMRIQNRD	DQRGIIVTVE
290	300	310	320
DELRVLSPQS	RGEEREERGL	DRDNGVEETF	CTLRRLKHNIA
330	340	350	360
DAQRADVYNE	RGGRITSLNS	FNLPILRYIQ	LSAERGVLYR
370	380	390	400
NALVAPHYNL	NSHSVIYVIR	GNARLQVVGE	NGQNVFDGEI
410	420	430	440
REGQALTVPQ	NFAVIKKAGN	QGF EWVAFKT	NDNAKINALA
450	460	470	480
GRLSTMRALP	EDVLINAYRI	DREEARRLKY	NRDETTLFSS
490			
DSSGSPRRRD			

Figure 26. Sequence of Walnut 11S Globulin (A0A2I4GEH1) Identified by MS

Matched peptides shown in bold with a sequence coverage of 79%.

In LMW, the L1 band exhibited the highest sequence similarity with 2S albumin protein (A0A2I4EX90), with a sequence coverage of 47% (Figure 27), indicating that 2S albumin is the predominant component of LMW. To further explore the immunoreactivity of these proteins, ELISA was conducted using the above samples.

10	20	30	40
MTRLATLAAL	LVALLFVANA	AAFRTTITTM	EIDEDIDNPR
50	60	70	80
RRGEGCRE EQI	QRQQNLNHCQ	YYLRQQSRSG	GYDEDNQRQH
90	100	110	120
FR QCCQQLSQ	MDEQCQCEGL	RQVVR RQQQQ	QGLRGEEMEE
130	140	146	
MVQSARDLPN	ECGISSQRCE	IRRSWF	

Figure 27. Sequence of Walnut 2S Albumin (A0A2I4EX90) Identified by MS

Matched peptides shown in bold with a sequence coverage of 47%.

ELISA

As in Chapter 2, the purified and concentrated walnut proteins were tested via ELISA. Figure 28 presents the IgE reactivity levels of sera from 24 walnut-allergic patients to different walnut allergens (Jug r 5, Jug r 7, HMW, and LMW). The IgE reactivity of each serum was compared based on multiples of the blank value, with each point in the figure representing the IgE reactivity of an individual patient. A value exceeding 10 times the blank value was considered positive.

Figure 28A shows that approximately 33.3% of patients exhibited positive IgE reactivity to Jug r 5. Some patients' serum, such as P4 and P22, displayed extremely high IgE reactivity, suggesting that Jug r 5 may be an important allergen for some patients. Approximately 62.5% of patients showed positive reactivity to Jug r 7.

Figure 28B reveals that 100.0% of patients showed positive IgE reactivity to HMW, while 95.8% exhibited positive reactivity to LMW. Overall, the reactivity levels were higher than those observed for Jug r 5 and Jug r 7. Correlation analysis between total IgE levels and the four proteins showed

significant positive correlations between HMW, and LMW, with r-values of 0.702, and 0.749, respectively ($p < 0.001$). These findings suggest that the storage proteins: 11S globulin and 7S globulin (HMW), and 2S albumin (LMW) may play a primary role in walnut allergy in these patients. In contrast, no significant correlations observed between Jug r 5 and Jug r 7, suggesting that these proteins may function in specific subsets of patients.

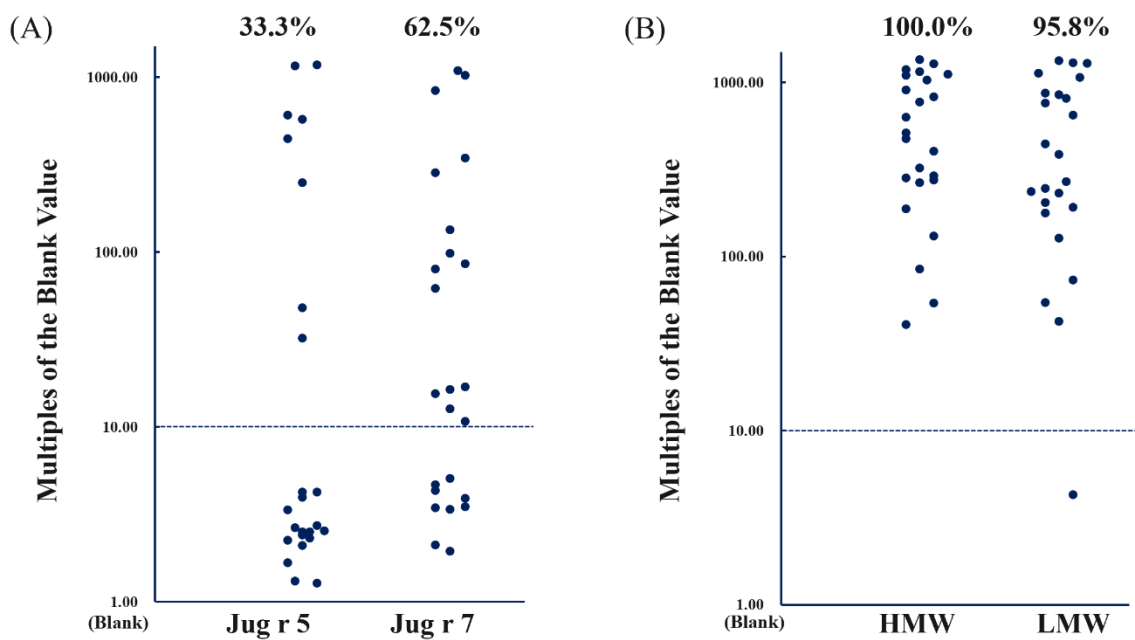


Figure 28. IgE Reactivity of Sera to Walnut Proteins

(A) IgE reactivity to Jug r 5 and Jug r 7 in walnut-allergic patients' sera.

(B) IgE reactivity to walnut storage protein fractions (HMW and LMW).

The y-axis represents the multiples of the blank value, and the horizontal dashed line represents the threshold for a positive reaction (10 times the blank value). Each dot represents the IgE reactivity level of an individual patient's serum. Sera from patients with reactivity levels above the threshold are considered positive for specific IgE against the respective allergen.

Figure 29 analyzes the IgE reactivity to multiple allergens in walnut-allergic patients, revealing individual variations among the patients. The Venn diagram shows that all patients reacted to two or

more proteins. In the central overlapping area, 7 patients exhibited IgE reactivity to all four proteins (Jug r 5, Jug r 7, HMW, and LMW), suggesting a broad sensitization profile and a high degree of cross-reactivity. This indicates that both storage proteins and non-storage proteins could be contributing sensitizing factors in this subset of patients.

Additionally, 7 patients reacted to HMW, LMW, and Jug r 7, but showed no reactivity to Jug r 5. Furthermore, 8 patients reacted to HMW and LMW, but not to Jug r 5 or Jug r 7, highlighting that their allergic responses were primarily driven by storage proteins, with minimal contributions from Jug r 5 and Jug r 7.

The remaining 2 patients exhibited unique IgE reactivity patterns: P13 was negative for Jug r 5, and LMW but positive for Jug r 7 and HMW. P14 was negative for Jug r 7 but positive for Jug r 5, HMW, and LMW. These individual cases will be further analyzed in the discussion section.

In conclusion, the key role of storage proteins is evident: all patients' sera reacted to HMW containing 11S globulin and 7S globulin, and most patients exhibited strong IgE reactivity to HMW, and LMW, confirming the dominant role of storage proteins in walnut allergy. The significant cross-reactivity observed among these proteins suggests their strong sensitizing potential in certain patients. On the other hand, the individualized role of non-storage proteins: while only a few patients reacted to the non-storage proteins Jug r 5 and Jug r 7, these reactions were more prominent in specific individuals, indicating that non-storage proteins may contribute to the sensitization mechanisms in a more complex and personalized manner.

