



HOKKAIDO UNIVERSITY

Title	Structural and Stability Analysis of GRP Family Allergens Pru p 7 and Cry j 7, Which Cause Pollen and Food Allergy Syndrome
Author(s)	鄭, 靖康
Degree Grantor	北海道大学
Degree Name	博士(ソフトマター科学)
Dissertation Number	甲第16296号
Issue Date	2025-03-25
DOI	https://doi.org/10.14943/doctoral.k16296
Doc URL	https://hdl.handle.net/2115/97704
Type	doctoral thesis
File Information	Jingkang_Zheng.pdf



Doctoral Dissertation

Structural and Stability Analysis of GRP Family Allergens Pru p 7 and Cry

j 7, Which Cause Pollen and Food Allergy Syndrome

(花粉食物アレルギー症候群の原因となるGRPファミリーアレルゲン

Pru p 7およびCry j 7の構造および安定性の解析)

Zheng Jingkang

Graduate School of Life Science, Hokkaido University

2025.3

Content

Abstract.....	4
1. Introduction.....	5
2. Materials and methods.....	9
2.1 Transformation of <i>P. pastoris</i> and positive transformant selection.....	9
2.2 High cell-density fermentation of <i>P. pastoris</i>	9
2.3 Purification of Cry j 7 and Pru p 7.....	10
2.4 Isotope-labeled GRP proteins reparation	11
2.5 Proteins preparation in the reduction state	12
2.6 Characterization of GRP proteins by mass spectrometry.....	13
2.7 NMR analysis	13
2.8 Circular Dichroism Spectroscopy	14
2.9 ¹ H-NMR spectroscopy analysis of GRP proteins in thermal denaturation	15
2.10 Proteolytic degradation assay	15
2.11 Stability to the Simulated Gastrointestinal Digestion	16
2.12 Epitope prediction.....	16
3. Results.....	18
3.1 Expression and purification of GRP proteins.....	18
3.2. Characterization of GRP proteins.....	23

3.3 Effects of pH and temperature on the secondary structure of GRP proteins	29
3.4 NMR chemical shift assignment and three-dimensional structural analysis.....	31
3.5 ¹ H NMR analysis of thermal denatured GRP proteins structural stability.....	37
3.6 Proteolytic degradation of Cry j 7 and Pru p 7	40
3.7 Stimulate gastrointestinal digestion of GRP proteins	43
3.8 Endosomal degradation of GRP proteins is mediated by Cathepsin S	47
3.9. Epitope prediction of Cry j 7 and Pru p 7.....	52
4. Discussion	56
5. Conclusion	60
6. References.....	61
7. Supplemental.....	72
7.1 Native Pru p 7 sample preparation.....	72
7.2 Molecular dynamics simulation.....	72
7.3 References	73
Acknowledgments.....	88

Abstract

Cry j 7 is a 7 kDa cysteine-rich gibberellin regulatory protein (GRP) with six disulfide bonds. It was isolated from Japanese cedar as the pollen allergen in this study. It exhibits cross-reactivity with food allergens such as Pru p 7 from peach and causes pollen-food allergy syndrome (PFAS). In this work, recombinant Cry j 7 and Pru p 7 were successfully overexpressed using *Pichia pastoris* in a high-cell-density fermentation culture, and pure proteins were purified by reverse phase-HPLC. The characterization of Cry j 7 and Pru p 7 were performed by MS, CD, and ¹H-NMR experiments, to confirm the correct native conformation of Cry j 7 as well as Pru p 7. When compared to the results showed that Cry j 7 exhibits excellent stability in disulfide linkages and preserves its original structure up to 90 °C in various pH buffers in comparison to Pru p 7. Notably, NMR analyses indicated the greater mobility in the α -helix and loop regions of S38-C47 in Pru p 7 compared to those of Cry j 7. Furthermore, our results showed that the sensitivity of Cry j 7 to enzyme digestion differed from that of Pru p 7: the Cry j 7 was more susceptible to proteolysis, while Pru p 7 displayed better resistance in the gastrointestinal tract. These variations in structural stability and sensitivity to proteolysis provide valuable insights into the allergenicity within the GRP family.

1. Introduction

Trees of the Cupressaceae family are distributed on all continents except Antarctica, and 162 species have been identified so far in this family [1]. The pollen from several trees is known to be highly sensitized and is associated with significant allergic diseases [2]. For example, more than 30% of the population in the Mediterranean area is hypersensitive to pollen from European cypress (*Cupressus sempervirens*) [3]. In Japan, the most common species is the Japanese cedar (*Cryptomeria japonica*), and an epidemiological investigation revealed that the prevalence of cedar pollinosis has surged from 10% to 40% in the last 40 years [4,5,6]. It is well known that Cupressaceae pollen allergens are composed of several allergen groups, some of which may be involved cross-reactivity with allergens contained in fruits causing pollen-food allergy syndrome (PFAS) [7].

PFAS is a prevalent immunoglobulin E (IgE)-mediated allergic condition that occurs through a cross-reaction between pollens and foods [8]. The pathogenesis of PFAS was observed to occur in patients who inhale pollen allergens via the respiratory tract and are subsequently exposed to raw plant-derived foods (class 2 food allergy [9]) that share the identical or similar epitopes to pollen proteins, corresponding to the pollen specific IgE antibody [10]. Cross-reactions to apple-fruit-derived Mal d 1 and peach-fruit-derived Pru p 1, etc., due to sensitization to Bet v 1 contained in birch pollen are typical examples of this PFAS caused by the RP-10 family [11]. Besides the main pathogenic protein types, RP-10 or LTP allergens, the recently described gibberellin-

regulated proteins (GRPs), which are widely distributed among all vascular plant species, including seed plants, and essential for plant defense as well as phytohormonal responses [12,13], have also been discovered as allergens involved in PFAS [14,15].

Cry j 7 is a pollen allergen from Japanese cedar, belonging to the GRPs, also known as the Snakin/Gibberellic acid-stimulated Arabidopsis (GASA) protein family [16]. It exhibits antimicrobial properties, and its amino-acid sequence contains 12 cysteine residues that form six disulfide bridges in conserved positions and confer thermostability and proteolytic resistance to the protein [17]. Pru p 7, another GRP family protein, was the first GRP to be identified as a food allergen and isolated from peach fruit [18]. To date, allergenic GRPs found in other plant-foods have been described, namely Pru m 7 from Japanese apricot [19], Cit s 7 from orange [20], Pun g 7 from pomegranate [21], and Cap a 7 from bell pepper [22], while subsequent extensive investigations have confirmed the cross-reactivity with pollen GRPs reported in the Cupressaceae family, such as Cup s 7 from Europe cypress (*Cupressus sempervirens*) and Jun a 7 from Mountain cedar (*Juniperus ashei*) [6,16,23]. Despite coming from distinct plant taxonomic orders, there is a significant degree of sequence conservation between the primary structure of GRPs in plant food and Cupressaceae pollen, according to a comparative analysis [24].

The actual prevalence of food allergies linked to the GRP family has been less widely studied, and studies on reactions to pollen GRP are also rare. In Japan, the incidence of food allergy mediated by GRPs was 65 in 100 patients with responses to

peach, and approximately 20% of patients were sensitized to Pru p 7 from Japanese cedar [25]. Another study on Cry j pollen and fruit GRPs sensitization in Japan revealed that 46% patients were sensitized to Cry j 7 [26]. The connection between Cry j 7 sensitization and Pru p 7 sensitization appears to be related to severe fruit allergies [12]. Further experiments are necessary to explore the effect of GRP allergens in Japanese cedar pollinosis [24].

Previously, we have successfully identified Cry j 7 sequence through a homolog search on the ForestGen website by using Cup s 7 sequence as a query. The *Pichia pastoris* over-expression system was applied to obtain active Cry j 7 and Pru p 7, which can be used in NMR studies. *P. pastoris* is a kind of methylotrophic yeast and qualified with efficient extracellular secretion pathway, as well as post-translational modification to form disulfide bonds [27]. It also contains a methanol inducible alcohol oxidase (AOX1) promoter that regulates the protein induction after adding methanol to the medium [28]. We have already successfully expressed Cup s 7 at high levels using this approach.

In this study, we performed a three-dimensional structure analysis of Cry j 7 and Pru p 7, as well as predicted the flexibility regions to obtain new insights into the structure of the GRP family. We further analyzed the effects of pH, thermal denaturation, and proteolytic enzymes on the structural stability of the GRP allergens. This information is critical for studying the GRP allergens and facilitates investigation of the complex relationship between stability and allergenicity. Comparing the characteristics

of pollen allergen Cry j 7 and food allergen Pru p 7 will help us understand the allergenicity mechanisms of PFAS.

2. Materials and methods

2.1 Transformation of *P. pastoris* and positive transformant selection

The pPIC9-Cry j 7 and pPIC9-Pru p 7 recombinant plasmid vectors were introduced in our experiments within the methylotrophic yeast *Pichia. pastoris* expression system [29,30]. The DNA fragments encoding Cry j 7 and Pru p 7 were amplified by inverse PCR and the gene was simultaneously introduced into the pPIC9 vectors (Figure. S1). The ligated pPIC9-Cry j 7 and pPIC9-Pru p 7 recombinant plasmids were introduced into *E. coil* DH5 α and the presence of the gene inserts was determined by colony PCR and DNA sequencing (3100-Avant Genetic Analyzer; Applied Biosystems/Hitachi, Tokyo, Japan). The recombinant construction was linearized with Sal I enzyme (Thermo Fisher, Waltham, USA) and expressed into the competent *P. pastoris* GS115 strain (Thermo Fisher, Waltham, MA, USA) according to the instructions provided in the *P. pastoris* expression kit v.3.0 manual (Thermo Fisher, Waltham, USA) [31]. Recombinant Cry j 7 and Pru p 7 clones were obtained from MD plates (1.34% YNB, 4 $\times 10^{-5}$ % biotin, 2% dextrose, 1.5% agar).

2.2 High cell-density fermentation of *P. pastoris*

High-cell-density fermentation of *P. pastoris* was carried out in a 5.0 L jar fermenter (MBF-500; Tokyo Rikakikai, Tokyo, Japan), following the established procedure [32]. A single colony from a positive transformant was incubated in 5 mL of YPD medium overnight at 30 °C and 200 rpm. About 400 μ L of overnight cell culture was inoculated into 200 mL BMG medium (100 mM potassium phosphate pH 6.0, 1.34%

YNB, $4 \times 10^{-5}\%$ biotin, 1% glycerol) and cultured for 24 h at 30 °C, and then the resulting culture was inoculated into 2.4 L first medium (50% glycerol, 10×basal salt 840 mL, H₂O 1380 mL, PTM1 10 mL, 500×biotin 10 mL) in jar fermenter. During the fermentation, the temperature and agitation rate were maintained at 30 °C and 800 rpm, respectively; the pH was adjusted at 5.0 by adding 10% (v/v) ammonia water; the foam was decreased by an antifoaming agent (10% propylene glycerol). When the glycerol phase was completed, as identified by the sudden increase in the level of dissolved oxygen, then delivered second medium (50% glycerol 300 mL, PTM1 3.6 mL, 500×biotin 3.6 mL) at a speed of 30 mL/h to increase the cell mass. Within 7 h of the second medium feeding, a third medium (MeOH 900 mL, PTM1 10.8 mL, 500×biotin 10.8 mL) started at 3 mL/h rate. The feed was gradually increased until the maximum feed rate at 15 mL/h, and the whole cultivation time was about 42 hours. After putting methanol, 1 mL samples were collected every 12 h to confirm the cell density of target proteins. At the end of the fermentation process, the cells were centrifuged at 6000 rpm for 20 minutes, and the supernatant was frozen and stored at -20 °C [33]. The expression levels of Cry j 7 and Pru p 7 were evaluated by tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis (tricine SDS-PAGE) followed by Coomassie brilliant blue staining.

2.3 Purification of Cry j 7 and Pru p 7

The pH and conductivity of supernatants containing expressed Cry j 7 and Pru p 7 were adjusted to 6.0 and 4.9 mS/cm, respectively, using 1 M NaOH and Milli-Q water.

The filtered products were loaded onto a 5 mL pre-packed HiTrap SP FF cation exchange column (Cytiva, Marlborough, USA) equilibrated with buffer (50 mM potassium phosphate, pH 6.0). Then, the proteins were eluted with a linear gradient of 15–36% elution buffer (50 mM potassium phosphate and 1 M NaCl, pH 6.0). Solutions were filtration (Millex-HV 0.22 μ m, Merck, Rahway, NJ, USA) and mixed with trifluoroacetic acid (TFA, Nacalai Tesque Inc., Kyoto, Japan) to achieve a pH of 2.0–3.0 for reverse phase HPLC (RP-HPLC). Proteins were purified using a Cosmosil 5C18AR-300 column (Nacalai Tesque Inc., Kyoto, Japan). Elution was performed by using a linear gradient of 15–25% acetonitrile with 0.1% TFA. The contents of each fraction were assessed using 15% tri-cine-SDS-PAGE without reducing reagent [33]. The yield of Cry j 7 and Pru p 7 were estimated by measuring the absorbance at 280 nm. The purity and molecular weight of Cry j 7 and Pru p 7 were verified by MALDI-TOF MS (UltrafleXtreme; USA). In addition, we also prepared natural Pru p 7 samples as control, and the specific information can be found in Supplemental.

2.4 Isotope-labeled GRP proteins preparation

Baffled-flask cultivation as the culture method for isotope-labeled samples preparation [34]. The ^{15}N isotope labeling reagent was used as the nitrogen source from CIL company (Cambridge Isotope Laboratories, Inc. Japan). A single transformant colony was grown at 30 °C about 28 h in a 300 mL shaking flask containing 50 mL BMD1 medium (100 mM potassium phosphate buffer, pH 6.0; 0.34%, w/v, YNB w/o amino acids and ammonium sulfate; 1% w/v $^{15}\text{NH}_4\text{Cl}$; $4 \times 10^{-5}\%$, w/v, biotin and 1%,

w/v, D-glucose) until OD600 = 2~6. The culture was centrifuged (3000 rpm, 5 min), then cells were resuspended in the growth medium BMD2 (100 mM potassium phosphate buffer, pH 6.0; 0.34%, w/v, YNB w/o amino acids and ammonium sulfate; 0.2% w/v $^{15}\text{NH}_4\text{Cl}$; $4 \times 10^{-5}\%$, w/v, biotin and 0.5%, w/v, D-glucose). The growth phase was pursued for 12 h and stopped by centrifugation (3000 rpm, 5 min). Cells were washed with 0.2% glycerol solution for twice and collected by centrifugation before transferred into 500 mL induction medium BMM (100 mM potassium phosphate buffer, pH 6.0; 0.34%, w/v, YNB w/o amino acids and ammonium sulfate; 1% w/v $^{15}\text{NH}_4\text{Cl}$; $4 \times 10^{-5}\%$, w/v, biotin and 0.5% methanol) [35]. The culture was shaken at 30 °C and 160-200 rpm for 72 hours induction period, with methanol added every 12 h to a final concentration of 0.5% (v/v). The purification processes were the same as the previously described steps (2.3), and the expression level was determined by tricine SDS-PAGE analysis.

2.5 Proteins preparation in the reduction state

To obtain the reduction state of Cry j 7 and Pru p 7, 1 M dithiothreitol (DTT) was added to 2–5 mg/mL Cry j 7oxi or Pru p 7oxi solution. The solution was then incubated at room temperature. After 16 h, the reaction solution was adjusted to pH 2.0~3.0 and purified by RP-HPLC through a linear gradient of 15–40% acetonitrile with 0.1% TFA, then lyophilized and stored at –30 °C before use. The procedure for getting the ^{15}N -isotope labeled reduced sample was described as above.

2.6 Characterization of GRP proteins by mass spectrometry

Matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF) was carried out using a Bruker Autoflex Speed mass spectrometer (Bruker Daltonics). The protein samples were combined with sinapic acid matrix (α -Cyano-4-hydroxycinnamic acid in 0.1% TFA solvent) on a MALDI sample plate, and then the plate was loaded onto the spectrometer for analysis.