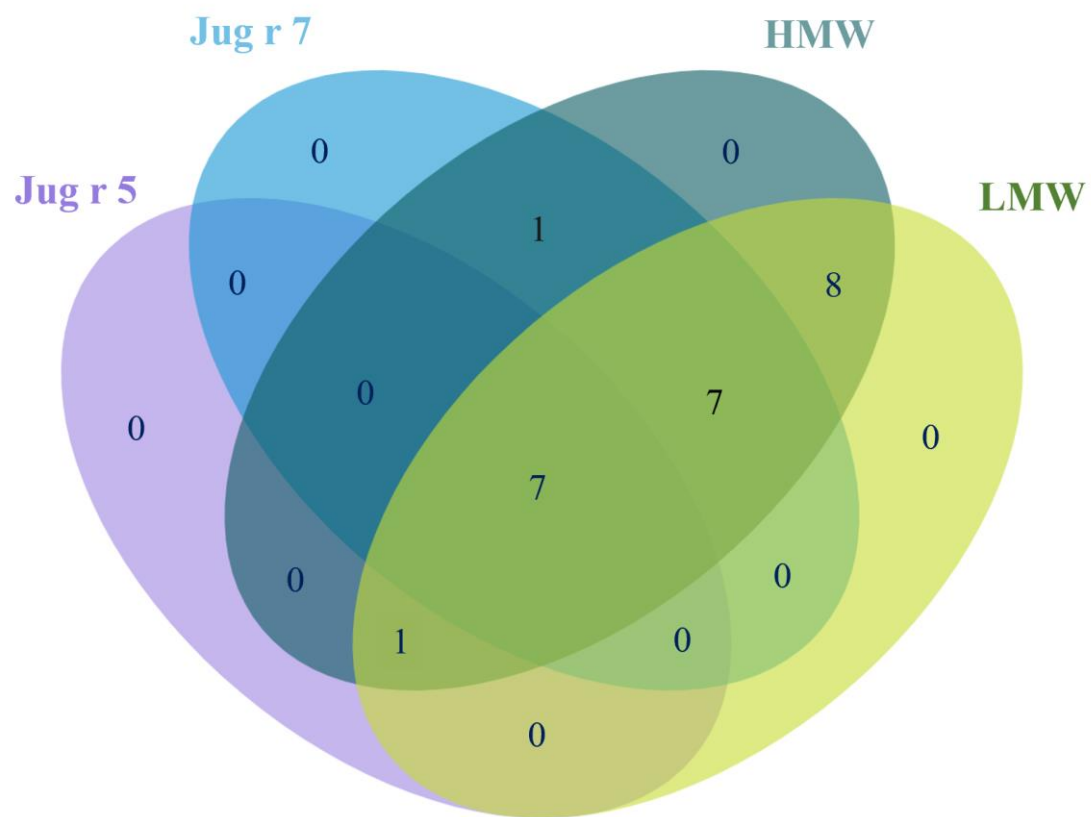


Figure 29. Classification of IgE Reactivity to Walnut Proteins in Sera

The Venn diagram shows the overlap in IgE reactivity to different walnut proteins in sera from 24 patients with walnut allergy. The numbers represent the count of sera reactive to specific combinations of these proteins.

Discussion

This chapter examined the sensitization to Jug r 5, Jug r 7, HMW (11S globulin and 7S globulin), and LMW (2S albumin) in walnut-allergic patients by analyzing the IgE reactivity of their sera to various recombinant and natural walnut proteins. The results showed that the storage proteins in HMW and LMW were the primary allergens for most patients, whereas Jug r 5 and Jug r 7 contributed to allergic reactions in specific subsets of individuals.

To express walnut PR-10 (Jug r 5) and profilin (Jug r 7), different hosts and vectors were utilized. The system for Jug r 5 was established in *E. coli* using the vector pColdGST and the host BL21, while for Jug r 7, the vector pGEX and the host SHuffle T7 were employed. After inducing protein expression with IPTG at different temperature, the cells were lysed via sonication and purified through a GSTrap FF column. To remove the GST tag, the eluted fractions were treated with HRV 3C protease, followed by additional purification using a HisTrap FF 5 mL column. SDS-PAGE and Western blot confirmed the integrity and purity of the recombinant proteins. These proteins were then tested for IgE reactivity against patient sera. Additionally, HMW and LMW protein fractions were extracted from walnut seeds. MS analysis identified the HMW as containing 11S globulin and 7S globulin, while the LMW was primarily composed of 2S albumin. ELISA tests were then performed on these samples using patient sera to further investigate allergenic responses.

The ELISA results showed that HMW and LMW showed positive responses in 100.0% and 95.8% of the patients, respectively. In contrast, Jug r 5 and Jug r 7 had positive rates of 33.3% and 62.5%. All patients' sera reacted to at least two proteins. Previous studies have identified that storage proteins

Jug r 1, Jug r 2, and Jug r 4 are the major allergens in walnut, with IgE binding rates of 75%, 60%, and 57%, respectively (S. Teuber, Dandekar, et al., 1998; S. Teuber, Jarvis, et al., 1999; Wallowitz et al., 2006; Wangorsch et al., 2017).

In this study, the higher positive rates for storage proteins and their significant correlation with total IgE levels suggest that they play a critical sensitizing role in most patients. Additionally, a small subset of patients exhibited more pronounced positive responses to Jug r 5, Jug r 7, suggesting that these proteins may be associated with specific allergic responses, such as PFAS. The subsequent section focuses on individual case analyses:

P4: Although P4 tested positive for all four walnut allergens studied, Jug r 5 showing one of the most pronounced responses among all 24 participants (Figure S2). The IgE response to Jug r 5 was 4 to 13 times higher than the responses to the other allergens, suggesting a dominant sensitization to Jug r 5.

P8: Same to P4, P8 demonstrated positive IgE reactivity to all four tested proteins. Notably, P8 exhibited strong IgE responses to Jug r 5 and Jug r 7, with a pronounced reactivity to Jug r 5. These findings suggest that P8 may exhibit heightened sensitivity to non-storage proteins such as Jug r 5 and Jug r 7, potentially associated with PFAS. Although P8 also reacted to storage proteins, the observed sensitization mechanism appears predominantly directed towards non-storage proteins, thereby distinguishing this patient from the typical walnut allergy profile.

P13: P13 showed positive responses only to Jug r 7 and HMW, with an extremely strong IgE reaction to Jug r 7, one of the highest positive level among the 24 serum samples. This response was

nearly 19 times higher than the response to HMW. These findings suggest that P13 is predominantly sensitized to Jug r 7, a member of the PR-10 protein family, while the response to the storage protein HMW was comparatively weaker.

P14: In contrast to P13, P14 exhibited positive responses to all tested proteins except for Jug r 7. Notably, P14 showed a remarkably high response to Jug r 5, with absorbance values approximately three times higher than those of the other proteins, indicating Jug r 5 may be the primary allergen for this patient.

The analysis of patients P4, P8, P13, and P14 revealed notable interindividual differences in sensitization mechanisms. P8 exhibited strong IgE responses to Jug r 5 and Jug r 7, indicating a sensitization mechanism more focused on non-storage proteins, possibly linked to PFAS. P13 showed predominant sensitization to Jug r 7, with one of the highest IgE level among all 24 participants, nearly 19 times higher than the response to HMW. In contrast, P14 exhibited strong reactivity to Jug r 5, despite showing no response to Jug r 7. P4 exhibited the highest IgE reactivity to Jug r 5, with levels 4 to 13 times higher than other allergens, suggesting a dominant role for Jug r 5 in this patient.