2.7 NMR analysis

The ^{15}N -isotope-labeled Cry j 7 and Pru p 7 were dissolved in a 10% D_2O /90% H_2O mixture at a final concentration of 0.8 mM at pH 3.0. NMR experiments were carried out at 25 °C on Bruker 800 MHz (AVANCE NEO; Bruker, USA) and 600 MHz (AVANCE III HD; Bruker, Billerica, MA, USA). The main chain assignment of ^1H -, ^{13}C - and ^{15}N -resonance were performed using the following spectra: [^1H - ^1H] TOCSY and [^1H - ^1H] NOESY with ^{15}N -decoupling, [^1H - ^{15}N] HSQC, [^1H - ^{13}C] HSQC, HNCOC, HNCA, ^{15}N -TOCSY and ^{15}N -NOESY. All chemical shift values were referenced against DSS and the following frequency ratios: ($^{15}\text{N}/^1\text{H}$) = 0.101329118, ($^{13}\text{C}/^1\text{H}$) = 0.251449519 [36]. Moreover, hetero-nuclear [^1H - ^{15}N] NOESY-HSQC with a mixing time of 100 ms for ^1H - ^1H NOESY mixing time provided inter-proton distance restraints for structural analysis. The experimental information is shown in Table S1. The 2D and 3D NMR spectra were processed using NMRPipe [37], and the data were analyzed using the Sparky program (<http://www.cgl.ucsf.edu/home/sparky/> access on 30th May 2008).

Nuclear overhauser effect (NOE) distance constraints were automatically assigned and computed in 7 cycles of the 'noeassign' macro of the CYANA 2.1 software package [38]. To rectify the automatic NOE signal assignment, we extracted restraints between the medium and long-range residues and used a pseudo-atom for all optical isomeric protons. As a result, the distance restraints had a fixed upper limit value of 6 Å. The final ensemble of 20 structures was chosen based on CYANA target function values after the 30,000 simulated annealing steps that were used to generate 100 structures in each "noeassign" cycle [39].

2.8 Circular Dichroism Spectroscopy

All CD measurements were performed via a Jasco J-725 spectropolarimeter (Jasc.o, Japan) using cuvettes with a 1 mm path length. Freshly made 0.1 mg/mL (14.63 µM) Cry j 7 and 0.1 mg/mL (14.47 µM) Pru p 7 proteins were dissolved in 50 mM PBS at pH 7.4. Spectra were acquired at a 50 nm/min scan rate with 4 scans taken from 260 nm to 190 nm registered and averaged at room temperature, 25 °C.

For pH and thermal denaturation treatments, pure Cry j 7 and Pru p 7 powders were diluted to a final concentration of 5 mM in a 200 mM Na-citrate (pH 3.0) stock solution. At pH 6.0, 8.0, and 10.0, Cry j 7 and Pru p 7 samples were directly dialyzed against the respective buffers (pH 6.0 and 8.0: 50 mM Na-phosphate buffer, pH 10.0: 50 mM Na-carbonate buffer). The temperature of samples at 10 µM was increased progressively at a constant rate of 2 °C/min from 20 °C to 90 °C and subsequently cooled back down to 20 °C. Signals at a wavelength of 210 nm were monitored with a

step resolution of 0.2 °C [40]. The mean residue ellipticity values, θ , were calculated using the formula below [41]:

$$\theta = \frac{\theta_{\text{observed}}}{10 \times n \times C \times l}$$

where n is the number of amino acid residues, C is the peptide concentration, and l is the optical pass length of the cell.

2.9 ¹H-NMR spectroscopy analysis of GRP proteins in thermal denaturation

Freeze-dried Cry j 7 and Pru p 7 samples were dissolved in a mixture of 10% D₂O/90% H₂O at pH 3.0. The temperature was incrementally adjusted from 10 °C to 75 °C using the different Chiller mode (0 °C ~ 20 °C: maximum; 20 °C ~ 30 °C: medium; 30 °C ~: off). NMR experiments were performed on a Bruker Avance III HD 600 MHz instrument.

2.10 Proteolytic degradation assay

GRPs at a 0.5 mg/mL final concentration were individually incubated with 0.1 mg/mL thermolysin at 30 °C in a solution containing 20 mM HEPES, 200 mM NaCl, and 5 mM CaCl₂, at pH 7.4. [42]. Reaction aliquots from the proteolytic degradation assay were taken out at different intervals, and then proteolysis was quenched by the addition of 12.5 mM EDTA. The aliquot samples of GRPs were compared to those under tricine SDS-PAGE and gel bands were quantified using Image J software 1.45 (<http://rsbweb.nih.gov/ij/> access on 29 October 2011) [43].

Alternatively, the proteolytic stability against endo-lysosomal proteases was

assessed by employing pure Cathepsin S enzyme (human, recombinant from *E.coil*, Sigma. Aldrich, USA) [44]. Proteins were incubated with cathepsin S in a digestion buffer (0.1 M sodium acetate, pH 5.0, 0.1 M NaCl, 5 mM EDTA, and 2 mM DTT) at 37 °C for various durations times. At each time point, the reaction was halted by boiling for 5 minutes at 95 °C [45]. The degradation characteristics were analyzed by tricine SDS-PAGE and MS.

2.11 Stability to the Simulated Gastrointestinal Digestion

In vitro gastrointestinal digestion was carried out as previously described [46]. Briefly, GRP proteins were subjected to pepsin (Sigma. Aldrich, St. Louis, MOUSA) digestion in a 1:20 w/w enzyme-to-substrate ratio in 0.15 M NaCl, the pH of which was adjusted to pH 2.0 with 1 M HCl. The incubation lasted for 2 h at 37 °C and it was stopped by raising the pH to 7.4 by adding PBS [47]. For intestinal digestion, the product mixture, resulting from the gastric digestion reaction was further subjected to trypsin (Porcine, Sigma. Aldrich, St. Louis, MO, USA) and α -chymotrypsin (Sigma. Aldrich, St. Louis, MO, USA) digestion in an enzyme-to-protein ratio of 1:50 (w/w) in 0.05 M potassium phosphate buffer, the pH of which was adjusted to 6.8. After 48 hours of reaction at 37 °C, the process was terminated by heating the mixture to 95 °C for 2 minutes [48]. The degree of gastrointestinal digestion was monitored through tri-cine SDS-PAGE and RP-HPLC.

2.12 Epitope prediction

To computationally predict the epitopes of Cry j 7 and Pru p 7, we primarily

employed epitope prediction methods from the immune Epitope Database (IEDB) [49]. We utilized IEDB MHC class II binding sites and online tool ProPred [50] to predict Th-cell epitopes. To acquire the high scoring peptides, we set a 5% threshold for ProPred. Our analyses focused on common human alleles (DRB1*0101, DRB1*0301, DRB1*0401, DRB1*0701, DRB1*0801, DRB1*1101, DRB1*1301, DRB1*1501 and DRB5*0101), which are alleles considered as representative in humans. Fewer than 15 residues were evaluated for T-cell epitopes. Simultaneously, the latest version of the BepiPred 3.0 website was considered for B-cell linear epitope prediction, with the threshold set to 0.15 [51].

Additionally, the Disco Tope 2.0 dataset is a method for predicting discontinuous epitopes from 3D structures of proteins in PDB format [52]. By uploading the PDB file of the NMR structure to the website (<http://tools.iedb.org/discotope/> access on 27 December 2012), the PyMOL 4.6 software displayed the corresponding epitope regions.

3. Results

3.1 Expression and purification of GRP proteins

To achieve high yields, we performed high-cell-density fermentation in a jar fermenter, beginning with basal salt medium and involving medium supplementation with ammonia and glycerol. The total methanol induction period spanned 42 hours. Following culture, the yeast cell density of Cry j 7 and Pru p 7 increased to 220 g/L and 290 g/L, respectively. A 7 kDa protein band was identified by tricine SDS-PAGE analysis of the culture supernatant (Figure 1). Two bands with similar molecular weights were observed, but upon reduction, these bands merged into one, suggesting that they are probably different molecular species of the same protein with different disulfide bridges' content. We also observed that protein expression started gradually 12 hours after methanol induction and reached its peak at 42 hours. These findings suggested that the secreted expression of Cry j 7 and Pru p 7 secretions in *P. pastoris* was effectively consistent with observations made in past reports.

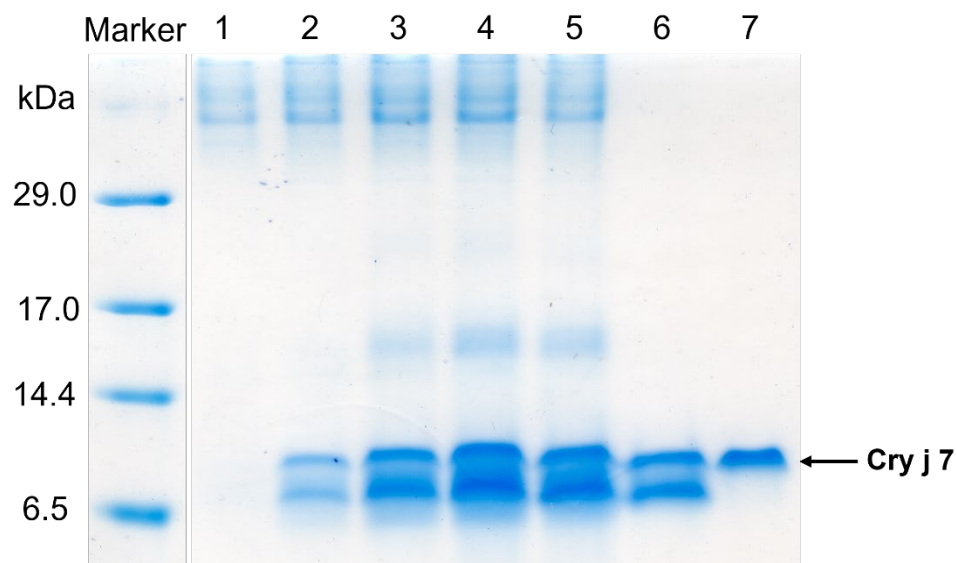


Figure 1. Tricine SDS PAGE analysis of Cry j 7 in fermentation supernatants from *P. pastoris* and purification steps. Lane marker: protein molecular weight mark. Lanes 1-5: a total of 20 μ L of supernatant samples taken at 0, 12, 24, 36, and 42 hours of induction, respectively. Lane 6: the fraction eluted at 28.0-29.0 min via cation exchange chromatography (CIEX). Lane 7: the peak fraction (17.5-18.0 min) obtained from RP-HPLC.

At the end of the fermentation process, the supernatants of the culture were collected through centrifugation and applied to a cation exchange column as the first purification stage. The fractions containing the expressed proteins were further separated by using a PR-HPLC C18 column, and the purified products in each step were analyzed by tricine SDS-PAGE. Cry j 7 appeared as a single peak in the HPLC profile, with a retention time between 17.0 and 19.0 minutes (Figure 2A), confirming that it was isolated with high purity. By performing complete reduction, the hydrophobicity of these proteins increased, and they showed a slower elution time in RP-HPLC (Figure 3). In the case of Pru p 7, the main peak that was present from 23.0 to 27.0 minutes likewise showed good separation, even though other peaks occurred after 25.0 minutes (Figure 2B). The yield estimated from the absorbance at 280 nm for purified Cry j 7 was 20 mg from the 1 L fermentation cultivation medium, while it was 45 mg for Pru p 7 obtained from the 1 L culture. The purified samples were subsequently freeze-dried in liquid nitrogen and stored at -30 °C for further use in various experiments, including structural stability analysis and proteolysis research.

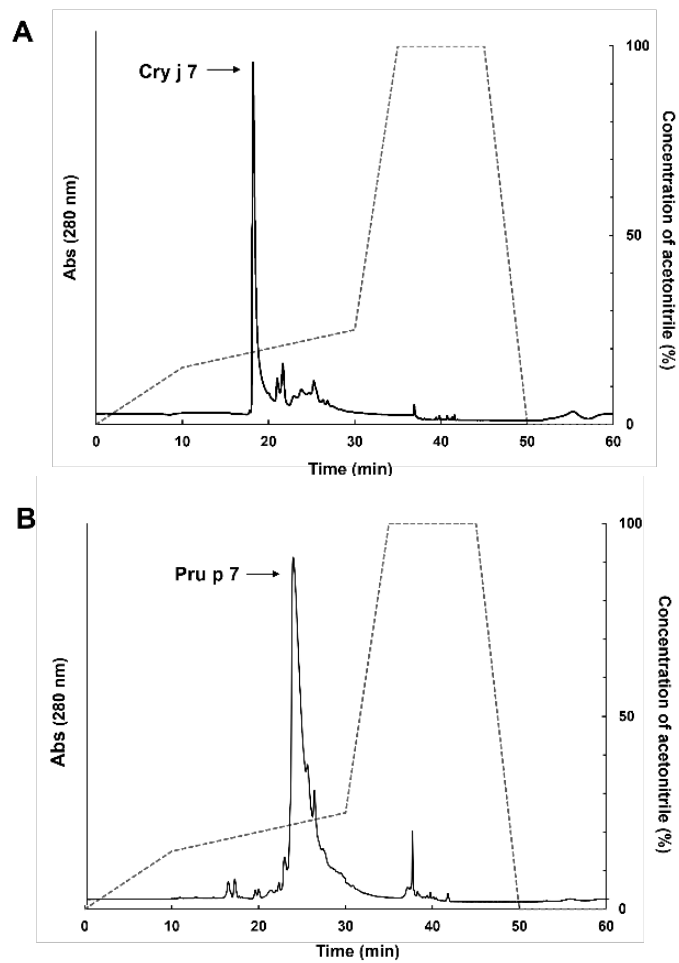


Figure 2. Purification of Cry j 7 (A) and Pru p 7 (B) by an RP-HPLC C18 column that was eluted with a linear gradient of 15-25% acetonitrile with 0.1% TFA. The black line shows the absorbance at 280 nm. The gray dotted line indicates the acetonitrile gradient.

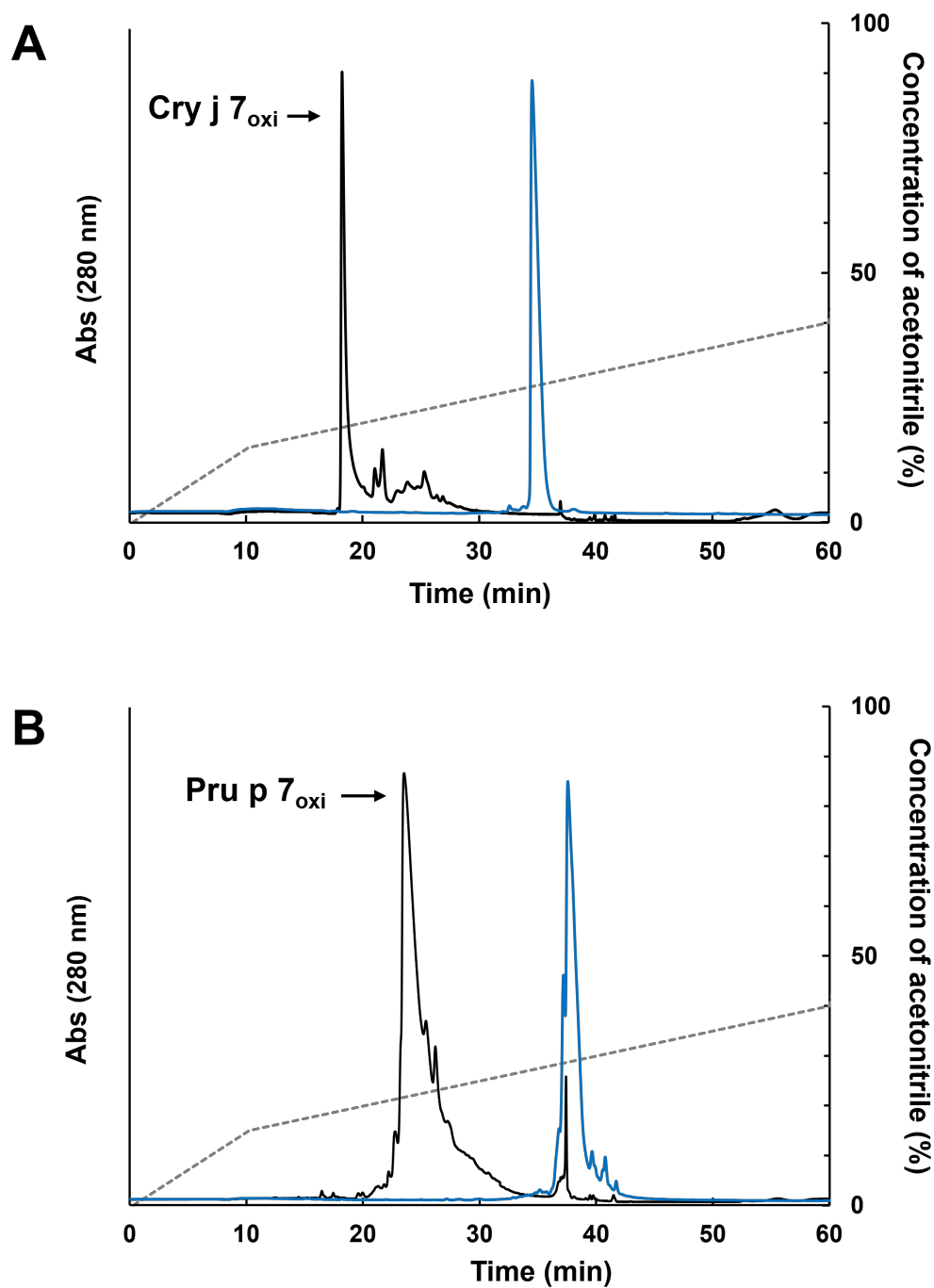


Figure 3. The RP-HPLC analysis of Cry j 7 (A) and Pru p 7 (B) under oxidation or reduction conditions. Purification of recombinant samples by a C18 column that was eluted with a linear gradient of 15-40% acetonitrile with 0.1% TFA. Black line: oxidation state. Blue line: reduction state.

3.2. Characterization of GRP proteins

Cry j 7 and Pru p 7, after being purified by RP-HPLC, were evaluated for their molecular weights via MALDI-TOF MS. The measured spectrum of the Cry j 7 and Pru p 7 with six disulfide bonds (Figure 4) exhibited the desired peaks, corresponding to $[M+H]^+$ at m/z 6889.68 and m/z 6896.59, respectively, which were within the expected error range.

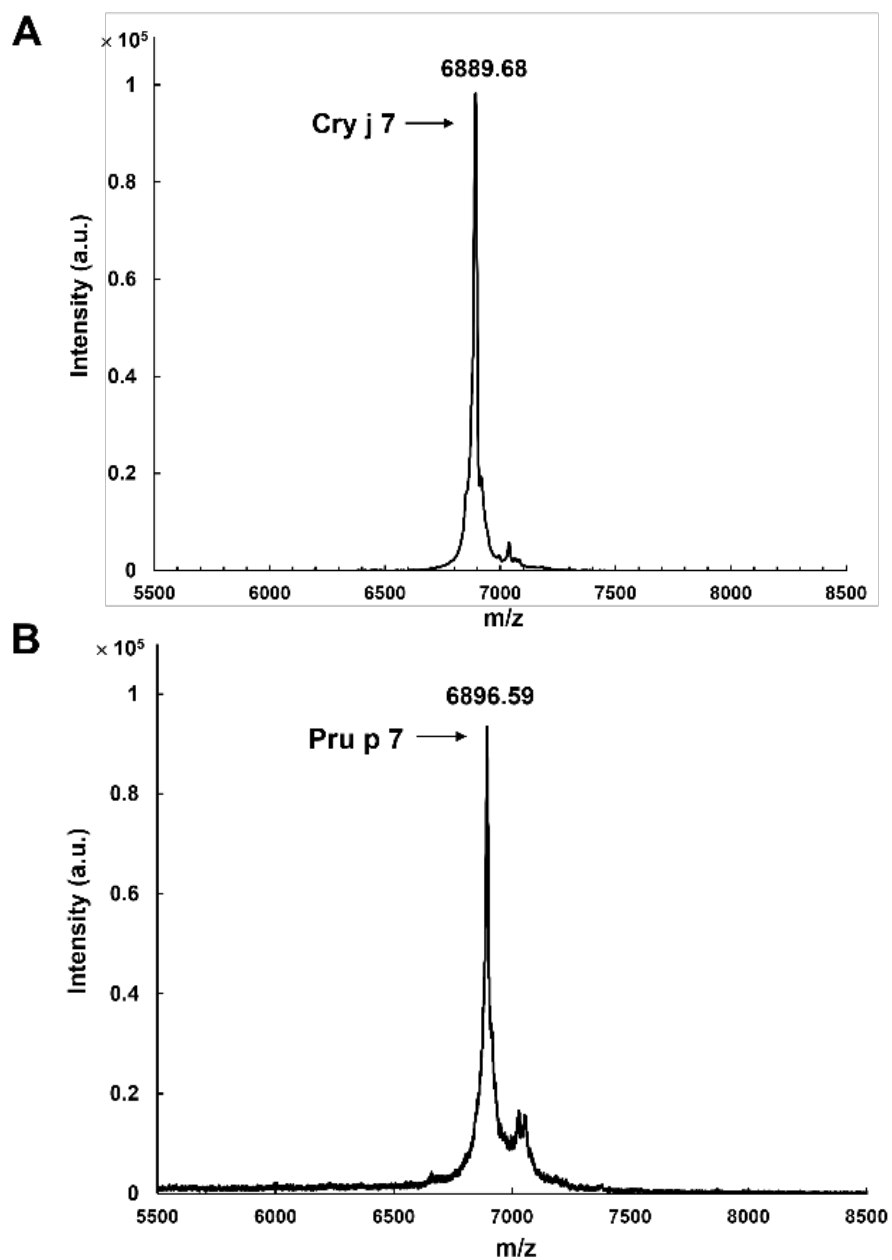


Figure 4. MALDI-TOF mass spectrum of the purified Cry j 7 (A) and Pru p 7 (B)

The far-UV CD spectrum of Cry j 7 revealed a broad negative band with double peaks at approximately 208 nm and 222 nm, along with a minimum at 211 nm (Figure 5A). In contrast to the reduced state, the oxidized Cry j 7 displayed a primarily α -helical structure, whereas reduced Cry j 7 apparently transformed into a typical random-coil pattern, indicating a loss of its secondary structure. The Pru p 7 demonstrated comparable CD findings as well (Figure. 5B), indicating similar secondary structure content and folding of the oxidized and reduced states of proteins.

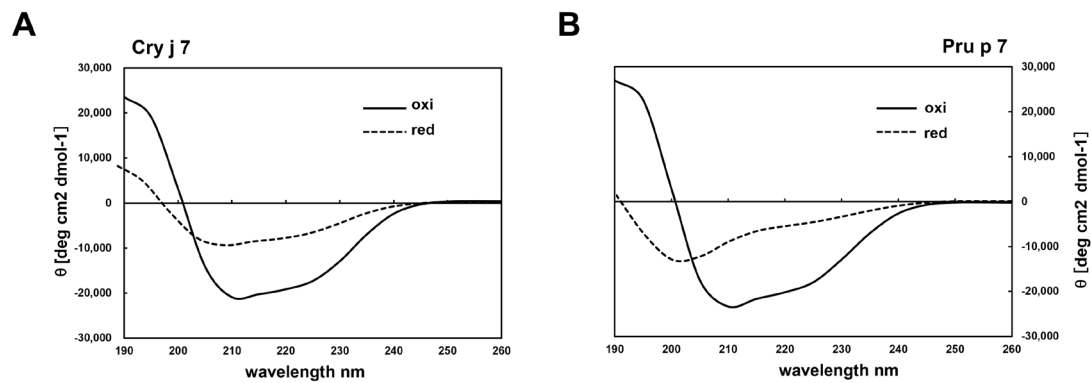


Figure 5. Circular dichroism spectra of Cry j 7 (A) and Pru p 7 (B). The proteins were dissolved in 50 mM PBS (pH 7.4) at 25 °C. Black solid line: oxidation state; black dotted line: reduction state.

Simultaneously, the ^1H -NMR spectrum of Pru p 7 was also identical to the native one, which was extracted from peach pulp and purified by RP-HPLC (Figure.6). It is well known that the NMR chemical shifts of proteins are quite sensitive to tertiary structures [53]. The results strongly suggested that the disulfide bonds and three-dimensional structure of the purified GRP were identical to those of the natural GRP, consistent with the results of previous other GRP studies.

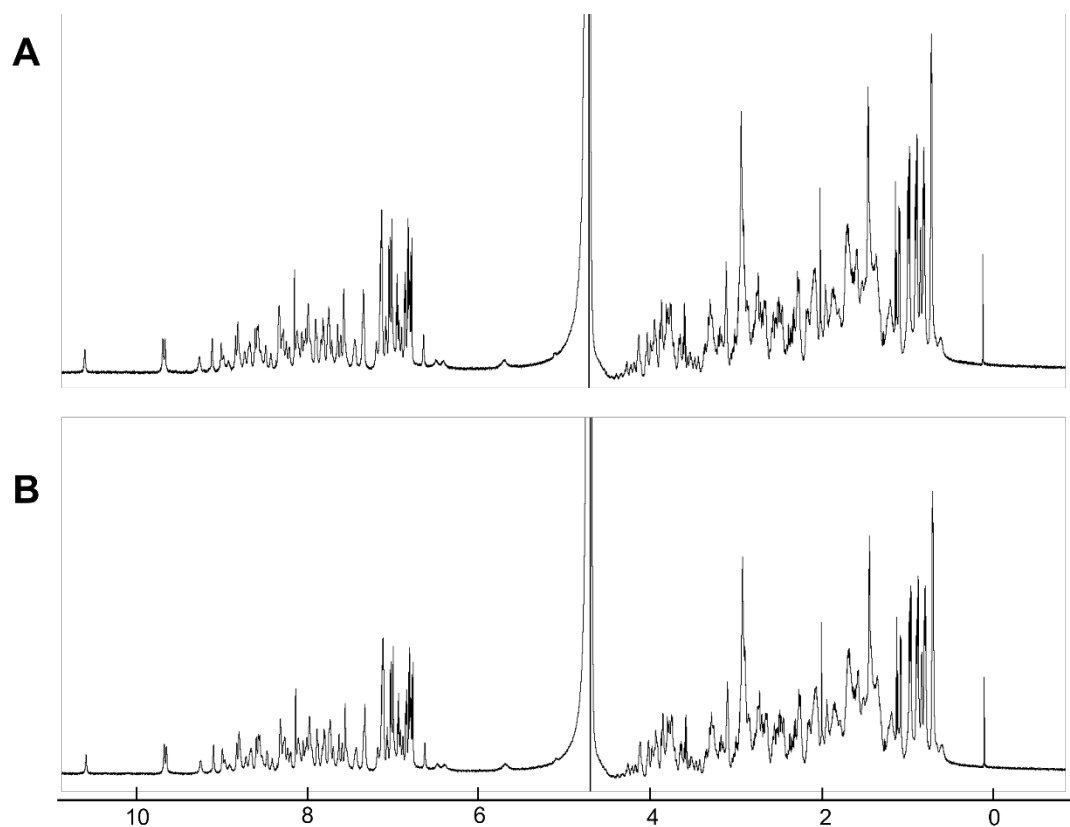


Figure 6. ^1H -NMR analysis of the recombinant Pru p 7 and native Pru p 7. The ^1H NMR spectra of recombinant Pru p 7 (A) and native Pru p 7 (B) in 50 mM potassium phosphate buffer, pH 6.0, with 10% D_2O , 298K.

3.3 Effects of pH and temperature on the secondary structure of GRP proteins

Allergens usually contain better stability in certain biophysical properties than non-allergens, for instance heat, pH, and chemical modification, etc. [54]. The pH and thermal stability of Cry j 7 and Pru p 7 were examined by CD spectroscopy. At various pH values (3.0, 6.0, 8.0, and 10.0), both Cry j 7 and Pru p 7 formed structures rich in α -helical structures at 20 °C. However, as the temperature increased, GRPs gradually underwent structural changes (Figure 7). Under acidic or neutral pH conditions, the pattern of the CD spectra did not change significantly even when the temperature increased, and when cooled to 20 °C, they re-turned to their original spectra. On the other hand, irreversible structural alterations were observed under basic pH conditions. In the case of Cry j 7 at pH 10.0, the profile of re-cooling at 20 °C displayed a similar result to that observed at 90 °C, maintaining the α -helix structure to a certain extent. For Pru p 7, the temperature values above 80 °C led to the protein denaturation, and the CD spectrum showed an unfolded structure which was not recovered after cooling down to 20 °C. These results show the differences in the stability of Cry j 7 and Pru p 7 again heating and show that Pru p 7 is particularly unstable under alkaline conditions. Additionally, the calculated denaturation rate of the helical structure (Figure S2) suggested that Pru p 7 showed a greater degree of secondary structure transformation than Cry j 7 did at different pH levels. Overall, the secondary structure stability of Cry j 7 was higher than that of Pru p 7.

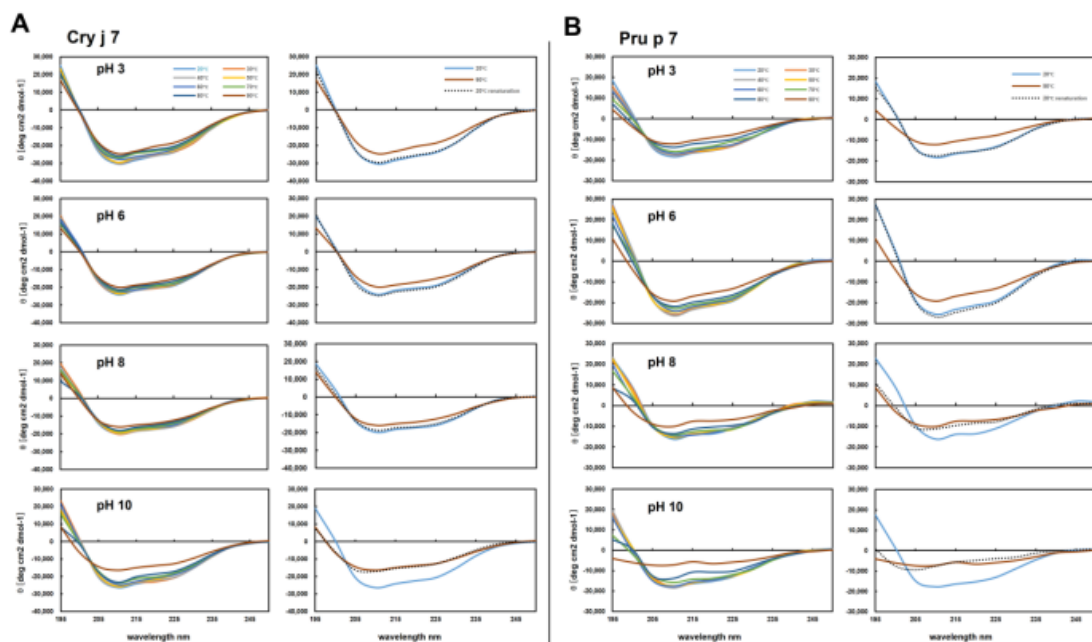


Figure 7. Secondary structural stability of Cry j 7 and Pru p 7 at different conditions, determined through circular dichroism. CD spectra of Cry j 7 (A) and Pru p 7 (B) heated in a stepwise manner in the range 20–90 °C and cooled back down to 20 °C at pH 3.0, pH 6.0, pH 8.0 and pH 10.0. The CD spectra at 20 °C, 30 °C, 40 °C, 50 °C, 60 °C, 70 °C, 80 °C, and 90 °C and after re-cooling to 20 °C (renaturation) are shown by light-blue, red, gray, orange, blue, green, dark-blue, dark-red, and black dotted lines, respectively.

3.4 NMR chemical shift assignment and three-dimensional structural analysis

^{15}N -isotope labeled Cry j 7 and Pru p 7 were expressed using the *P. pastoris* system in the BMM medium (Figure 8). The same purification procedure described in the supplemental materials and methods was used to purify 1.5 mg of Cry j 7 and 3.5 mg of Pru p 7 from 1 L BMM medium. NMR spectroscopy provides extensive protein structural information [55]. The wide dispersion of the chemical shifts of amide protons in the ^1H - ^{15}N HSQC spectra of oxidized Cry j 7 and Pru p 7 strongly suggested that these two proteins were correctly folded in the oxidized condition (Figure 9).