Although 11S globulin, 7S globulin and 2S albumin storage proteins are considered major allergens in most walnut-allergic individuals, these cases demonstrate that non-storage proteins, such as Jug r 5 and Jug r 7, may play a more significant role in specific patients (e.g., P4, P8, P13 and P14). Furthermore, previous studies from European regions reported higher positivity rates for Jug r 3, Jug r 5, and Jug r 7 in walnut-allergic individuals, differing from the patterns observed in this study (Lyons

et al., 2021). These findings suggest potential regional differences in walnut allergy mechanisms that warrant further investigation.

Summary

In this study, for the analysis of walnut allergy, allergens from walnut, pecan, and cashew were prepared and analyzed to better understand the mechanisms underlying walnut allergy.

1. Purification of vicilin N-terminal fragments from walnut, cashew, and pecan seeds

At first, proteins containing N-terminal fragments were extracted from seeds of walnut, pecan and cashew. Due to the varying solubility of N-terminal fragments among different tree nuts, specific buffers were utilized to optimize protein extraction from each type of seed. Following stepwise ammonium sulfate fractionation, the vicilin N-terminal fragments from walnut, cashew, and pecan seeds underwent further purification by HPLC, and their amino acid sequences were subsequently confirmed.

2. Expression, purification, and characterization of recombinant vicilin N-terminal fragments of walnut, pecan, and cashew

The vicilin N-terminal fragments of walnut (Jug r 2 and Jug r 6), pecan (Car i 2), and cashew (Ana o 1) were expressed using the *Pichia pastoris* expression system. DNA fragments encoding these vicilin N-terminal fragments were cloned into the pPICZ α vector, and the resulting expression vectors were transformed into the *Pichia* X33 strain via electroporation. Transformed colonies were inoculated into culture medium and grown under appropriate conditions. Following methanol induction, supernatants were collected by centrifugation. The proteins were then precipitated using 80% ammonium sulfate saturation, dialyzed, and further purified using a Superdex 75 pg 16/60 column.

The presence of vicilin N-terminal fragments in the purified fractions was confirmed by SDS-PAGE and Western blot.

These recombinant proteins were employed to assess their IgE reactivity with sera from 24 walnut-allergic patients, revealing notable cross-reactivity and significant variability in allergenic potential among the different proteins. Specifically, the Jug r 2 and Jug r 6 N-terminal fragments demonstrated IgE reactivity exceeding the threshold in 70.8% of the samples, while the Car i 2 N-terminal fragment exhibited the second positivity rate at 66.7%. In contrast, the Ana o 1 N-terminal fragment showed a much lower positivity rate of 12.5%. These findings underscore the immunogenic variability among vicilin N-terminal fragments, suggesting that their allergenic properties may be intricately linked to their structural features. Further analysis of amino acid sequences and structure revealed that, despite their relatively low sequence identity—falling below the commonly accepted 70% threshold for linear epitope cross-reactivity, these fragments share at least one conserved sequence motif capable of forming disulfide-bonded hairpin structures. This structural feature likely facilitates cross-reactivity. Conformational epitope predictions indicated a high degree of similarity between the J2N-1 and C2N-3 epitopes, which may play a key role in mediating cross-reactivity between the vicilin N-terminal fragments of walnut and pecan. Additionally, the unique localization of J6N-2 outside the typical hairpin structure, coupled with its increased surface accessibility, may contribute to its enhanced immunogenicity.

3. Expression, purification, and characterization of recombinant PR-10 and profilin in walnut

The *E. coli* expression systems for walnut PR-10 (Jug r 5) and profilin (Jug r 7) were constructed using BL21 with pColdGST for Jug r 5, and SHuffle T7 with pGEX for Jug r 7. After culturing and IPTG induction at different temperatures, the cells were harvested, disrupted by sonication, and subjected to affinity purification via a GSTrap FF column. The eluted fractions from the GSTrap FF column were treated with HRV 3C protease and then applied to a HisTrap FF column to remove the GST tag. The purity of the recombinant proteins was confirmed by SDS-PAGE and Western blotting. These purified proteins were subsequently used to assess their reactivity with sera from walnut-allergic patients.

To further investigate the impact of different walnut allergens on patients, walnut storage proteins were purified. HMW (11S globulin and 7S globulin) and LMW (2S albumin) were extracted from walnut seeds. ELISA experiments were then conducted using sera from 24 patients. The results revealed that, unlike European walnut-allergic patients—who predominantly react to non-storage proteins such as Jug r 3, Jug r 5, and Jug r 7—the sera from patients at the National Hospital Organization Sagamihara Hospital exhibited the highest positivity rate for storage proteins. This suggests possible regional differences in walnut allergy mechanisms.

Furthermore, all patients showed positive reactions to two or more allergens, with significant individual variation. While 11S globulin, 7S globulin, and 2S albumin were identified as major allergens, some patients exhibited unique response patterns. For instance, P4 exhibited one of the highest IgE reactivity levels to Jug r 5, with values 4 to 13 times higher than other allergens. Similarly, P14 displayed strong reactivity to Jug r 5 but showed no response to Jug r 7. In contrast, P13

demonstrated one of the highest IgE reactivity levels to Jug r 7 among all patients, while P8 displayed strong reactions to both Jug r 5 and Jug r 7.

These findings indicate that, although storage proteins are the primary allergens for many walnut-allergic patients, non-storage proteins such as Jug r 5 and Jug r 7 may play a more significant role in certain individuals (e.g., P4, P8, P13 and P14). This highlights the importance of considering individual sensitization profiles in the diagnosis and treatment of walnut allergy.

Overall, this research advances our understanding of walnut allergy, providing key insights into the structural basis of allergenicity and cross-reactivity, and laying the groundwork for improved diagnostic methods and targeted therapies for tree nut allergies.

Supplemental Information

Table S1: IgE Reactivity Data of Nuts Vicilin N-terminal Fragments

	Jug r 2 N- Terminal Fragment	Jug r 6 N- Terminal Fragment	Car i 2 N- Terminal Fragment	Ana o 1 N- Terminal Fragment
P1	6396	6684	6465	6016
P2	99591	67760	109054	16485
P3	477662	293601	498602	30288
P4	92035	122781	166684	6634
P5	815649	627744	930639	9358
P6	10087	8495	5656	5350
P7	655775	32297	703055	55981
P8	6771	6531	6381	5829
P9	214587	189394	277273	7120
P10	1128592	723236	361153	7480
P11	10683	11944	8418	6706
P12	8059	7394	35172	5970
Blank	2887			
P13	8926	8243	8254	7894
P14	249806	239572	352878	25010
P15	150099	116856	9240	6443
P16	266024	193272	7348	12739
P17	7117	7715	48744	3792
P18	3153335	2481641	3113185	6419
P19	131045	282590	10086	6082
P20	2592071	2075678	2760973	8788
P21	1178740	790909	1219770	13282
P22	1907852	1940405	2003993	33538
P23	2064023	1563913	2130075	7991
P24	388161	308942	440432	4885
Blank	2755			