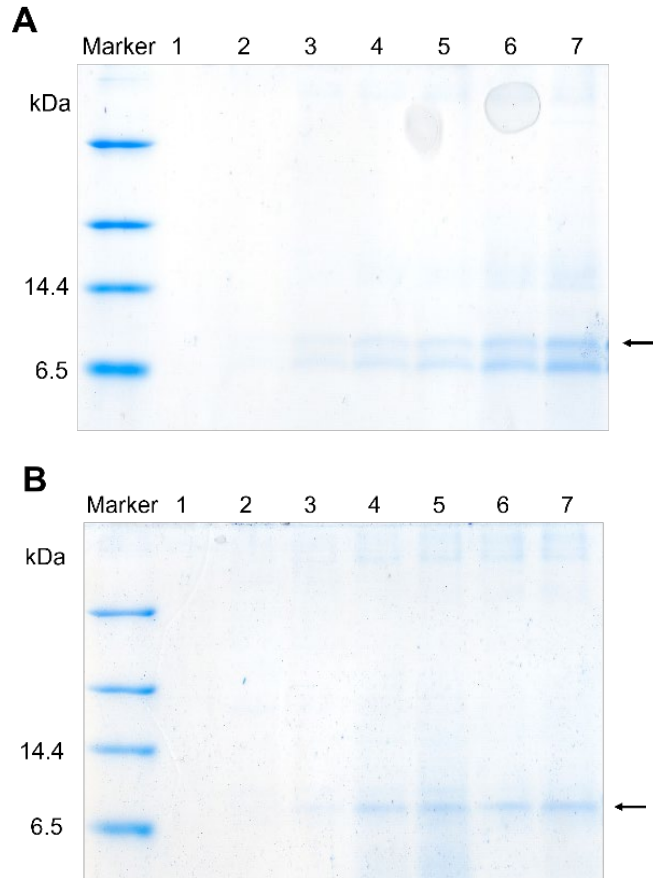


Figure 8. The expression and purification of ^{15}N -isotope labeled Cry j 7 and Prup7. Tricine SDS PAGE analysis of ^{15}N -isotope labeled Cry j 7 (A) and ^{15}N -isotope labeled Pru p 7 (B) in supernatants from *P. Pastoris* through baffled flask culture. Lane marker: protein molecular weight mark. Lanes 1-7: a total of 20 μL supernatant samples taken at 0, 12, 24, 36, 48, 60, 72 hours of methanol-induction, respectively.

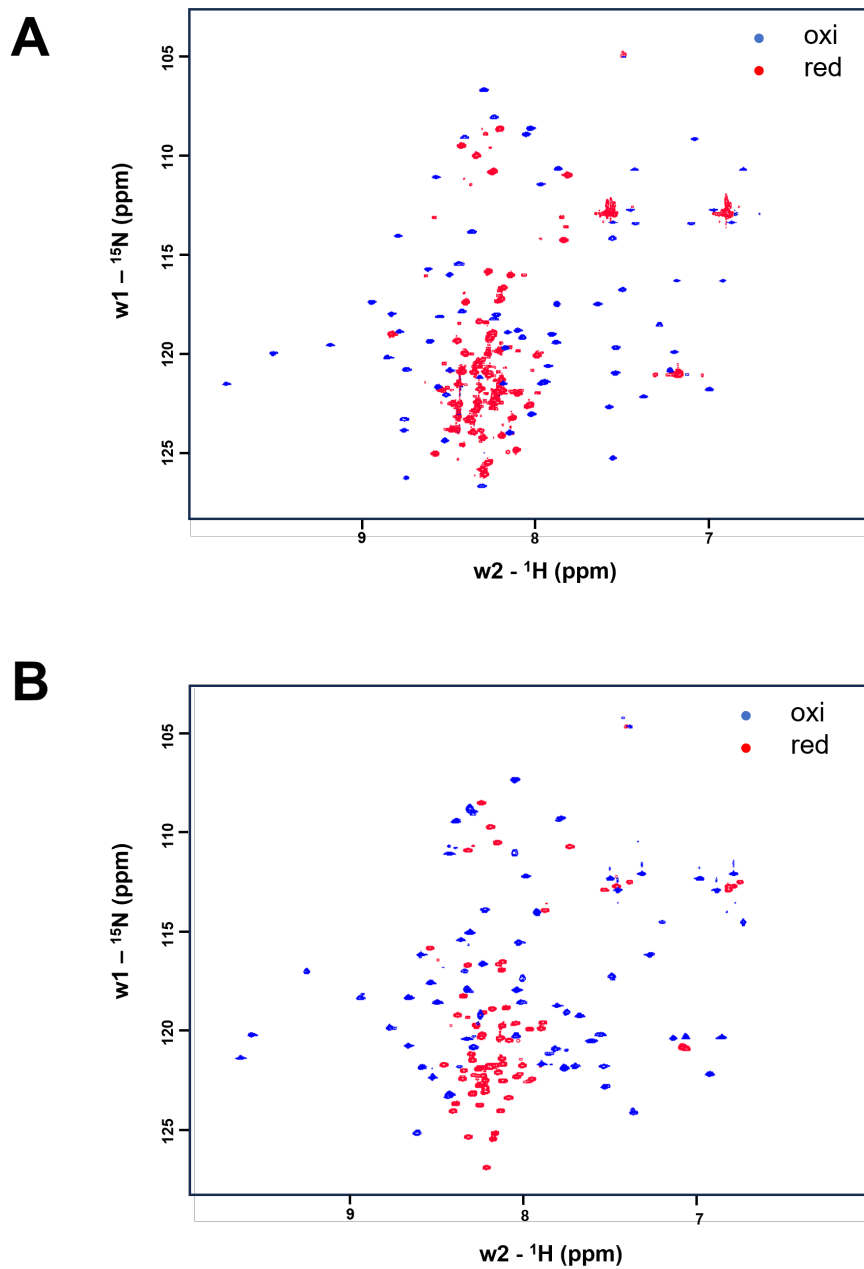


Figure 9. ^1H - ^{15}N HSQC NMR spectra of Cry j 7 (A) and Pru p 7 (B) in 10% $\text{D}_2\text{O}/\text{H}_2\text{O}$ solution, pH 3.0, 298K. blue signal points: oxidation state; red signal points: reduction state.

To investigate the three-dimensional structure of Cry j 7 and Pru p 7 in the aqueous solution, the chemical shift assignment was performed using the standard NMR experiments (Table S1). All backbone amide chemical shifts were assigned and visually represented in the ^1H - ^{15}N HSQC spectrum (Figure 10A and 10B). The set of 100 structures of Cry j 7 and Pru p 7 were calculated in the CYANA program, and the top 20 structures were extracted from the set based on the CYANA score function (Figure 10C and 10D). The structural calculation statistics are summarized in Table S2. The NMR structure and chemical shift information of Cry j 7 and Pru p 7 were deposited in the PDB with the codes 8X0R and 8X67, respectively. Both proteins formed well-converged structures rich in α -helices, in agreement with the CD data. The chemical shift values of the β -carbons confirmed that all cysteine residues formed disulfide bonds, which possibly contributed to the structural stability. Furthermore, the disulfide bond restriction, which was based on the X-ray crystallographic structure of another GRP, was evaluated in structural simulations to confirm the correct disulfide bond pairs between the 12 cysteine residues. The most energetically stable bond pairs were as follows: Cys5-Cys30, Cys9-Cys26, Cys13-Cys22, Cys29-Cys62, Cys33-Cys49 and Cys35-Cys47. The presence of a wide patch of positive charge in the C-terminal of the protein surface on the electrostatic potential surface can be observed (Figure S3A and S4A). The maps of molecular hydrophobic potentials on the surface of proteins (Figure S3B and S3B) indicated that Cry j 7 and Pru p 7, which were highly hydrophilic in the aqueous solution play an important role in maintaining protein conformation.

The three-dimensional structures of Cry j 7 and Pru p 7 closely resemble each other, with a confirmed RMSD of 0.83 Å. The heteronuclear steady-state $\{^1\text{H}\}-^{15}\text{N}$ NOE values revealed subtle differences between the helical segments of Cry j 7 and those of Pru p 7 (Figure 10E). A lower NOE value indicates greater flexibility in a particular region, and this was particularly evident in the P38/S38-C47 region, which is composed of α -helices and loops. In that specific region, Pru p 7 exhibited lower values, suggesting more mobility compared to the results of Cry j 7, according to the results for it. Cry j 7 demonstrated a more converged structure in aqueous solutions at pH 3.0 and at room temperature, which was probably also related to this feature..

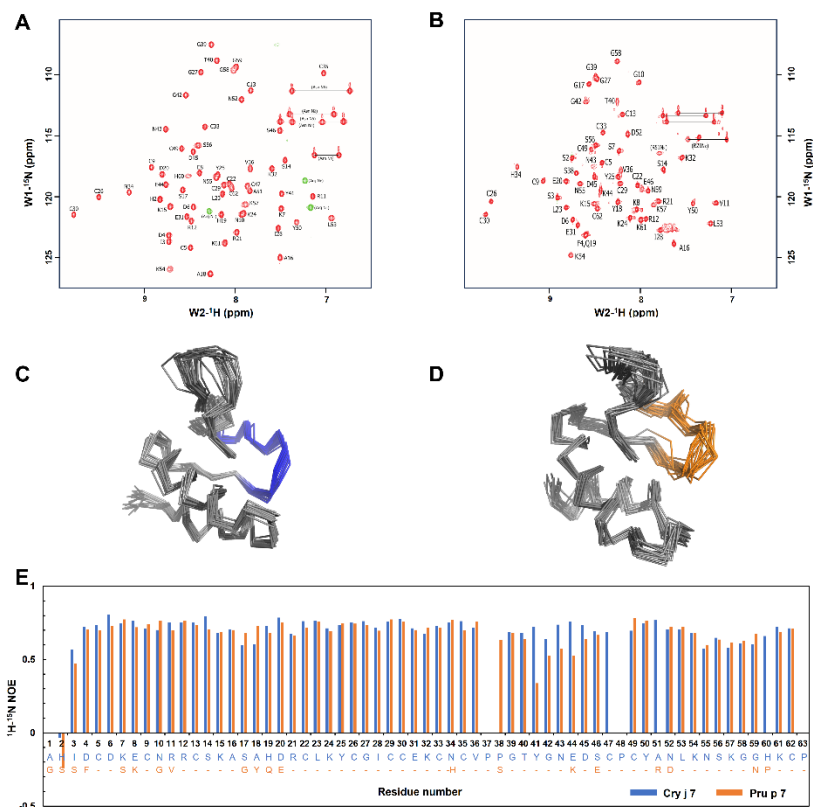


Figure 10. NMR investigation of Cry j 7 and Pru p 7. (A) ^1H - ^{15}N HSQC NMR spectrum assignment of Cry j 7 in 10% $\text{D}_2\text{O}/\text{H}_2\text{O}$ solution at pH 3.0, 25 °C. Asn N ϵ and Arg N ϵ are peaks from side chains. (B) ^1H - ^{15}N HSQC NMR spectrum assignment of Pru p 7 in 10% $\text{D}_2\text{O}/\text{H}_2\text{O}$ solution at pH 3.0, 25 °C. R21 N ϵ and R51 N ϵ are peaks from side chains. NMR structure of main chain of Cry j 7 (C) and Pru p 7 (D) in aqueous solution and overlay of the ensemble of 20 final energy-minimized CYANA structures. The upper side is the C-terminus. (C) Gray: NMR structure of Cry j 7; blue: P38-C47 amino acid sequence. (D) Gray: NMR structure of Pru p 7; orange: the low-NOE region, i.e., the S38-C47 amino acid sequence. (E) Comparison of the heteronuclear steady-state ^1H - ^{15}N nuclear Overhauser effect (NOE) values for Cry j 7 (blue) and Pru p 7 (orange).

3.5 ^1H NMR analysis of thermally denatured GRP structural stability

The signal of one-dimensional NMR can reflect the folding state of proteins; thus, the structural stability of proteins can be evaluated. This enables a reliable distinction between various folding states of proteins and a discernment of subtle differences in structural stability [56]. Therefore, ^1H -NMR was employed to assess the structural stability of Cry j 7 and Pru p 7 following thermal denaturation at acidic an acidic pH levels, pH 3.0 (Figure S5). As a result, the spectra of Cry j 7 and Pru p 7 exhibited distinct change in chemical shifts in the amide region at 25 °C and 75 °C, although the chemical shifts obtained while cooling them down to 25 °C aligned well with the initial spectrum at 25 °C, signifying that Cry j 7 and Pru p 7 renatured to the same structural conformation as that before the thermal denaturation at 75 °C (Figure 11 A and B). In the case of Cry j 7, the peak at a lower field than 10 ppm disappeared as the temperature reached 75°C, and its absence in ^1H - ^{15}N HSQC suggests that it is not an NH group. In our signal assignment, it was discovered that the side chain related to Y50 moved into the hydrophobic core of the structure (Figure 11 C and D). This hydrophobic core protected the OH proton, resulting in a slower proton exchange rate, which allowed the OH signal to be easily detected in the ^1H NMR spectrum. Besides the peaks around 9-10 ppm, the signals of C26 and C30 in the NH region also clearly changed. In particular, the peak signal intensities in Cry j 7 weakened whereas the signal in Pru p 7 vanished. NOE signals were observed between R21 H ϵ and P63 H γ for Y50 OH (H η), suggesting that they are located close to each other. Therefore, the mobility of this region in Pru p 7 was likely suggested to be increased under heating to 75 °C. It is also

suggested that the structural stability of disulfide bonds between C26 (distance 3.0 Å) and C30 residues (distance 4.9 Å) close to the region was affected. As compared to Cry j 7, the conformation of Pru p 7 might have changed more as temperature increased, leading to a high exchange rate with water molecules, affecting signal detection in solutions. In general, the structural stability of Cry j 7 was higher than that of Pru p 7 as shown by thermal denaturation experiments at a low pH value.

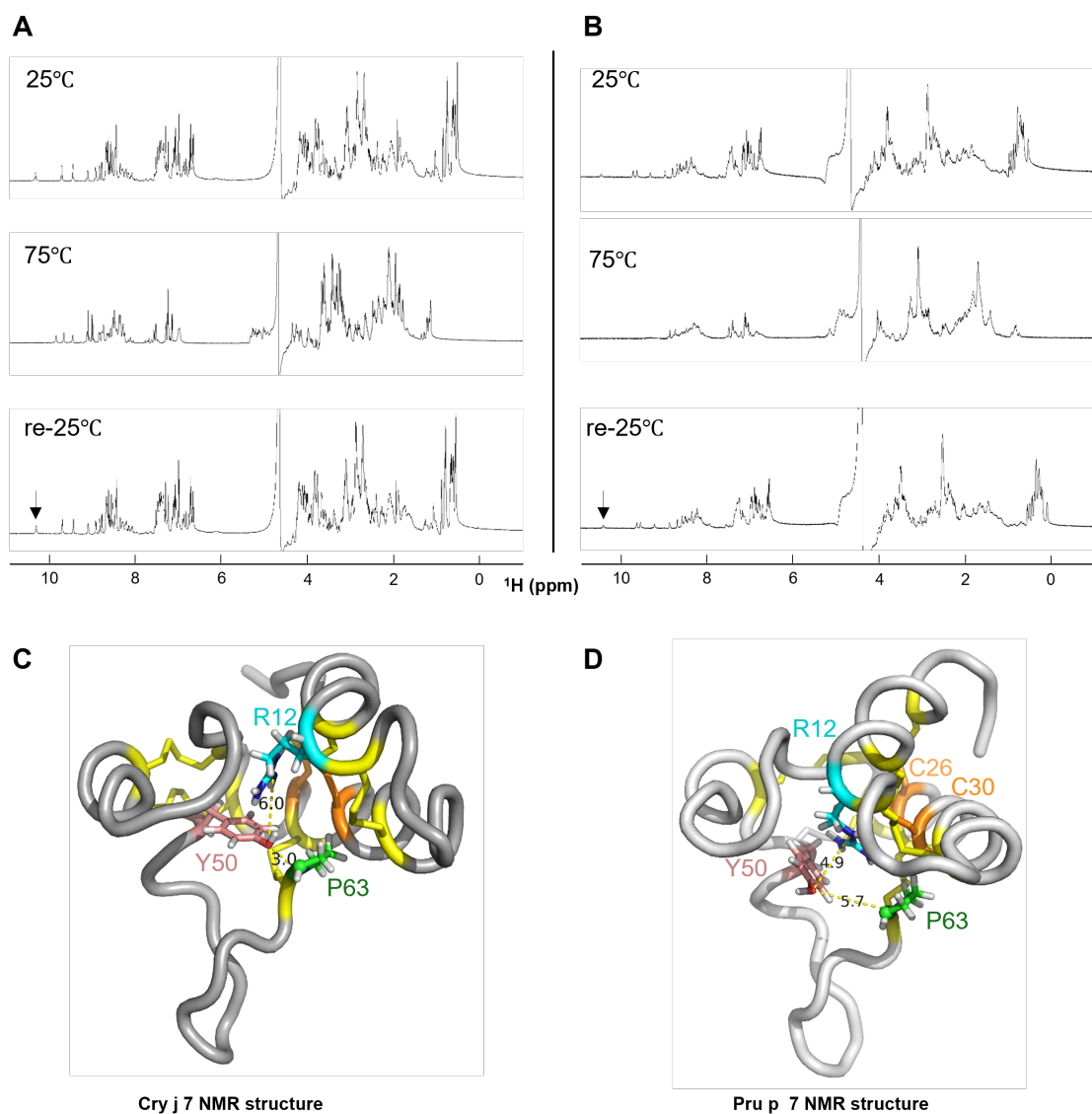


Figure 11. ^1H -NMR analysis of the Cry j 7 and Pru p 7 in 10% $\text{D}_2\text{O}/\text{H}_2\text{O}$ solution (pH 3.0) at various temperatures. The ^1H NMR spectrum of Cry j 7 (A) and Pru p 7 (B) at 25 °C, and 75 °C, and when re-cooled down to 25 °C. The arrow points to the peak of Tyr-OH. The atom distance (Å) between Y50 (pink color) and R12 (blue color)/P63 (green color) present in the NMR structure of Cry j 7 (C) and Pru p 7 (D). Gray: main chain. Yellow: disulfide linkage. Orange: C26 and C30.

3.6 Proteolytic degradation of Cry j 7 and Pru p 7

Proteolysis is a valuable method with which to investigate protein dynamics and stability [57]. For this study, common thermolysin, a common enzyme, was chosen to examine the proteolysis kinetics of Cry j 7 and Pru p 7 at 30 °C, and Snakin-1 (GRP from potato) was used as a control. In contrast to the result, showing that Pru p 7 had one band, the appearance of the second band of Cry j 7 indicated that its native structure was digested into different fragments, and the primary band of Snakin-1 was the only one to vanish by the end of the experiment (Figure 12A). All three proteins exhibited slow degradation, showing a decrease in half-integrated intensity over several hours (Figure 12B).

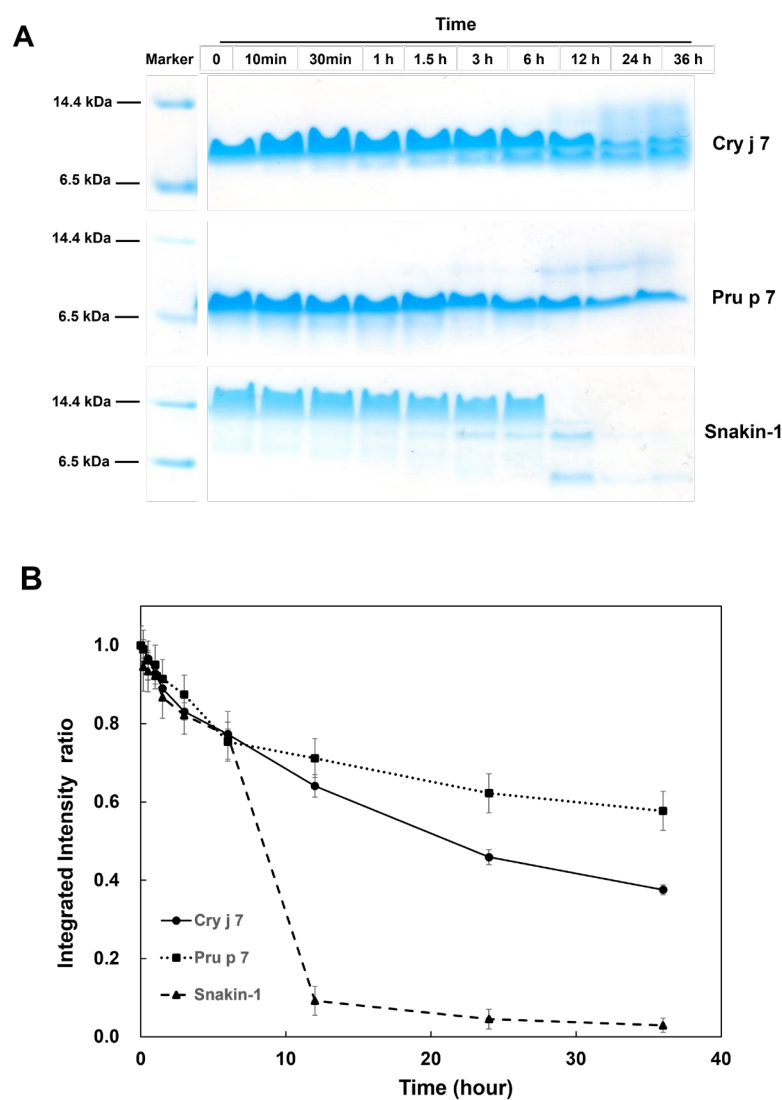


Figure 12. Kinetics of proteolysis of Cry j 7 and Pru p 7 (0.5 mg/mL) at 30 °C in the presence of 0.1 mg/mL thermolysin. (A) Representative analysis by tricine SDS PAGE gels. Lane marker: protein molecular weight mark. (B) Proteolysis kinetics. The data for Cry j 7, Pru p 7, and Snakin-1 are shown in circles, squares, and triangles, respectively.

The three proteins demonstrated different degrees of resistance to thermolysin, while the peptide Snakin-1 was shown to be the most sensitive, followed by Cry j 7 and Pru p 7. The resistance of GRP proteins to thermolysin revealed that Cry j 7 and Pru p 7 were almost completely resistant under these conditions, suggesting their high stability against proteolysis.

3.7 Stimulation of gastrointestinal digestion of GRP proteins

Food allergens generally have high structural stability, which allows them to retain an almost native structure and expose the surface epitopes even under environmental conditions capable of unfolding most globular proteins [47]. In our study, the peach GRP allergen Pru p 7 and the potato GRP peptide Snakin-1 demonstrated remarkable stability against gastric digestion by pepsin. Only one evident band was revealed by tricine SDS-PAGE, even after 2 hours of digestion (Figure 13A). Subsequent treatment mimicking intestinal digestion resulted in effective Snakin-1 degradation, with the emergence of the digestion products and near-destruction after 40 hours of reaction. In contrast, Pru p 7 was able to maintain an intact structure throughout the enzymatic degradation process, revealing that it could be highly stable in gastrointestinal digestion, with similar results to native Pru p 7 [48]. However, Cry j 7 is a pollen allergen that causes cross-reactivity with Pru p 7, sharing the same epitope regions. Therefore, Cry j 7 was also evaluated by gastrointestinal digestion for comparison. Following 4 hours of trypsin and chymotrypsin enzyme digestion, a new band with a molecular weight of approximately 6.5 kDa appeared and remained consistent in the subsequent hours.

The results of the gastrointestinal digestion of Cry j 7 and Pru p 7 were also analyzed by RP-HPLC (Figure S6) and mass spectrometry (Figure S7). For Cry j 7, the peak in the chromatograms at 27 min, corresponding to the intact form, was detected during digestive digestion by pepsin. After 4 h of trypsin and chymotrypsin digestion, an additional peak at 30 min corresponding to the proteolytic product was observed.

Mass spectrometry analysis also revealed a peak at 27 min in all the samples, belonging to intact Cry j 7 stabilized by six disulfide bonds (m/z 6890.8). It also appears that one internal peptide bond was cleaved, i.e., the carboxyl end of K (15, 57 or 61), and a molecule (m/z 6906.4) with added water was produced. The other produce, which was most likely the disulfide-stabilized protein without an amino acid sequence from K54 to K57 (Figure 13B), with a molecular weight of m/z 6451.9, was present at the peak at 30 min after 4 hours of intestinal digestion (Figure S7A). For Pru p 7, a peak at 33 min was observed following intestinal digestion, and it was difficult to separate from the peak at 34 min, which was associated with gastric digestion and maintained the intact protein structure (m/z 6902.4). At the same time, a protein fragment with m/z 6917.5 was detected in the peak at 33 min (Figure S7B), implying that one internal amino acid in the loop area was digested, such as Y41 and K (54, 57 or 61). Y41 might be a cleavage site, given the reduction of the trypsin digestion fragment (Figure 13C).

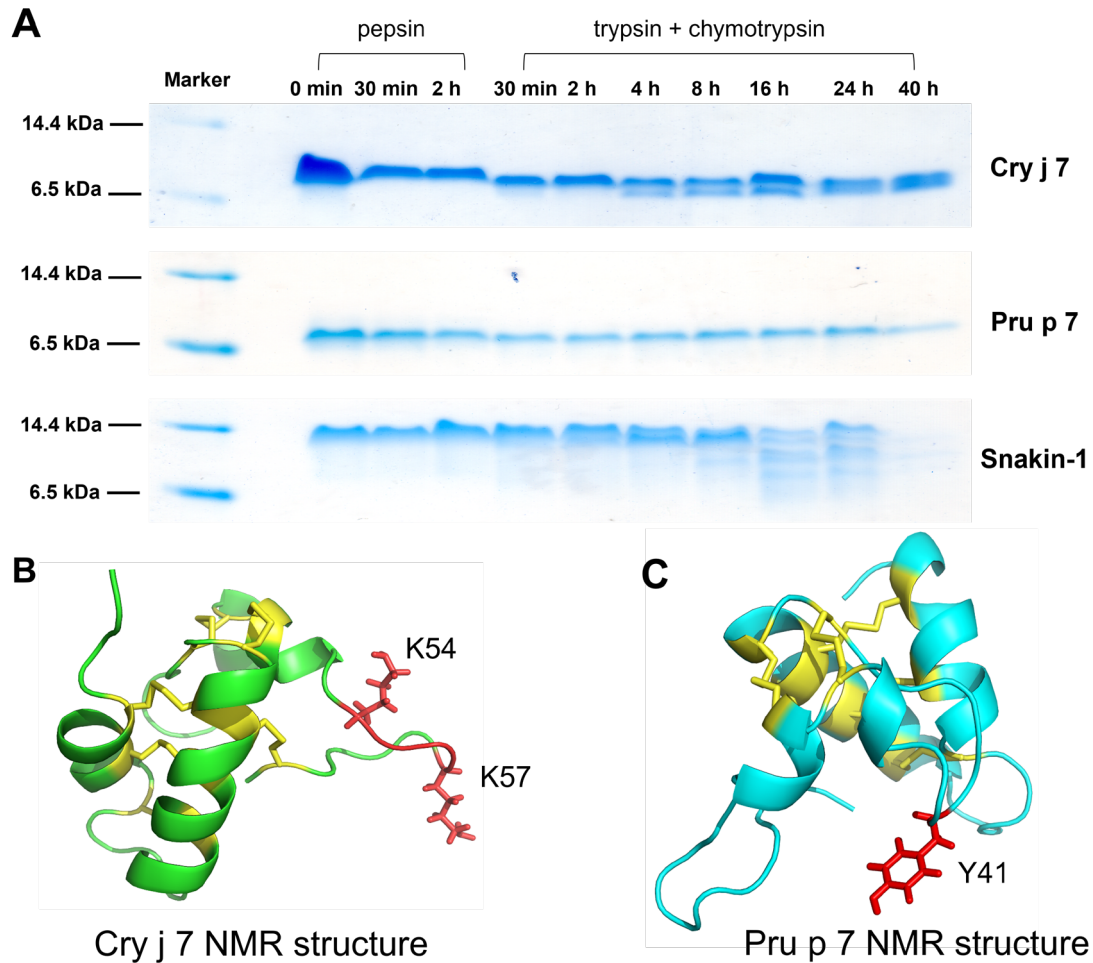


Figure 13. The simulated gastrointestinal digestion of Cry j 7 and Pru p 7. (A) Gastrointestinal digestion detected by tricine SDS PAGE. Pepsin digestion was performed in a pepsin-to-protein mass ratio of 1:20, and subsequent trypsin/chymotrypsin digestion was performed in a ratio of 1:50. Lane marker: protein molecular weight. Cleavage sites of trypsin/chymotrypsin digestion were presented in the NMR structure of Cry j 7 (B) and Pru p 7 (C). Green: Cry j 7 NMR structure. Blue: Pru p 7 NMR structure. Red: cleaved residues. Yellow: 6 disulfide bridges.

GRP allergens exhibited sufficient structural stability in resisting degradation by gastrointestinal enzymes for a long time, which pointed to Pru p 7 as a food allergen that could elicit immune responses through its conformational epitopes. In addition, Cry j 7 lost a short peptide (K54-K57) in its loop region and showed greater sensitivity to pepsin or trypsin and chymotrypsin digestion in contrast to Pru p 7.

3.8 Endosomal degradation of GRP proteins is mediated by Cathepsin S

The capacity of potential allergens to withstand endosomal digestion can reflect the sensitization process, providing essential details about the structural stability of allergens [54]. Thus, to mimic endosomal degradation, the Cry j 7 and Pru p 7 proteins were subjected to digestion by Cathepsin, a lysosomal cysteine protease that acts on substrates with glutamine or lysine [58]. The proteins were fragmented into several pieces when incubated in a reaction buffer containing DTT at 37 °C (Figure 14). After 2 hours of digestion, an obvious low-molecular weight band appeared in Cry j 7, indicating that was digested into further fragments, while Pru p 7 displayed lower degradation. All proteins were observed to be thoroughly digested after 12 hours. Cry j 7 and Pru p 7 took a longer time for half-life degradation compared to Snakin-1, showing better stability. It was possible that they retained a small proportion of their ordered secondary structure to resist degradation under reduced conditions (Figure 5). In the absence of DTT, the proteins preserved a relatively intact band even after 48 hours of degradation (Figure. S8). These results revealed that the oxidized proteins exhibited a high ability to resist digestion, whereas the reduced proteins lost their stable structure, and were more susceptible to breakdown. In summary, the sensitization of Cry j 7 was higher than that of Pru p 7 in Cathepsin S enzyme digestion.