Table S2: IgE Reactivity Data of Walnuts Proteins

	Jug r 5	Jug r 7	HMW	LMW
P1	7608	13391	811404	587323
P2	6031	177949	1471254	708366
P3	1642972	2416011	1806879	1269575
P4	3338918	246367	787812	552216
P5	9608	44572	2603360	2185080
P6	3769	5563	760853	366569
P7	6915	9848	117059	156584
P8	715059	282036	242870	122048
P9	6633	9686	927305	662446
P10	7813	48510	3678119	3711343
P11	92272	3135585	2953706	2321365
P12	4780	988755	376539	210746
Blank	2887			
P13	6954	2804048	147938	11737
P14	1658467	5770	516250	487384
P15	131232	44690	3158294	2916577
P16	11574	218354	2999197	2376926
P17	3498	11843	1303362	1053653
P18	1222723	776475	3244964	3094139
P19	6858	368250	3046654	3516334
P20	6879	13846	2115336	2333817
P21	11588	34688	3690834	3653981
P22	3230872	29495	2258603	1779344
P23	10828	10713	1098781	735357
P24	6129	9558	793715	644897
Blank	2755			

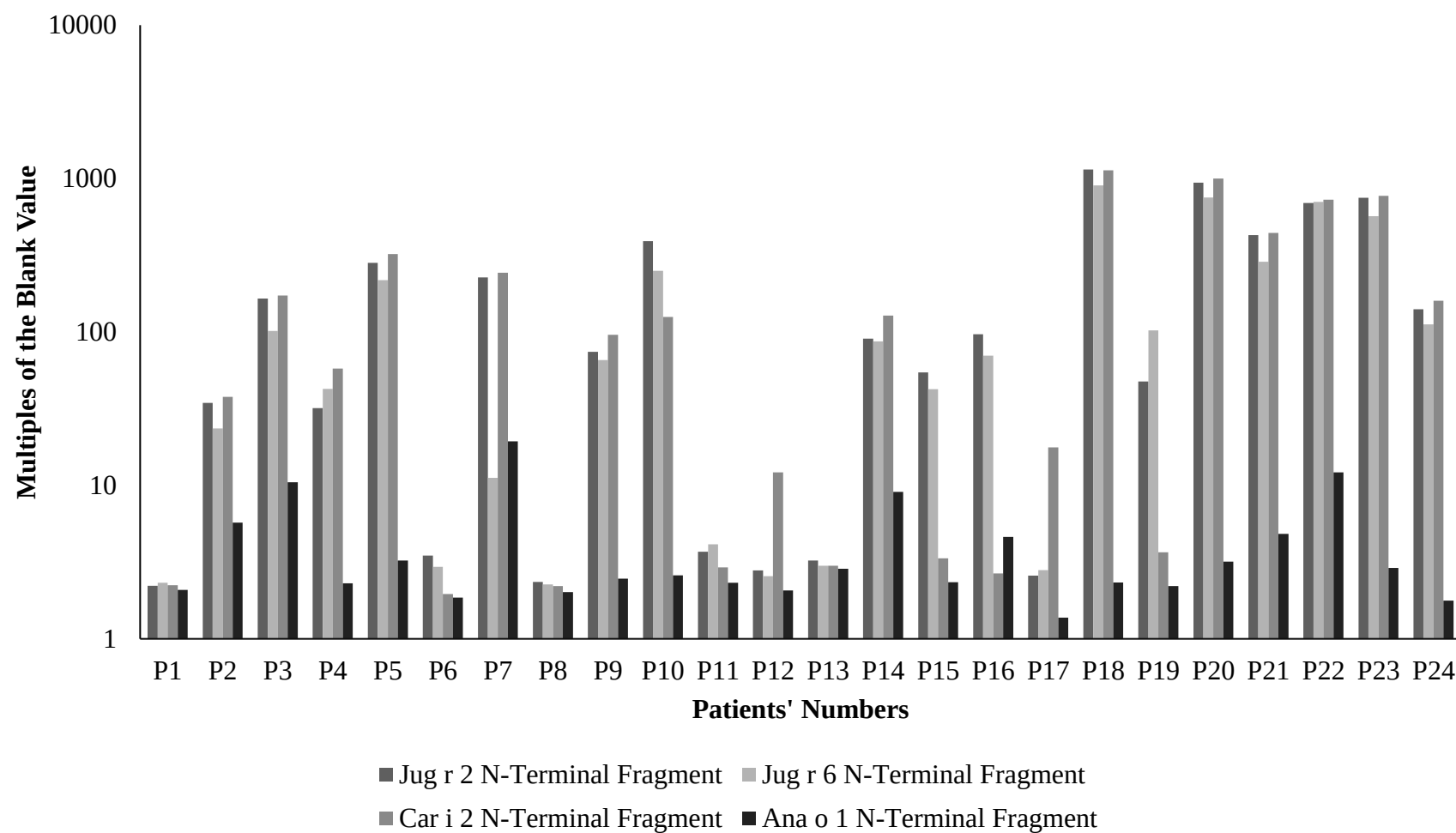


Figure S1. IgE Reactivity of Sera from 24 Patients to Vicilin N-Terminal Fragments from Nuts