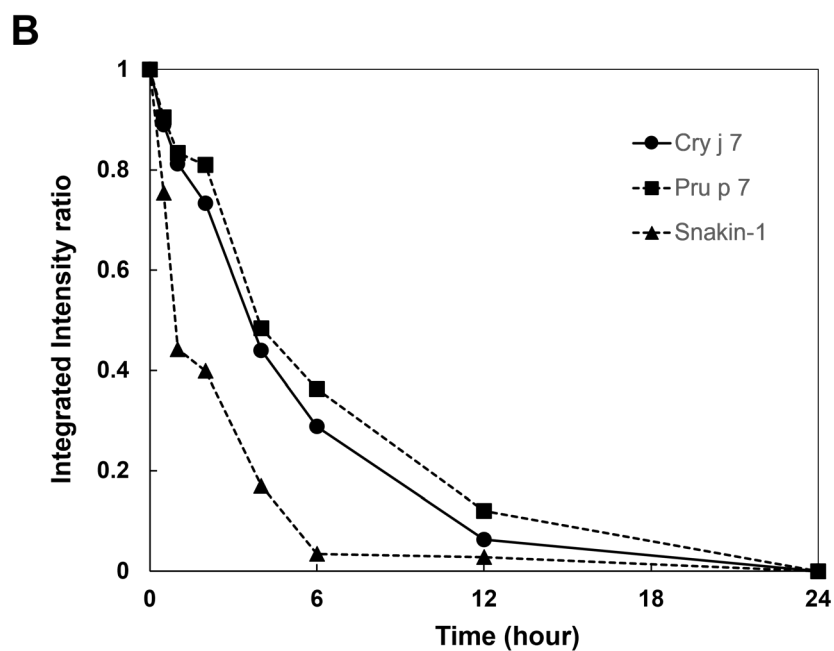
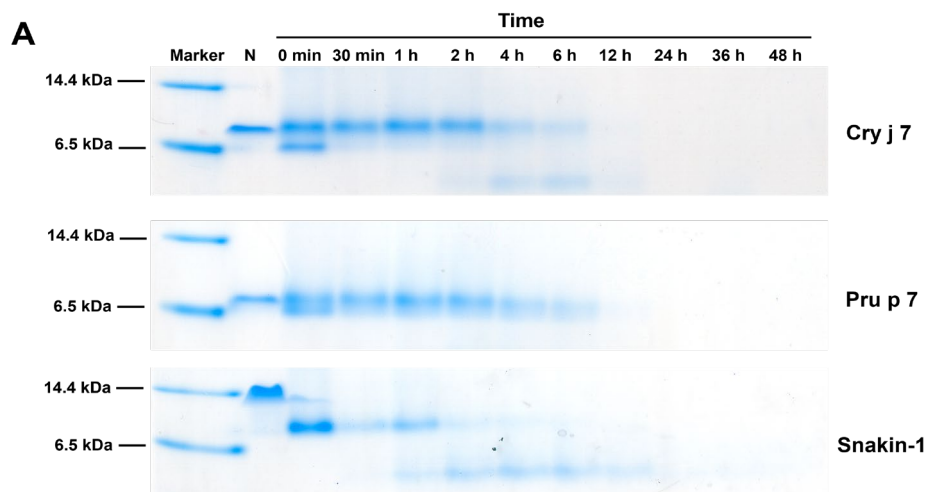


Figure 14. Degradation analysis of Cry j 7 and Pru p 7 in vitro. (A) Endosomal degradation assays by Cathepsin S enzyme in sodium acetate buffer with DTT, and the degradation profile was analyzed by tricine SDS PAGE. Lane marker: protein molecular weight marker. Lane N: sample without DTT and Cathepsin S enzyme. (B) Proteolytic stability analysis: Symbols are Cry j 7 (circles), Pru P 7 (squares), and Snakin-1 (triangle).

After RP-HPLC analysis (Figure S9), many of the degradation fragments of Cry j 7 and Pru p 7 were determined to correspond to the amino acid sequence regions: I3/S3-R12, K24-V36, and L53-K61, based on the analysis of the MS data (Figure S10). The flexible loop regions of GRPs were inferred as the primary locations of Cathepsin S cleavage. Simultaneously, we used the ProPred web tool and IEDB to predict major MHC II binding regions of target allergens to identify T-cell epitopes. The most likely binding location in Cry j 7 and Pru p 7 among these epitopes was the sequence Y50-G58, which connected to HLA-DRB1*04:01 (Table 1). This obtained MHC II binding region may correspond to the sequence of degradation fragments, suggesting that relevant immune responses of the GRP proteins might be mediated by the epitope sequence of Y50-G58.

Table 1 (A). T-cell epitopes prediction of Cry j 7 allergen

HLA type	Range	T-cell epitope	ProPred	IEDB
DRB1*01:01	41-49	YGNEDSCPC	+	+
	50-58	YANLKNSKG	-	+
DRB1*03:01	3-11	IDCDKECNR	-	+
DRB1*04:01	41-49	YGNEDSCPC	+	+
	50-58	YANLKNSKG	+	+
DRB1*07:01	50-58	YANLKNSKG	-	+
DRB1*08:01	23-31	LKYCGICCE	+	-
DRB1*11:01	50-58	YANLKNSKG	-	+
DRB1*13:01	23-31	LKYCGICCE	+	-
DRB1*15:01	50-58	YANLKNSKG	-	+
DRB5*01:01	50-58	YANLKNSKG	-	+
	53-61	LKNSKGGHK	+	+

Table 1 (B). T-cell epitopes prediction of Pru p 7 allergen

HLA type	Range	T-cell epitope	ProPred	IEDB
DRB1*01:01	25-32	YCGICCEKC	+	-
	50-58	YRDLKNSKG	+	+
DRB1*03:01	11-19	VRCSKAGYQ	-	+
	53-61	LKNSYGNPK	-	+
DRB1*04:01	25-32	YCGICCEKC	+	-
	50-58	YRDLKNSKG	+	+
DRB1*07:01	53-61	LKNSYGNPK	-	+
DRB1*08:01	41-49	YGNKEDCPC	+	-
DRB1*11:01	11-19	VRCSKAGYQ	+	-
	50-58	YRDLKNSKG	-	+
DRB1*13:01	23-31	LKYCGICCE	+	-
DRB1*15:01	50-58	YRDLKNSKG	-	+
DRB5*01:01	23-31	LKYCGICCE	+	-
	53-61	LKNSYGNPK	-	+

*Our analyses focused on the 9 common human alleles. The identified sequence length for the IEDB prediction method is 15, we extended the input sequence when its length was shorter than 15 residues.

3.9. Epitope prediction of Cry j 7 and Pru p 7

The binding of IgE antibodies to specific allergen epitopes is an essential stage in the onset of allergic responses. To assess the IgE epitope prediction, we examined publicly accessible epitope tools, selecting B-cell epitope prediction methods provided by IDEB. Based on amino acid sequences, Bepipred 3.0 predicted two linear epitopes regions: S14-R21, and N52/P52-K61 (Figure 15). In general, Bepipred is more effective at predicting nonspecific epitopes. Therefore, 3D structures of allergic proteins were used for predicting specific IgE epitopes with higher accuracy [49].

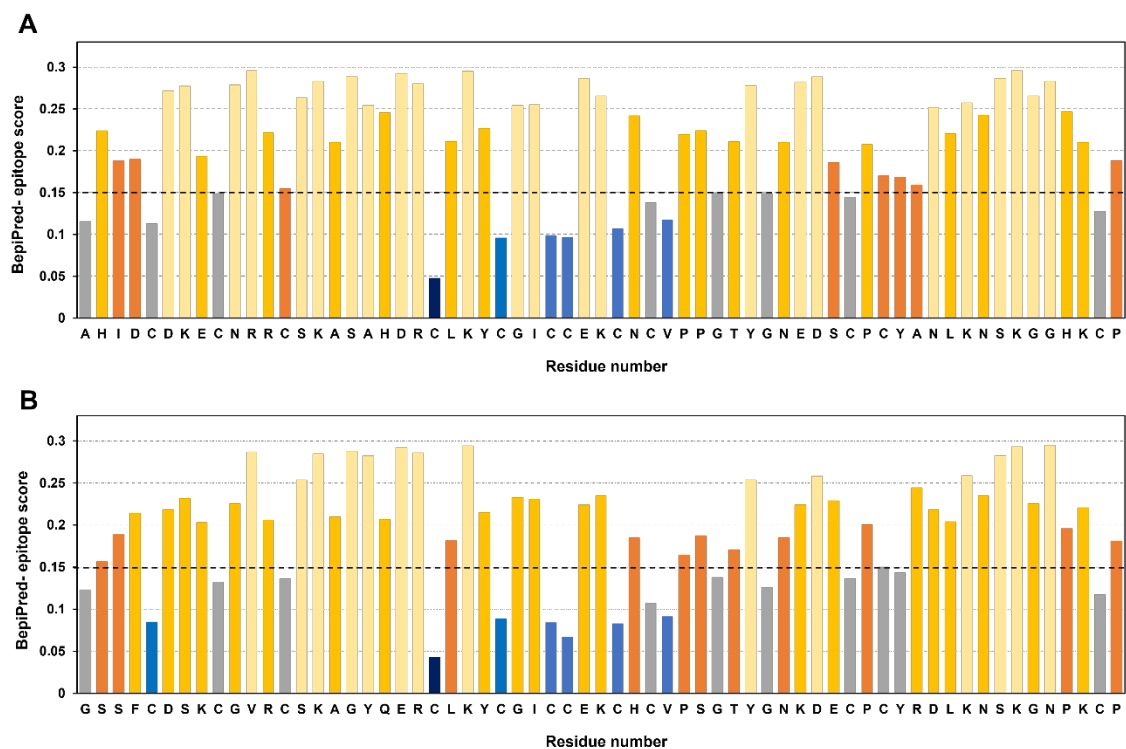


Figure 15. B-cell epitope prediction using the BepiPred-3.0 online software. (A) The B-cell epitope prediction graphic of pollen allergen Cry j 7. (B) The B-cell epitope prediction graphical of food allergen Pru p 7. In these diagrams, the x-axis and y-axis are protein sequence positions and BepiPred -3.0 epitope scores. Residues with higher scores are more likely to be part of a B-cell epitope. In ascending order: dark blue color (0-0.5), blue color (0.5-1.0), gray color (1.0-1.5), orange color (1.5-2.0), gold color (2.0-2.5), and light gold color (more than 2.5). The gray dotted line: threshold (0.15).

Software DiscoTope 2.0 can be employed to predict the epitopic areas of GRP proteins. Despite the amino acid sequence being conserved, the predictions for these three potential epitope regions will differ due to their prediction based on the three-dimensional structure from NMR analysis (Figure 16). Three distinct epitopic regions were identified in Cry j 7 and Pru p 7. These regions were located as follows: the loop in the C-terminal region (around K54 to G59/N59, the upper right part of the molecule in the 3D model), and the loop region at the bottom of the molecule (Y41 to N43) and top part of the molecule (K15 to A18/Y18). Additionally, these three regions displayed greater flexibility in MD stimulation (Figure S11) and NOE results (Figure 10E) compared to other locations.

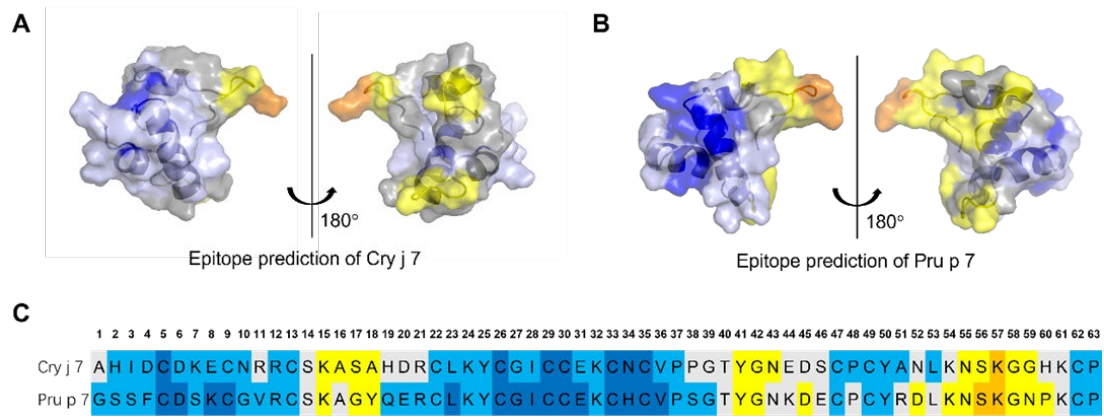


Figure 16. The epitope predictions for Cry j 7 and Pru p 7. Three-dimensional structure of Cy j 7 (A) and Pru p 7 (B) inferred by the NMR experiment. At least three conformational epitope regions were predicted using the software DiscoTope 2.0. They are highlighted in yellow and orange. (C) Sequences of Cry j 7 and Pru p 7 with predicted AA involved in epitope regions. The color codes correspond to DiscoTope propensity scores. The higher the score, the higher the propensity to be an epitope. In following colors represent the values in descending order: blue, -20 or less; light blue, -20 to -15 ; gray, -15 to -12.5 ; yellow, -12.5 to -5.0 ; and orange, greater than -5.0 .

4. Discussion

Many organizations around the world are engaged in molecular allergology research, to acquire insight into the sensitization process of food and inhaled allergens. Following recent research, it has been shown that food allergies associated with pollen allergies referred to as pollen-food allergy syndrome (PFAS), have become a public health concern as the prevalence has increased year by year. As the GRP allergens responsible for PFAS, the allergenicity and sensitization procedures for Cry j 7 and Pru p 7 are of great importance in investigations. Notably, we found that the allergic potential of GRP was not exclusively dictated by its quantitative content in the cypress family or the frequency of food consumption. Additional factors, encompassing allergen physicochemical properties, their biological activity, effects of structural stability and dynamics under stressful situations, as well as a patient's individuality, play an important role in comprehensively exploring the mechanism of PFAS [11,17]. As research progresses, a more in-depth understanding of these factors should be developed to inform enhanced diagnostic and therapeutic strategies for PFAS.

In NMR research, the solution structures of Cry j 7 and Pru p 7 were determined. This was the first report of the NMR structural characterization of the Japanese cedar pollen allergen Cry j 7, and it also improved the intact three-dimensional structure of the peach allergen Pru p 7. Compared with the previously published three-dimensional structural model of Pru p 7 using a natural sample [48], we performed NMR experiments on stable isotope-labeled GRP proteins to obtain more accurate three-

dimensional structure and flexibility information. In our experiments, all back-bone resonances of Pru p 7 were completely assigned, and the presence of two brief helical stretches encompassing residues K44-E46 and P48-D52 were also identified in the C-terminal region. Furthermore, we determined the proper connection of six disulfide linkages among 12 cysteine residues. The disulfide bond between C13 and C22 was the same as that found in Pru p 7 NMR analysis, but C26 showed a tendency to form disulfide bond with C9 rather than C62 based on the CYANA calculation. Fortunately, the disulfide bridges involving C47, C30, C33, and C49 [48] were specifically distinguished: Cys5-Cys30, Cys33-Cys49, and Cys35-Cys47. The Cry j 7 and Pru p 7 displayed high similarity in mainly helical structures in their N-terminal region but showed differences in the C-terminal flexible region according to the NMR results. In ^1H - ^{15}N hetero-nuclear steady-state NOE, the significant difference in the P38/S38-C47 region revealed the high mobility of Pru p 7 and low mobility of Cry j 7, which corresponded to the molecular dynamic simulation results (Figure S11). In these flexible regions composed of loops and a short α -helical structure, there are differences in charge residues and hydrogen bonds, and these might affect the movement trajectories of atoms [59]. Furthermore, the impact of significant differences in the flexibility of this region on immunogenicity to allergens necessitates further exploration. These results are consistent with the results of thermal denaturation experiments using CD and ^1H -NMR (Figures 7 and S5), and clearly support the structural characteristics of Cry j 7, which is more stable than Pru p 7 in terms of physical chemistry. In particular, the finding that heating under alkaline conditions promotes the irreversible structural

denaturation of Pru p 7 in contrast to what is observed for heating under acidic conditions is important when considering the removal of allergenicity through processing.

As is well known, stability in the digestive tract is an important factor related to the presentation of food allergen epitopes [60]. In contrast to pollen allergens, which are sensitized through the respiratory route, many food allergens are known to be more resistant to pepsin, a feature of a classical or true food allergen [61]. Interestingly, the degradation by enzymes, which is thought to occur in the digestive tract shows the opposite trend to thermal denaturation, and Pru p 7 was more stable. Therefore, the difference in stability between these digestive enzymes may be mainly due to the specificity of the enzymes rather than structural stability. It is also worth noting that the results of the conformational epitope prediction for B cells include flexible loop regions that are susceptible to enzymatic cleavage (Figures 13 and 16), although further IgE immunological experiments are still required to verify the true epitope region.

Sensitizations occur via the cutaneous or airway exposure of allergens, or when food allergens enter the human gastrointestinal tract to induce the immune response. In this study, we also evaluated the resistance to endosomal degradation associated with sensitization [54]. The characteristics of sensitization to allergens are strongly associated with endosomal digestion, the generation of allergen-derived peptide, and MHC II presentation [62]. Therefore, our investigating of the stability of Cry j 7 and Pru p 7 against endosomal degradation was expected to provide valuable insights into

the sensitization in the GRP family. The observations in this work indicated that Cry j 7 and Pru p 7 were easily digested by Cathepsin S enzyme in the reduction state. This allowed the antigenic peptides to be effectively coupled with MHC II, inducing the Th2 response and promoting the production of specific IgE by B-cells [63]. However, the structure of Cry j 7 and Pru p 7 in the oxidation state remained hyperstable (Figure S8). The kinetics of proteolytic degradation were significantly influenced by the stability of antigens, which determined the density of appropriate antigenic peptides / MHC II complexes and eventually, immunogenicity [64]. In certain instances, allergens with greater proteolysis resistance led to the lower production of T-cell epitopes, causing a decrease in immunogenicity [65]. Therefore, the efficiency of reduction of the GRP family after endo-cytosis may also be an important factor related to sensitization.