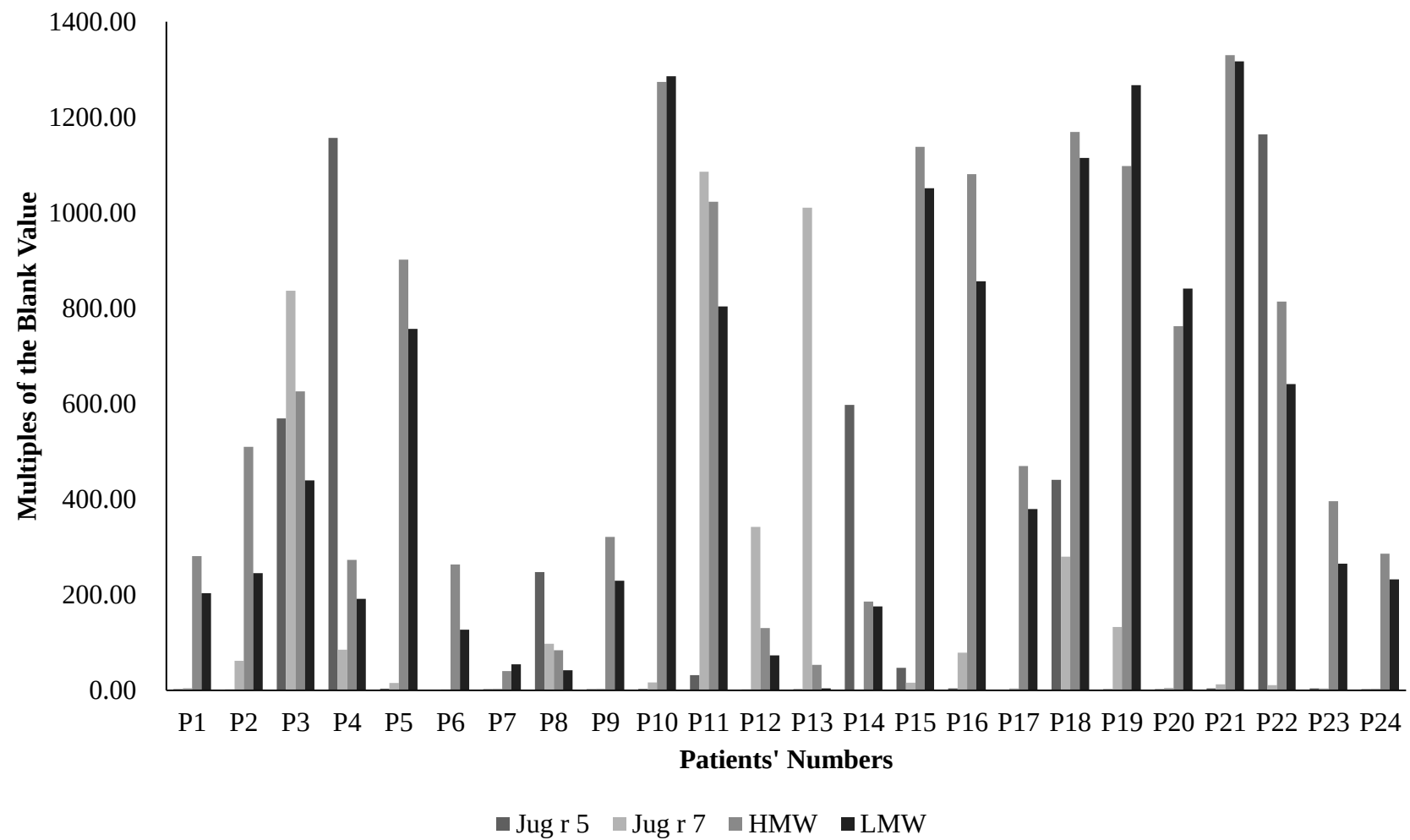


Figure S2. IgE Reactivity of Sera from 24 Patients to Walnut Proteins

References

- Aalberse, R. C. (2000). Structural biology of allergens. *Journal of Allergy and Clinical Immunology*, 106(2), 228-238.
- Aalberse, R. C., Mueller, G. A., Derksen, N. I. L., Aalberse, J. A., Edwards, L. L., Pomés, A., Lidholm, J., Rispens, T., & Briza, P. (2020). Identification of the amino-terminal fragment of Ara h 1 as a major target of the IgE-binding activity in the basic peanut protein fraction. *Clinical and Experimental Allergy*, 50(3), 401–405.
- Acharya, K. R., & Lloyd, M. D. (2005). The advantages and limitations of protein crystal structures. *Trends in Pharmacological Sciences*, 26(1), 10–14.
- Ahmed, H. (2017). *Principles and reactions of protein extraction, purification, and characterization*. CRC press.
- Amin, F., Masoodi, F. A., Baba, W. N., Khan, A. A., & Ganie, B. A. (2017). Effect of different ripening stages on walnut kernel quality: Antioxidant activities, lipid characterization and antibacterial properties. *Journal of Food Science and Technology*, 54, 3791-3801.
- Andersson, K., Ballmer-Weber, B. K., Cistero-Bahima, A., Östling, J., Lauer, I., Vieths, S., & Lidholm, J. (2007). Enhancement of hazelnut extract for IgE testing by recombinant allergen spiking. *Allergy: European Journal of Allergy and Clinical Immunology*, 62(8), 897–904.
- Andorf, S., Borres, M. P., Block, W., Tupa, D., Bollyky, J. B., Sampath, V., Elizur, A., Lidholm, J., Jones, J. E., Galli, S. J., Chinthrajah, R. S., & Nadeau, K. C. (2017). Association of clinical

reactivity with sensitization to allergen components in multifood-allergic children. *The Journal of Allergy and Clinical Immunology: In Practice*, 5(5), 1325-1334.

Asaranoj, A., Movérare, R., Östblom, E., Poorafshar, M., Lilja, G., Hedlin, G., Van Hage, M., Ahlstedt, S., & Wickman, M. (2010). IgE to peanut allergen components: Relation to peanut symptoms and pollen sensitization in 8-year-olds. *Allergy: European Journal of Allergy and Clinical Immunology*, 65(9), 1189–1195.

Asaranoj, A., Nilsson, C., Lidholm, J., Glaumann, S., Östblom, E., Hedlin, G., Van Hage, M., Lilja, G., & Wickman, M. (2012). Peanut component Ara h 8 sensitization and tolerance to peanut. *Journal of Allergy and Clinical Immunology*, 130(2), 468–472.

Baek, M., DiMaio, F., Anishchenko, I., Dauparas, J., Ovchinnikov, S., Lee, G. R., Wang, J., Cong, Q., Kinch, L. N., Schaeffer, R. D., Millán, C., Park, H., Adams, C., Glassman, C. R., DeGiovanni, A., Pereira, J. H., Rodrigues, A. V., van Dijk, A. A., Ebrecht, A. C., Opperman, D. J., Sagmeister, T., Buhlheller, C., Pavkov-Keller, T., Rathinaswamy, M. K., Dalwadi, U., Yip, C. K., Burke, J. E., Garcia, K. C., Grishin, N. V., Adams, P. D., Read, R. J., & Baker, D. (2021). Accurate prediction of protein structures and interactions using a three-track neural network. *Science*, 373(6557), 871–876.

Bechthold, A., Boeing, H., Schwedhelm, C., Hoffmann, G., Knüppel, S., Iqbal, K., De Henauw, S., Michels, N., Devleesschauwer, B., Schlesinger, S., & Schwingshackl, L. (2017). Food groups and risk of coronary heart disease, stroke and heart failure: A systematic review and dose-

- response meta-analysis of prospective studies. *Critical Reviews in Food Science and Nutrition*, 59(7), 1071–1090.
- Borges, J. P., Barre, A., Culerrier, R., Archimbaud, N., Didier, A., & Rougé, P. (2007). How reliable is the structural prediction of IgE-binding epitopes of allergens? The case study of plant lipid transfer proteins. *Biochimie*, 89(1), 83-91.
- Borres, M. P., Sato, S., & Ebisawa, M. (2022). Recent advances in diagnosing and managing nut allergies with focus on hazelnuts, walnuts, and cashew nuts. *World Allergy Organization Journal*, 15(4), 100641.
- Careri, M., & Mangia, A. (2003). Analysis of food proteins and peptides by chromatography and mass spectrometry. *Journal of Chromatography A*, 1000(1-2), 609-635.
- Carlson, G., & Coop, C. (2019). Pollen food allergy syndrome (PFAS): A review of current available literature. *Annals of Allergy, Asthma & Immunology*, 123(4), 359-365.
- Che, H., Zhang, Y., Lyu, S. C., Nadeau, K. C., & McHugh, T. (2019). Identification of Almond (*Prunus dulcis*) Vicilin as a Food Allergen. *Journal of Agricultural and Food Chemistry*, 67(1), 425–432.
- Choi, S. K., Adachi, M., Yoshikawa, M., Maruyama, N., & Utsumi, S. (2004). Soybean glycinin A1aB1b subunit has a molecular chaperone-like function to assist folding of the other subunit having low folding ability. *Bioscience, Biotechnology and Biochemistry*, 68(9), 1991-1994.
- Coates, A. M., Hill, A. M., & Tan, S. Y. (2018). Nuts and cardiovascular disease prevention. *Current Atherosclerosis Reports*, 20, 1-9.

Consumer Affairs Agency, Government of Japan. (2022). *Report on Food Labeling Standards Related to Food Allergies*.

Costa, J., Carrapatoso, I., Oliveira, M. B. P. P., & Mafra, I. (2014). Walnut allergens: Molecular characterization, detection and clinical relevance. *Clinical & Experimental Allergy*, 44(3), 319-341.

Crespo, J. F., James, J. M., Fernandez-Rodriguez, C., & Rodriguez, J. (2006). Food allergy: Nuts and tree nuts. *British Journal of Nutrition*, 96(S2), S95-S102.

Datema, M. R., Zuidmeer-Jongejan, L., Asero, R., Barreales, L., Belohlavkova, S., de Blay, F., Bures, P., Clausen, M., Dubakiene, R., Gislason, D., Jedrzejczak-Czechowicz, M., Kowalski, M. L., Knulst, A. C., Kralimarkova, T., Le, T. M., Lovegrove, A., Marsh, J., Papadopoulos, N. G., Popov, T., del Prado, N., Purohit, A., Reese, G., Reig, I., Seneviratne, S. L., Sinaniotis, A., Versteeg, S. A., Vieths, S., Zwinderman, A. H., Mills, C., Lidholm, J., Hoffmann-Sommergruber, K., Fernández-Rivas, M., Ballmer-Weber, B., & van Ree, R. (2015). Hazelnut allergy across Europe dissected molecularly: A EuroPrevall outpatient clinic survey. *Journal of Allergy and Clinical Immunology*, 136(2), 382–391.

De Brabander, P., Uitterhaegen, E., Delmulle, T., De Winter, K., & Soetaert, W. (2023). Challenges and progress towards industrial recombinant protein production in yeasts: A review. *Biotechnology Advances*, 64, 108121.

Downs, M. L., Semic-Jusufagic, A., Simpson, A., Bartra, J., Fernandez-Rivas, M., Rigby, N. M., Taylor, S. L., Baumert, J. L., & Mills, E. N. C. (2014). Characterization of low molecular weight

- allergens from English walnut (*Juglans regia*). *Journal of Agricultural and Food Chemistry*, 62(48), 11767-11775.
- Dubiela, P., Kabasser, S., Smargiasso, N., Geiselhart, S., Bublin, M., Hafner, C., Mazzucchelli, G., & Hoffmann-Sommergruber, K. (2018). Jug r 6 is the allergenic vicilin present in walnut responsible for IgE cross-reactivities to other tree nuts and seeds. *Scientific Reports*, 8(1), 11366.
- Dunlop, J. H., & Keet, C. A. (2018). Epidemiology of food allergy. *Immunology and Allergy Clinics of North America*, 38(1), 13-25.
- Elizur, A., Appel, M. Y., Nachshon, L., Levy, M. B., Epstein-Rigbi, N., Pontoppidan, B., Lidholm, J., & Goldberg, M. R. (2020). Clinical and molecular characterization of walnut and pecan allergy (NUTCRACKER Study). *The Journal of Allergy and Clinical Immunology: In Practice*, 8(1), 157-165.
- Enrique, E. (2002). Platanus acerifolia pollinosis and food allergy. *Allergy*, 57(4).
- Fernandes, H., Michalska, K., Sikorski, M., & Jaskolski, M. (2013). Structural and functional aspects of PR-10 proteins. *The FEBS Journal*, 280(5), 1169-1199.
- Foo, A. C., Nesbit, J. B., Gipson, S. A., Cheng, H., Bushel, P., DeRose, E. F., Schein, C. H., Teuber, S. S., Hurlburt, B. K., Maleki, S. J., & Mueller, G. A. (2022). Structure, immunogenicity, and IgE cross-reactivity among walnut and peanut vicilin-buried peptides. *Journal of Agricultural and Food Chemistry*, 70(7), 2389-2400.
- Foo, A. C., Nesbit, J. B., Gipson, S. A., DeRose, E. F., Cheng, H., Hurlburt, B. K., Kulis, M. D., Kim, E. H., Dreskin, S. C., Mustafa, S., Maleki, S. J., & Mueller, G. A. (2023). Structure and IgE cross-

- reactivity among cashew, pistachio, walnut, and peanut vicilin-buried peptides. *Journal of Agricultural and Food Chemistry*, 71(6), 2990–2998.
- Fraser, J. S., van den Bedem, H., Samelson, A. J., Lang, P. T., Holton, J. M., Echols, N., & Alber, T. (2011). Accessing protein conformational ensembles using room-temperature X-ray crystallography. *Proceedings of the National Academy of Sciences*, 108(39), 16247–16252.
- Gendel, S. M. (2012). Comparison of international food allergen labeling regulations. *Regulatory Toxicology and Pharmacology*, 63(2), 279–285.
- Glaumann, S., Nilsson, C., Johansson, S. G. O., Asarnoj, A., Wickman, M., Borres, M. P., & Nopp, A. (2015). Evaluation of basophil allergen threshold sensitivity (CD-sens) to peanut and Ara h 8 in children IgE-sensitized to Ara h 8. *Clinical and Molecular Allergy*, 13, 1-6.
- Hauser, M., Roulias, A., Ferreira, F., & Egger, M. (2010). Panallergens and their impact on the allergic patient. *Allergy, Asthma & Clinical Immunology*, 6, 1-14.
- Hoffmann-Sommergruber, K., Mills, E. C., & Vieths, S. (2008). Coordinated and standardized production, purification and characterization of natural and recombinant food allergens to establish a food allergen library. *Molecular Nutrition & Food Research*, 52(S2), S159-S165.
- INC International Nut and Dried Fruit Council. (2023). *Nuts & Dried Fruits Statistical Yearbook 2022/2023*. INC International Nut and Dried Fruit Council. Retrieved from <https://inc.nutfruit.org/wp-content/uploads/2023/05/Statistical-Yearbook-2022-2023.pdf>
- Jeon, Y. H. (2020). Pollen-food allergy syndrome in children. *Clinical and Experimental Pediatrics*, 63(12), 463.

- Jiang, R., Manson, J. E., Stampfer, M. J., Liu, S., Willett, W. C., & Hu, F. B. (2002). Nut and peanut butter consumption and risk of type 2 diabetes in women. *Jama*, 288(20), 2554-2560.
- Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., Back, T., Petersen, S., Reiman, D., Clancy, E., Zielinski, M., Steinegger, M., Pacholska, M., Berghammer, T., Bodenstein, S., Silver, D., Vinyals, O., Senior, A. W., Kavukcuoglu, K., Kohli, P., & Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, 596(7873), 583–589.
- Kabasser, S., Pratap, K., Kamath, S., Taki, A. C., Dang, T., Koplin, J., Perrett, K., Hummel, K., Radauer, C., Breiteneder, H., Lopata, A. L., & Bublin, M. (2022). Identification of vicilin, legumin and antimicrobial peptide 2a as macadamia nut allergens. *Food Chemistry*, 370, 131028.
- Kapingidza, A. B., Pye, S. E., Hyduke, N., Dolamore, C., Pote, S., Schlachter, C. R., Commins, S. P., Kowal, K., & Chruszcz, M. (2019). Comparative structural and thermal stability studies of Cuc m 2.0101, Art v 4.0101 and other allergenic profilins. *Molecular Immunology*, 114, 19–29.
- Kattan, J. D., & Wang, J. (2013). Allergen component testing for food allergy: Ready for prime time? *Current Allergy and Asthma Reports*, 13(1), 58–63.
- Kong, X., Zhang, L., Lu, X., Zhang, C., Hua, Y., & Chen, Y. (2019). Effect of high-speed shearing treatment on dehulled walnut proteins. *Lwt*, 116, 108500.

- Krikeerati, T., Rodsaward, P., Nawiboonwong, J., Pinyopornpanish, K., Phusawang, S., & Sompornrattanaphan, M. (2023). Revisiting fruit allergy: Prevalence across the globe, diagnosis, and current management. *Foods*, 12(22), 4083.
- Liebler, D. C., & Zimmerman, L. J. (2013). Targeted quantitation of proteins by mass spectrometry. *Biochemistry*, 52(22), 3797–3806.
- Liu, C., Gong, J. S., Su, C., Li, H., Li, H., Rao, Z. M., Xu, Z. H., & Shi, J. S. (2022). Pathway engineering facilitates efficient protein expression in *Pichia pastoris*. *Applied Microbiology and Biotechnology*, 106(18), 5893-5912.
- Lorenz, A. R., Scheurer, S., & Vieths, S. (2015). Food allergens: Molecular and Immunological aspects, allergen databases and cross-reactivity. In *Food Allergy: Molecular Basis and Clinical Practice*. Karger Publishers.
- Lyons, S. A., Datema, M. R., Le, T. M., Asero, R., Barreales, L., Belohlavkova, S., de Blay, F., Clausen, M., Dubakiene, R., Fernández-Perez, C., Fritsche, P., Gislason, D., Hoffmann-Sommergruber, K., Jedrzejczak-Czechowicz, M., Jongejan, L., Kowalski, M. L., Kralimarkova, T. Z., Lidholm, J., Papadopoulos, N. G., Pontoppidan, B., Popov, T. A., del Prado, N., Purohit, A., Reig, I., Seneviratne, S. L., Sinaniotis, A., Vassilopoulou, E., Versteeg, S. A., Vieths, S., Zwinderman, A. H., Welsing, P. M. J., Mills, E. N. C., Ballmer-Weber, B. K., Knulst, A. C., Fernández-Rivas, M., & Van Ree, R. (2021). Walnut allergy across Europe: Distribution of allergen sensitization patterns and prediction of severity. *The Journal of Allergy and Clinical Immunology: In Practice*, 9(1), 225-235.