5. Conclusion

Analysis of the three-dimensional structure and thermal stability using CD and NMR in this study revealed that Cry j 7 is more stable than Pru p 7. The differences in the mobility of the characteristic loop regions of the two proteins were clarified, and it was suggested that these loop regions not only contribute to stability but may also be epitope sites. In contrast, Pru p 7 had higher resistance to degradation by digestive enzymes, and analysis of the degradation fragments suggested that the difference in the primary sequence contributed to the resistance. In addition, although the resistance of endosomal enzymes involved in the establishment of sensitization was greatly reduced by the reduction in disulfide bonds in both proteins, the fact that the loop region in the native structure was cleaved to some extent preferentially suggests that the remaining structure after reduction may also have been related to the resistance. A detailed comparison of the structure and stability of the GRP family derived from pollen and fruit in this study will be useful as basic information for elucidating the mechanisms of sensitization and the development of allergies in PFAS in the future.

6. References

1. K. Mao, R.I. Milne, L. Zhang, Y. Peng, J. Liu, P. Thomas, R.R. Mill, S.S. Renner, Distribution of living Cupressaceae reflects the breakup of Pangea, *Proc. Natl. Acad. Sci. U. S. A.* 109 (2012) 7793–7798. https://doi.org/10.1073/PNAS.1114319109/SUPPL_FILE/ST06.
2. D. Charpin, C. Pichot, J. Belmonte, J.P. Sutra, J. Zidkova, P. Chanez, Y. Shahali, H. Sénéchal, P. Poncet, Cypress Pollinosis: from Tree to Clinic, *Clin. Rev. Allergy Immunol.* 20562. 56 (2017) 174–195. <https://doi.org/10.1007/S12016-017-8602-Y>.
3. G. D’Amato, L. Cecchi, S. Bonini, C. Nunes, I. Annesi-Maesano, H. Behrendt, G. Liccardi, T. Popov, P. Van Cauwenberge, Allergenic pollen and pollen allergy in Europe, *Allergy.* 62 (2007) 976–990. <https://doi.org/10.1111/J.1398-9995.2007.01393.X>.
4. K. Ohashi-Doi, D. Utsumi, Y. Mitobe, K. Fujinami, Japanese Cedar Pollen Allergens in Japan, *Curr. Protein Pept. Sci.* 23 (2022) 837–850. <https://doi.org/10.2174/1389203723666220930155719>.
5. Y. Saito, M. Otsuka, Japanese cedar pollinosis: Discovery, nomenclature, and epidemiological trends, *Proc. Japan Acad. Ser. B.* 90 (2014) 203–210. <https://doi.org/10.2183/PJAB.90.203>.
6. T. Osada, M. Okano, Japanese cedar and cypress pollinosis updated: new allergens, cross-reactivity, and treatment, *Allergol. Int.* 70 (2021) 281–290. <https://doi.org/10.1016/J.ALIT.2021.04.002>.
7. A. Hofmann, A.W. Burks, Pollen food syndrome: update on the allergens, *Curr. Allergy Asthma Rep.* 8 (2008) 413–417. <https://doi.org/10.1007/S11882-008-0080-0>.

8. T. Werfel, R. Asero, B.K. Ballmer-Weber, K. Beyer, E. Enrique, A.C. Knulst, A. Mari, A. Muraro, M. Ollert, L.K. Poulsen, S. Vieths, M. Worm, K. Hoffmann-Sommergruber, Position paper of the EAACI: food allergy due to immunological cross-reactions with common inhalant allergens, *Allergy*. 70 (2015) 1079–1090. <https://doi.org/10.1111/ALL.12666>.
9. C. Caffarelli, M. Garrubba, C. Greco, C. Mastroilli, C.P. Dascola, Asthma and food allergy in children: Is there a connection or interaction? *Front. Pediatr.* 4 (2016) 190754. <https://doi.org/10.3389/FPED.2016.00034/BIBTEX>.
10. C. Mastroilli, F. Cardinale, A. Giannetti, C. Caffarelli, Pollen-Food Allergy Syndrome: A not so Rare Disease in Childhood, *Med.* 55 (2019) 641. <https://doi.org/10.3390/MEDICINA55100641>.
11. O. Elisyutina, E. Fedenko, R. Campana, A. Litovkina, N. Ilina, D. Kudlay, E. Egorenkov, V. Smirnov, R. Valenta, C. Lupinek, and M. Khaitov. Bet v 1-specific IgE levels and PR-10 reactivity discriminate silent sensitization from phenotypes of birch allergy. *Allergy*, 74 (2019): 2525-2528. <https://doi.org/10.1111/all.13931>.
12. N. Inomata, Gibberellin-regulated protein allergy: Clinical features and cross-reactivity, *Allergol. Int.* 69 (2020) 11–18. <https://doi.org/10.1016/J.ALIT.2019.10.007>.
13. I. Muhammad, W.Q. Li, X.Q. Jing, M.R. Zhou, A. Shalmani, M. Ali, X.Y. Wei, R. Sharif, W.T. Liu, K.M. Chen, A systematic in silico prediction of gibberellic acid stimulated GASA family members: A novel small peptide contributes to floral architecture and transcriptomic changes induced by external stimuli in rice, *J. Plant Physiol.* 234–235 (2019) 117–132. <https://doi.org/10.1016/J.JPLPH.2019.02.005>.
14. H. Sénéchal, J. Šantrůček, M. Melčová, P. Svoboda, J. Zídková, D. Charpin, L. Guilloux, Y. Shahali, M.A. Selva, R. Couderc, T. Aizawa, P. Poncet, A new

- allergen family involved in pollen food-associated syndrome: Snakin/gibberellin-regulated proteins, *J. Allergy Clin. Immunol.* 141 (2018) 411-414.e4. <https://doi.org/10.1016/J.JACI.2017.06.041>.
15. H. Sénéchal, S. Keykhosravi, R. Couderc, M.A. Selva, Y. Shahali, T. Aizawa, J.M. Busnel, R. Arif, I. Mercier, N. Pham-Thi, D.A. Charpin, P. Poncet, Pollen/Fruit Syndrome: Clinical Relevance of the Cypress Pollen Allergenic Gibberellin-Regulated Protein, *Allergy. Asthma Immunol. Res.* 11 (2019) 143–151. <https://doi.org/10.4168/AAIR.2019.11.1.143>.
 16. T. Iizuka, M. Takei, Y. Saito, F. Rumi, J. Zheng, X. Lu, P. Chafey, C. Broussard, L. Guilloux-Assalet, D. Charpin, M. Ebisawa, H. Sénéchal, T. Aizawa, P. Poncet, Gibberellin-regulated protein sensitization in Japanese cedar (*Cryptomeria japonica*) pollen allergic Japanese cohorts, *Allergy.* 76 (2021) 2297–2302. <https://doi.org/10.1111/ALL.14816>.
 17. T. Iizuka, A. Barre, P. Rougé, D. Charpin, E. Scala, B. Baudin, T. Aizawa, H. Sénéchal, P. Poncet, Gibberellin-regulated proteins: Emergent allergens, *Front. Allergy.* 3 (2022) 105. <https://doi.org/10.3389/FALGY.2022.877553>.
 18. L. Tuppo, C. Alessandri, D. Pomponi, D. Picone, M. Tamburrini, R. Ferrara, M. Petriccione, I. Mangone, P. Palazzo, M. Liso, I. Giangrieco, R. Crescenzo, M.L. Bernardi, D. Zennaro, M. Helmer-Citterich, A. Mari, M.A. Ciardiello, Peamaclein – A new peach allergenic protein: similarities, differences and misleading features compared to Pru p 3, *Clin. Exp. Allergy.* 43 (2013) 128–140. <https://doi.org/10.1111/CEA.12028>.
 19. N. Inomata, M. Miyakawa, M. Aihara, Gibberellin-regulated protein in Japanese apricot is an allergen cross-reactive to Pru p 7, *Immunity, Inflamm. Dis.* 5 (2017) 469. <https://doi.org/10.1002/IID3.180>.

20. F. Huan, T.J. Han, M. Liu, M.S. Li, Y. Yang, Q.M. Liu, D. Lai, M.J. Cao, G.M. Liu, Identification and characterization of: *Crassostrea angulata* arginine kinase, a novel allergen that causes cross-reactivity among shellfish, *Food Funct.* 12 (2021) 9866–9879. <https://doi.org/10.1039/d1fo02042k>.
21. L. Tuppo, C. Alessandri, M.S. Pasquariello, M. Petriccione, I. Giangrieco, M. Tamburrini, A. Mari, M.A. Ciardiello, Pomegranate Cultivars: Identification of the New IgE-Binding Protein Peamaclein and Analysis of Antioxidant Variability, *J. Agric. Food Chem.* 65 (2017) 2702–2710. <https://doi.org/10.1021/ACS.JAFC.7B00092>.
22. M. Takei, C. Nin, T. Iizuka, M. Pawlikowski, M.A. Selva, Y. Chantran, Y. Nakajima, J. Zheng, T. Aizawa, M. Ebisawa, H. Sénéchal, P. Poncet, Capsicum Allergy: Involvement of Cap a 7, a New Clinically Relevant Gibberellin-Regulated Protein Cross-Reactive with Cry j 7, the Gibberellin-Regulated Protein from Japanese Cedar Pollen, *Allergy. Asthma Immunol. Res.* 14 (2022) 328. <https://doi.org/10.4168/AAIR.2022.14.3.328>.
23. L. Tuppo, C. Alessandri, I. Giangrieco, M. Ciancamerla, C. Rafaiani, M. Tamburrini, M.A. Ciardiello, A. Mari, Isolation of cypress gibberellin-regulated protein: Analysis of its structural features and IgE binding competition with homologous allergens, *Mol. Immunol.* 114 (2019) 189–195. <https://doi.org/10.1016/J.MOLIMM.2019.07.023>.
24. P. Poncet, H. Sénéchal, D. Charpin, Update on pollen-food allergy syndrome, *Expert Rev. Clin. Immunol.* 16 (2020) 561–578. <https://doi.org/10.1080/1744666X.2020.1774366>.
25. L. Cecchi, P. Poncet, L. Maltagliati, G. Carli, D. Macchia, L. Maggi, E. Meucci, P. Parronchi, A. Mazzoni, L. Salvati, E. Scala, H. Sénéchal, T. Aizawa, D. Villalta,

- F. Annunziato, L. Cosmi, A. Farsi, Optimization of the diagnosis and characterization of gibberellin-regulated protein sensitization: An Italian cohort study, *Ann. Allergy, Asthma Immunol.* (2023). <https://doi.org/10.1016/J.ANAI.2023.09.011>.
26. T. Koga, K. Tokuyama, S. Ogawa, E. Morita, Y. Ueda, T. Itazawa, A. Kamijo, Surveillance of pollen-food allergy syndrome in elementary and junior high school children in Saitama, Japan, *Asia Pac. Allergy.* 12 (2022). <https://doi.org/10.5415/APALLERGY.2022.12.E3>.
 27. Juturu V, Wu JC. Heterologous Protein Expression in *Pichia pastoris*: Latest Research Progress and Applications. *ChemBioChem.* 2018; 19(1):7-21.
 28. Wang J, Wang X, Shi L, et al. Methanol-Independent Protein Expression by AOX1 Promoter with trans-Acting Elements Engineering and Glucose-Glycerol-Shift Induction in *Pichia pastoris*. *Sci Rep.* 2017; 7:41850.
 29. K.M.S. Cabral, M.S. Almeida, A.P. Valente, F.C.L. Almeida, E. Kurtenbach, Production of the active antifungal *Pisum sativum* defensin 1 (Psd1) in *Pichia pastoris*: overcoming the inefficiency of the STE13 protease, *Protein Expr. Purif.* 31 (2003) 115–122. [https://doi.org/10.1016/S1046-5928\(03\)00136-0](https://doi.org/10.1016/S1046-5928(03)00136-0).
 30. P. Kant, W.Z. Liu, K.P. Pauls, PDC1, a corn defensin peptide expressed in *Escherichia coli* and *Pichia pastoris* inhibits growth of *Fusarium graminearum*, *Peptides.* 30 (2009) 1593–1599. <https://doi.org/10.1016/J.PEPTIDES.2009.05.024>.
 31. J.L. Cereghino, J.M. Cregg, Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*, *FEMS Microbiol. Rev.* 24 (2000) 45–66. <https://doi.org/10.1111/J.1574-6976.2000.TB00532.X>.

32. N. Koganesawa, T. Aizawa, H. Shimojo, K. Miura, A. Ohnishi, M. Demura, Y. Hayakawa, K. Nitta, K. Kawano, Expression and purification of a small cytokine growth-blocking peptide from armyworm *Pseudaletia separata* by an optimized fermentation method using the methylotrophic yeast *Pichia pastoris*, *Protein Expr. Purif.* 25 (2002) 416–425. [https://doi.org/10.1016/S1046-5928\(02\)00036-0](https://doi.org/10.1016/S1046-5928(02)00036-0).
33. M.R. Kuddus, F. Rumi, M. Tsutsumi, R. Takahashi, M. Yamano, M. Kamiya, T. Kikukawa, M. Demura, T. Aizawa, Expression, purification and characterization of the recombinant cysteine-rich antimicrobial peptide snak-in-1 in *Pichia pastoris*, *Protein Expr. Purif.* 122 (2016) 15–22. <https://doi.org/10.1016/j.pep.2016.02.002>.
34. F. De Lamotte, H. Boze, C. Blanchard, C. Klein, G. Moulin, M.F. Gautier, M.A. Delsuc, NMR Monitoring of Accumulation and Folding of ¹⁵N-Labeled Protein Overexpressed in *Pichia pastoris*, *Protein Expr. Purif.* 22 (2001) 318–324. <https://doi.org/10.1006/PREP.2001.1435>.
35. M. Zhang, Recent developments of methyl-labeling strategies in *Pichia pastoris* for NMR spectroscopy, *Protein Expr. Purif.* 166 (2020) 105521. <https://doi.org/10.1016/J.PEP.2019.105521>.
36. D.S. Wishart, C.G. Bigam, J. Yao, F. Abildgaard, H.J. Dyson, E. Oldfield, J.L. Markley, B.D. Sykes, ¹H, ¹³C and ¹⁵N chemical shift referencing in biomolecular NMR, *J. Biomol. NMR.* 6 (1995) 135–140. <https://doi.org/10.1007/BF00211777>.
37. F. Delaglio, S. Grzesiek, G.W. Vuister, G. Zhu, J. Pfeifer, A. Bax, NMRPipe: A multidimensional spectral processing system based on UNIX pipes, *J. Biomol. NMR.* 6 (1995) 277–293. <https://doi.org/10.1007/BF00197809>.
38. P. Güntert, Automated NMR structure calculation with CYANA., *Methods Mol. Biol.* 278 (2004) 353–378. <https://doi.org/10.1385/1-59259-809-9:353/COVER>.