- Ma, S., Sicherer, S. H., & Nowak-Wegrzyn, A. (2003). A survey on the management of pollen-food allergy syndrome in allergy practices. *Journal of Allergy and Clinical Immunology*, 112(4), 784-788.
- Maloney, J. M., Rudengren, M., Ahlstedt, S., Bock, S. A., & Sampson, H. A. (2008). The use of serum-specific IgE measurements for the diagnosis of peanut, tree nut, and seed allergy. *Journal of Allergy and Clinical Immunology*, 122(1), 145–151.
- Martínez-González, M. A., & Bes-Rastrollo, M. (2011). Nut consumption, weight gain and obesity: Epidemiological evidence. *Nutrition, Metabolism and Cardiovascular Diseases*, 21, S40-S45.
- Maruyama, N. (2021). Components of plant-derived food allergens: Structure, diagnostics, and immunotherapy. *Allergology International*, 70(3), 291-302.
- Maruyama, N., Matsuoka, Y., Yokoyama, K., Takagi, K., Yamada, T., Hasegawa, H., Terakawa, T., & Ishimoto, M. (2018). A vacuolar sorting receptor-independent sorting mechanism for storage vacuoles in soybean seeds. *Scientific Reports*, 8(1), 1108.
- Matsuoka, Y., Yamada, T., & Maruyama, N. (2022). Wheat α -gliadin and high-molecular-weight glutenin subunit accumulate in different storage compartments of transgenic soybean seed. *Transgenic Research*, 31(1), 43–58.
- McWilliam, V. L., Perrett, K. P., Dang, T., & Peters, R. L. (2020). Prevalence and natural history of tree nut allergy. *Annals of Allergy, Asthma & Immunology*, 124(5), 466-472.

- Mendes, C., Costa, J., Vicente, A. A., Oliveira, M. B. P., & Mafra, I. (2019). Cashew nut allergy: Clinical relevance and allergen characterisation. *Clinical Reviews in Allergy & Immunology*, 57, 1-22.
- Ministry of Agriculture, Forestry and Fisheries of Japan. (2023). *Agricultural, Forestry, and Fisheries Products Export and Import Overview 2023*.
- Mohammadifard, N., Salehi-Abargouei, A., Salas-Salvadó, J., Guasch-Ferré, M., Humphries, K., & Sarrafzadegan, N. (2015). The effect of tree nut, peanut, and soy nut consumption on blood pressure: A systematic review and meta-analysis of randomized controlled clinical trials. *American Journal of Clinical Nutrition*, 101(5), 966–982.
- Moore, P. B., Anfinsen, C. B., & Dobson, C. M. (2022). The protein-folding problem: Not yet solved. *Science*, 375(6580), 507–507.
- Moreira Da Silva, B., Myung, Y., Ascher, D. B., & Pires, D. E. V. (2022). Epitope3D: A machine learning method for conformational B-cell epitope prediction. *Briefings in Bioinformatics*, 23(1), 1–8.
- O’Neil, C. E., Nicklas, T. A., & Fulgoni III, V. L. (2015). Tree nut consumption is associated with better nutrient adequacy and diet quality in adults: National health and nutrition examination survey 2005–2010. *Nutrients*, 7(1), 595-607.
- Oparin, P. B., Mineev, K. S., Dunaevsky, Y. E., Arseniev, A. S., Belozersky, M. A., Grishin, E. V., Egorov, T. A., & Vassilevski, A. A. (2012). Buckwheat trypsin inhibitor with helical hairpin structure belongs to a new family of plant defence peptides. *Biochemical Journal*, 446(1), 69–77.

- Pastorello, E. A., Farioli, L., Pravettoni, V., Robino, A. M., Scibilia, J., Fortunato, D., Conti, A., Borgonovo, L., Bengtsson, A., & Ortolani, C. (2004). Lipid transfer protein and vicilin are important walnut allergens in patients not allergic to pollen. *Journal of Allergy and Clinical Immunology*, 114(4), 908–914.
- Pawel Dubiela. (2021). *Jug r 8 Allergen Details*. Retrieved from <https://allergen.org/viewallergen.php?aid=924>
- Payne, C. D., Vadlamani, G., Fisher, M. F., Zhang, J., Clark, R. J., Mylne, J. S., & Rosengren, K. J. (2020). Defining the familial fold of the vicilin-buried peptide family. *Journal of Natural Products*, 83(10), 3030-3040.
- Poncet, P., Sénéchal, H., & Charpin, D. (2020). Update on pollen-food allergy syndrome. *Expert Review of Clinical Immunology*, 16(6), 561-578.
- 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.
- Rajaram, S., & Sabaté, J. (2006). Nuts, body weight and insulin resistance. *British Journal of Nutrition*, 96(S2), S79-S86.
- Rasheed, F., Markgren, J., Hedenqvist, M., & Johansson, E. (2020). Modeling to understand plant protein structure-function relationships—implications for seed storage proteins. *Molecules*, 25(4), 873.
- Robotham, J. M., Wang, F., Seamon, V., Teuber, S. S., Sathe, S. K., Sampson, H. A., Beyer, K., Seavy, M., & Roux, K. H. (2005). Ana o 3, an important cashew nut (*Anacardium occidentale* L.)

- allergen of the 2S albumin family. *Journal of allergy and clinical immunology*, 115(6), 1284-1290.
- Serge Versteeg. (2017). *Jug r 7 Allergen Details*. Retrieved from <https://www.allergen.org/viewallergen.php?aid=912>
- Shao, C., Wang, Y., Xu, J., & Zhang, Y. (2022). Assessing PDB macromolecular crystal structure confidence at the individual amino acid residue level. *Structure*, 30(10), 1385–1394.e3.
- Sharma, G. M., Irsigler, A., Dhanarajan, P., Ayuso, R., Bardina, L., Sampson, H. A., Roux, K. H., & Sathe, S. K. (2011a). Cloning and characterization of 2s albumin, Car i 1, a major allergen in pecan. *Journal of Agricultural and Food Chemistry*, 59(8), 4130–4139.
- Sharma, G. M., Irsigler, A., Dhanarajan, P., Ayuso, R., Bardina, L., Sampson, H. A., Roux, K. H., & Sathe, S. K. (2011b). Cloning and characterization of an 11S legumin, Car i 4, a major allergen in pecan. *Journal of Agricultural and Food Chemistry*, 59(17), 9542–9552.
- Shewry, P. R., Napier, J. A., & Tatham, A. S. (1995). Seed storage proteins: Structures and biosynthesis. *The Plant Cell*, 7(7), 945.
- Sicherer, S. H., & Sampson, H. A. (2006). 9. Food allergy. *Journal of Allergy and Clinical Immunology*, 117(2), S470–S475.
- Smeeckens, J. M., Bagley, K., & Kulis, M. (2018). Tree nut allergies: Allergen homology, cross-reactivity, and implications for therapy. *Clinical & Experimental Allergy*, 48(7), 762-772.
- Stone, H., FitzGibbon, L., Millan, E., & Murayama, K. (2022). Curious to eat insects? Curiosity as a key predictor of willingness to try novel food. *Appetite*, 168, 105790.

- Sugizaki, C., Takahashi, K., Sato, S., Yanagida, N., & Ebisawa, M. (2023). National survey of immediate-type food allergies in Japan in 2020: A report supported by a grant from the Consumer Affairs Agency, Government of Japan. *Allergy*, 72(8), 1032-1037.
- Tapia, M. I., Sánchez-Morgado, J. R., García-Parra, J., Ramírez, R., Hernández, T., & González-Gómez, D. (2013). Comparative study of the nutritional and bioactive compounds content of four walnut (*Juglans regia* L.) cultivars. *Journal of Food Composition and Analysis*, 31(2), 232-237.
- Tawde, P., Venkatesh, Y. P., Wang, F., Teuber, S. S., Sathe, S. K., & Roux, K. H. (2006). Cloning and characterization of profilin (Pru du 4), a cross-reactive almond (*Prunus dulcis*) allergen. *Journal of Allergy and Clinical Immunology*, 118(4), 915–922.
- Terwilliger, T. C., Liebschner, D., Croll, T. I., Williams, C. J., McCoy, A. J., Poon, B. K., ... & Adams, P. D. (2024). AlphaFold predictions are valuable hypotheses and accelerate but do not replace experimental structure determination. *Nature Methods*, 21(1), 110–116.
- Teuber, S. S., Dandekar, A. M., Peterson, W. R., & Sellers, C. L. (1998). Cloning and sequencing of a gene encoding a 2S albumin seed storage protein precursor from English walnut (*Juglans regia*), a major food allergen. *Journal of Allergy and Clinical Immunology*, 101(6), 807-814.
- Teuber, S. S., Jarvis, K. C., Dandekar, A. M., Peterson, W. R., & Ansari, A. A. (1999). Identification and cloning of a complementary DNA encoding a vicilin-like proprotein, Jug r 2, from English walnut kernel (*Juglans regia*), a major food allergen. *Journal of Allergy and Clinical Immunology*, 104(6), 1311-1320.

USDA. (2024). *Food Allergies*. Retrieved from <https://www.fsis.usda.gov/food-safety/safe-food-handling-and-preparation/food-safety-basics/food-allergies>

Valenta, R., & Kraft, D. (1996). Type I allergic reactions to plant-derived food: A consequence of primary sensitization to pollen allergens. *Journal of Allergy and Clinical Immunology*, 97(4), 893–895.

Van Regenmortel, M. H. V. (2009). What is a b-cell epitope? *Methods in Molecular Biology*, 524, 3–20.

Vieths, S., Scheurer, S., Reindl, J., Lüttkopf, D., Wangorsch, A., Kästner, M., Haase, T., & Haustein, D. (2001). Optimized allergen extracts and recombinant allergens in diagnostic applications. *Allergy*, 56(Suppl 67), 78-82.

Wallowitz, M., Peterson, W. R., Uratsu, S., Comstock, S. S., Dandekar, A. M., & Teuber, S. S. (2006). Jug r 4, a legumin group food allergen from walnut (*Juglans regia* Cv. Chandler). *Journal of Agricultural and Food Chemistry*, 54(21), 8369–8375.

Wang, F., Robotham, J. M., Teuber, S. S., Sathe, S. K., & Roux, K. H. (2003). Ana o 2, a major cashew (*Anacardium occidentale* L.) nut allergen of the legumin family. *International archives of allergy and immunology*, 132(1), 27-39.

Wang, F., Robotham, J. M., Teuber, S. S., Tawde, P., Sathe, S. K., & Roux, K. H. (2002). Ana o 1, a cashew (*Anacardium occidentale*) allergen of the vicilin seed storage protein family. *Journal of Allergy and Clinical Immunology*, 110(1), 160–166.

- Wangorsch, A., Jamin, A., Lidholm, J., Gräni, N., Lang, C., Ballmer-Weber, B., Vieths, S., & Scheurer, S. (2017). Identification and implication of an allergenic PR-10 protein from walnut in birch pollen associated walnut allergy. *Molecular Nutrition & Food Research*, 61(4), 1600902.
- Waserman, S., & Watson, W. (2011). Food allergy. *Allergy, Asthma & Clinical Immunology*, 7, 1-7.
- Willison, L. A. N., Sathe, S. K., & Roux, K. H. (2014). Production and analysis of recombinant tree nut allergens. *Methods*, 66(1), 34–43.
- Yamada, K., Shimada, T., Kondo, M., Nishimura, M., & Hara-Nishimura, I. (1999). Multiple functional proteins are produced by cleaving Asn-Gln bonds of a single precursor by vacuolar processing enzyme. *Journal of Biological Chemistry*, 274(4), 2563-2570.
- Zeng, J., Jiang, F., & Wu, Y. D. (2016). Folding simulations of an α -helical hairpin motif $\alpha\alpha$ with residue-specific force fields. *The Journal of Physical Chemistry B*, 120(1), 33-41.
- Zhang, J., Payne, C. D., Pouvreau, B., Schaefer, H., Fisher, M. F., Taylor, N. L., Berkowitz, O., Whelan, J., Rosengren, K. J., & Mylne, J. S. (2019). An ancient peptide family buried within vicilin precursors. *ACS Chemical Biology*, 14(5), 979-993.
- Zhang, Y., Lee, B., Du, W. X., Lyu, S. C., Nadeau, K. C., Grauke, L. J., Zhang, Y., Wang, S., Fan, Y., Yi, J., & McHugh, T. H. (2016). Identification and characterization of a new pecan [*Carya illinoensis* (Wangenh.) K. Koch] allergen, Car i 2. *Journal of Agricultural and Food Chemistry*, 64(20), 4146-4151.

Acknowledgments

I would like to express my profound gratitude to Professor Nobuyuki Maruyama for his invaluable guidance, encouragement, and advice throughout my doctoral journey. His support, from the initial topic selection and material gathering to addressing the various challenges encountered during the experimental processes, has been instrumental in my research.

I am deeply grateful to Dr. Motohiro Ebisawa and Dr. Sakura Sato from the National Hospital Organization Sagamihara Hospital for providing the serum samples necessary for this research.

I would also like to extend my appreciation to Professor Akira Oikawa (Laboratory of Quality Analysis and Assessment) and Associate Professor Tomoyuki Katsube-Tanaka (Laboratory of Crop Science) for their valuable feedback and constructive advice throughout the examination process.

Special thanks go to Ms. Rie Suzuki for her kind assistance with experimental reagents and various administrative procedures. I sincerely thank Associate Professor Cabanos Cerrone Salamat and Dr. Misaki Ishibashi for their help and concern during my studies.

I am also grateful to all professors and staff at the Division of Agronomy and Horticultural Science, Graduate School of Agriculture, Kyoto University, where I experienced unwavering support, kindness, and inclusivity.

I deeply appreciate my colleagues in the Laboratory of Food Quality Design and Development for their support. Their collaboration and friendship have made my research journey in Japan truly unforgettable.

Finally, I wish to express my gratitude to my parents, my boyfriend, my cousin Bai Li, and my best friends Shubing and Fan. Their unwavering emotional support throughout my doctoral studies has been invaluable. Despite the physical distance between us, their constant encouragement and love have been my greatest source of strength. May I always carry with me the enduring ability to love and be loved, no matter where life may lead.

BINGYU CHEN