39. H. Kumeta, K. Ogura, Y. Nishimiya, A. Miura, F. Inagaki, S. Tsuda, NMR structure note: A defective isoform and its activity-improved variant of a type III antifreeze protein from *Zoarces elongatus* Kner, *J. Biomol. NMR.* 55 (2013) 225–230. <https://doi.org/10.1007/S10858-012-9703-9/FIGURES/3>.
40. T. Jacob, L. Vogel, A. Reuter, A. Wangorsch, C. Kring, V. Mahler, B.M. Wöhrle, Food Processing Does Not Abolish the Allergenicity of the Carrot Allergen Dau c 1: Influence of pH, Temperature, and the Food Matrix, *Mol. Nutr. Food Res.* 64 (2020) 2000334. <https://doi.org/10.1002/MNFR.202000334>.
41. N.J. Greenfield, Using circular dichroism spectra to estimate protein secondary structure, *Nat. Protoc.* 16. 1 (2007) 2876–2890. <https://doi.org/10.1038/nprot.2006.202>.
42. A.L. Pey, The interplay between protein stability and dynamics in conformational diseases: The case of hPGK1 deficiency, *Biochim. Biophys. Acta - Proteins Proteomics.* 1834 (2013) 2502–2511. <https://doi.org/10.1016/j.bbapap.2013.07.011>.
43. P.J. Robinson, M.A. Pringle, C.A. Woolhead, N.J. Bulleid, Folding of a single domain protein entering the endoplasmic reticulum precedes disulfide formation, *J. Biol. Chem.* 292 (2017) 6978–6986. <https://doi.org/10.1074/jbc.M117.780742>.
44. W.T. Soh, L. Aglas, G.A. Mueller, S. Gilles, R. Weiss, S. Scheiblhofer, S. Huber, T. Scheidt, P.M. Thompson, P. Briza, R.E. London, C. Traidl-Hoffmann, C. Cabrele, H. Brandstetter, F. Ferreira, Multiple roles of Bet v 1 ligands in allergen stabilization and modulation of endosomal protease activity, *Allergy Eur. J. Allergy Clin. Immunol.* 74 (2019) 2382–2393. <https://doi.org/10.1111/all.13948>.
45. M. Egger, A. Jürets, M. Wallner, P. Briza, S. Ruzek, S. Hainzl, U. Pichler, C. Kitzmüller, B. Bohle, C.G. Huber, F. Ferreira, Assessing protein immunogenicity

- with a dendritic cell line-derived endolysosomal degradome, *PLoS One*. 6 (2011).
<https://doi.org/10.1371/journal.pone.0017278>.
46. E.I. Finkina, D.N. Melnikova, I. V. Bogdanov, N.S. Matveevskaya, A.A. Ignatova, I.Y. Toropygin, T. V. Ovchinnikova, Impact of Different Lipid Ligands on the Stability and IgE-Binding Capacity of the Lentil Allergen Len c 3, *Biomol.* 10 (2020) 1668. <https://doi.org/10.3390/BIOM10121668>.
 47. M.L. Bernardi, I. Giangrieco, L. Camardella, R. Ferrara, P. Palazzo, M.R. Panico, R. Crescenzo, V. Carratore, D. Zennaro, M. Liso, M. Santoro, S. Zuzzi, M. Tamburrini, M.A. Ciardiello, A. Mari, Allergenic Lipid Transfer Proteins from Plant-Derived Foods Do Not Immunologically and Clinically Behave Homogeneously: The Kiwifruit LTP as a Model, *PLoS One*. 6 (2011) e27856. <https://doi.org/10.1371/JOURNAL.PONE.0027856>.
 48. L. Tuppo, R. Spadaccini, C. Alessandri, H. Wienk, R. Boelens, I. Giangrieco, M. Tamburrini, A. Mari, D. Picone, M.A. Ciardiello, Structure, stability, and IgE binding of the peach allergen peamaclein (Pru p 7), *Biopolym. - Pept. Sci. Sect.* 102 (2014) 416–425. <https://doi.org/10.1002/bip.22530>.
 49. J.K. Lee, S. Ko, K. Ha, Y. Choi, Assessing the Predictive Ability of Computational Epitope Prediction Methods on Fel d 1 Allergen, *BioRxiv*. (2023) 543222. <https://doi.org/10.1101/2023.06.01.543222>.
 50. H. Singh, G.P.S. Raghava, ProPred: Prediction of HLA-DR binding sites, *Bioinformatics*. 17 (2002) 1236–1237. <https://doi.org/10.1093/bioinformatics/17.12.1236>.
 51. J.N. Clifford, M.H. Høie, S. Deleuran, B. Peters, M. Nielsen, P. Marcatili, BepiPred-3.0: Improved B-cell epitope prediction using protein language models, *Protein Sci.* 31 (2022) e4497. <https://doi.org/10.1002/PRO.4497>.

52. J.V. Kringelum, C. Lundegaard, O. Lund, M. Nielsen, Reliable B Cell Epitope Predictions: Impacts of Method Development and Improved Benchmarking, *PLOS Comput. Biol.* 8 (2012) e1002829. <https://doi.org/10.1371/JOURNAL.PCBI.1002829>.
53. Shen Y, Lange O, Delaglio F, Rossi P, Aramini JM, Liu G, Eletsky A, Wu Y, Singarapu KK, Lemak A, Ignatchenko A, Arrowsmith CH, Szyperski T, Montelione GT, Baker D, Bax A. Consistent blind protein structure generation from NMR chemical shift data. *Proc Natl Acad Sci U S A.* 2008 Mar 25;105(12):4685-90. <https://doi.org/10.1073/pnas.0800256105>.
54. A.C.Y. Foo, J.B. Nesbit, S.A.Y. Gipson, H. Cheng, P. Bushel, E.F. DeRose, C.H. Schein, S.S. Teuber, B.K. Hurlburt, S.J. Maleki, G.A. Mueller, Structure, Immunogenicity, and IgE Cross-Reactivity among Walnut and Peanut Vicilin-Buried Peptides, *J. Agric. Food Chem.* 70 (2022) 2389–2400. <https://doi.org/10.1021/acs.jafc.1c07225>.
55. A. Chatterjee, A. Kumar, J. Chugh, S. Srivastava, N.S. Bhavesh, R. V. Hosur, NMR of unfolded proteins, *J. Chem. Sci.* 117 (2005) 3–21. <https://doi.org/10.1007/BF02704356>.
56. B. Hoffmann, C. Eichmüller, O. Steinhauser, R. Konrat, Rapid Assessment of Protein Structural Stability and Fold Validation via NMR, *Methods Enzymol.* 394 (2005) 142–175. [https://doi.org/10.1016/S0076-6879\(05\)94006-8](https://doi.org/10.1016/S0076-6879(05)94006-8).
57. T.A. Young, E. Skordalakes, S. Marqusee, Comparison of Proteolytic Susceptibility in Phosphoglycerate Kinases from Yeast and E. coli: Modulation of Conformational Ensembles Without Altering Structure or Stability, *J. Mol. Biol.* 368 (2007) 1438–1447. <https://doi.org/10.1016/J.JMB.2007.02.077>.

58. N. Lützner, H. Kalbacher, Quantifying cathepsin S activity in antigen presenting cells using a novel specific substrate, *J. Biol. Chem.* 283 (2008) 36185–36194. <https://doi.org/10.1074/jbc.M806500200>.
59. Z. Zhang, R. Ma, Y. Xu, L. Chi, Y. Li, G. Mu, X. Zhu, Investigation of the Structure and Allergic Potential of Whey Protein by Both Heating Sterilization and Simulation with Molecular Dynamics, *Foods* 11 (2022) p4050. <https://doi.org/10.3390/FOODS11244050>.
60. E.S. Foster, I. Kimber, R.J. Dearman, Relationship between protein digestibility and allergenicity: Comparisons of pepsin and cathepsin, *Toxicology*. 309 (2013) 30–38. <https://doi.org/10.1016/J.TOX.2013.04.011>.
61. J.D. Astwood, J.N. Leach, R.L. Fuchs, Stability of food allergens to digestion in vitro, *Nat. Biotechnol.* 14 (1996) 1269–1273. <https://doi.org/10.1038/nbt1096-1269>.
62. A. Langenkamp, M. Messi, A. Lanzavecchia, F. Sallusto, Kinetics of dendritic cell activation: impact on priming of TH1, TH2 and nonpolarized T cells, *Nat. Immunol.* 1 (2000) 311–316. <https://doi.org/10.1038/79758>.
63. Traidl-Hoffmann, A. Kasche, A. Menzel, T. Jakob, M. Thiel, J. Ring, H. Behrendt, Impact of Pollen on Human Health: More Than Allergen Carriers? *Int. Arch. Allergy Immunol.* 131 (2003) 1–13. <https://doi.org/10.1159/000070428>.
64. S. Scheiblhofer, J. Laimer, Y. Machado, R. Weiss, J. Thalhamer, Influence of protein fold stability on immunogenicity and its implications for vaccine design, *Expert Rev Vaccines*. 16 (2017) 479–489. <https://doi.org/10.1080/14760584.2017.1306441>.

65. T. So, H.O. Ito, T. Koga, S. Watanabe, T. Ueda, T. Imoto, Depression of T-cell Epitope Generation by Stabilizing Hen Lysozyme, *J. Biol. Chem.* 272 (1997) 32136–32140. <https://doi.org/10.1074/JBC.272.51.32136>.

7. Supplemental

7.1 Native Pru p 7 sample preparation

The native Pru p 7 from peach pulp was extracted and purified according to the procedures published previously with some modifications. Frozen peach material was broken and dissolved in buffer and added equal volume of NaCl to final concentration 2 M. The mixture was filtered by different size filter membranes and collected the supernatants after centrifugation. The solution was diluted 40 times by Milli-Q and applied onto a Ni-NTA column (TOYOPEARL CM-650C) equilibrated with wash buffer (20 mM sodium phosphate buffer, pH 7.1). The target protein was eluted using elution buffer (20 mM sodium phosphate buffer, pH 6.3), and all fractions were collected after centrifugation at 5000 rpm, 1 minutes. Then the sample was filtrated, loaded onto a Cosmosil 5C18-AR300 column with a linear gradient 0–40% acetonitrile with 0.1% TFA for 80 minutes, and purified by RP-HPLC. After that, the freeze-dried native Pru p 7 sample were dissolved in a mixture of 10% D₂O solution with 50mM potassium phosphate buffer, pH 6.0 at 25°C to analyze the structure by ¹H-NMR.

7.2 Molecular dynamics simulation

The molecular dynamics simulation (MD) of Cry j 7 and Pru p 7 was conducted in a water environment, using the TIP3 water model [1]. The system was prepared using CHARMM-GUI (<https://www.charmm-gui.org/>) with the appropriate protonation state representing pH 3.0 by calculating the pK_a value of each residue using H⁺⁺ website. The simulation was done using the CHARMM27 force field and computational

package GROMACS 2023.2 version [2]. The three-dimensional NMR structures of Cry j 7 and Pru p 7 were used as the starting structure for the simulation. Sodium chloride ions with a concentration of 0.15 M were added to neutralize the system charge. All H-bonds were constrained using the LINCS algorithm. The electrostatic calculation was done using Particle-Mesh-Ewald (PME) grids, with a cut-off radius of 1.2 nm to reduce computing time. The system underwent an energy minimization using 50,000 steps of calculation [3]. After that, the system was equilibrated to 300 K for 100 ps in NVT ensemble, and the pressure to 1 bar for 100 ps using the NPT ensemble both using V-rescale thermostat and barostat respectively [4]. The equilibrated system was then subjected to 200 ns production simulation in NPT, 300K ensembles using V-rescale thermostat and Parinello-Rahman barostat and three replicas [5]. Trajectories and flexibility were evaluated through RMSD and RMSF of C using the trajectory from the last 50 ns.

7.3 References

1. P. Mark, L. Nilsson, Structure and dynamics of the TIP3P, SPC, and SPC/E water models at 298 K, *J. Phys. Chem. A.* 105 (2001) 9954–9960. <https://doi.org/10.1021/JP003020W>.
2. A.D. MacKerell, J. Nilesch Banavali, N. Foloppe, Development and Current Status of the CHARMM Force Field for Nucleic Acids, (2001). <https://doi.org/10.1002/1097-0282>.

3. W.F. Porto, O.L. Franco, Theoretical structural insights into the snakin/GASA family, *Peptides*. 44 (2013) 163–167.
<https://doi.org/10.1016/J.PEPTIDES.2013.03.014>.
4. Q. Ke, X. Gong, S. Liao, C. Duan, L. Li, Effects of thermostats/barostats on physical properties of liquids by molecular dynamics simulations, *J. Mol. Liq.* 365 (2022) 120116. <https://doi.org/10.1016/J.MOLLIQ.2022.120116>.
5. M. Kim, E. Kim, S. Lee, J.S. Kim, S. Lee, New method for constant- NPT molecular dynamics, *J. Phys. Chem. A*. 123 (2019) 1689–1699.
<https://doi.org/10.1021/ACS.JPCA.8B09082>.

Table S1 (A). NMR experimental details of Cry j 7

NMR experiment	pulse program	scan number	f1 data size	f2 data size
[¹ H- ¹⁵ N] HSQC	hsqcetf3gpphgw	2	128	
[¹ H- ¹³ C] HSQC*	hsqcedetgppsp	16	512	
HNCO	hncogpwg3d	16	128	40
HNCA	hncagp3d	256	128	1
¹ H- ¹ H NOESY**	noesyegpph	8	1024	
¹ H- ¹ H TOCSY**	mlevesgpph	8	1024	
¹⁵ N-NOESY	noesyhsqcf3gpwg3d	8	256	64
¹⁵ N-TOCSY	dipsihsqcf3gpsi3d	8	256	64

Table S1 (B). NMR experimental details of Pru p 7

NMR experiment	pulse program	scan number	f1 data size	f2 data size
[¹ H- ¹⁵ N] HSQC	hsqcetf3gpsi2	2	128	
[¹ H- ¹³ C] HSQC*	hsqcedetgppsp	16	512	
HNCO	hncogpwg3d	32	128	40
HNCA	hncagp3d	64	128	40
¹ H- ¹ H NOESY**	noesyegpph	8	1024	
¹ H- ¹ H TOCSY**	mlevesgpph	8	1024	
¹⁵ N-NOESY	noesyhsqcf3gpwg3d	8	256	64
¹⁵ N-TOCSY	dipsihsqcf3gpsi3d	8	256	64

*The spectral width of the ¹³C dimension was set at 80 ppm.

**With ¹⁵N-decoupled.

Table S2. CYANA structure calculation statistics of Cry j 7 and Pru p 7

CYANA NOE assay	Cry j 7	Pru p 7
Upper distance limits		
Total	766	712
Short-range ($ i-j \geq 1$)	489	458
Medium-range ($1 < i-j < 5$)	201	198
Long-range ($ i-j \geq 5$)	76	56
Dihedral angle limits		
Total	91	81
Phi	46	40
Psi	45	41
Average target function value	0.02	0.05
Violations		
Distance $> 0.1 \text{ \AA}$	0	0
Angle $> 1^\circ$	0	0
RMSD (residue range 2..62)		
Backbone atoms	0.63	0.8
Heavy atoms	0.98	1.46
Ramachandran plot		
most favored region	66.0%	76.9%
additionally allowed region	32.1%	17.3%
generously allowed region	1.95%	3.8%
disallowed region	0%	1.9%

A

GACTGGTTCCAATTGACAAGCTTTGTATTTTAACGACTTTTAACGACAACCTTGAGAAG
ATCAAAAAACAACATAATTATTCGAAGGATCCAAACGATGAGATTTCCTTCAATTTTAC
TGCAGTTTTATTTCGCAGCATCCTCCGCATTAGCTGCTCCAGTCAACACTACAACAGAAG
ATGAAACGGCACAATTCGGGCTGAAGCTGTCATCGGTTACTCAGATTTAGAAGGGGA
TTTCGATGTTGCTGTTTTGCCATTTTCCAACAGCACAAATAACGGGTTATTGTTTATAAA
TACTACTATTGCCAGCATTGCTGCTAAAGAAGAAGGGGTATCT**CTCGA**GAAAAGA**GCT**
CACATCGACTGCGACAAGGAGTGCAACAGAAGATGCTCCAAGGCTTCCGCTCACGAC
AGATGCTTGAAGTACTGCGGTATCTGCTGCGAGAAGTGCAACTGTGTCCCACCCGGTA
CTTACGGTAACGAGGACTCTTGTCATGCTACGCCAAGTTGAAGAACTCCAAGGGTGG
TCACAAGTGTCCATAA**GAATTC**CCTAGGGCGGCCGCGAATTAATTCGCCTTAGACATGA
CTGTTCTCAGTTCAAGTTGGGCACTTACGAGAAGACCGGTCTTGCTAGATTCTAATCA
AGAGGATGTCAGAATGCCATTTCG
5' AOX 1 primer site α -factor **Xho I** KEX2 site **Cry j 7** **EcoR I** **3' AOX 1 primer site**

B

GACTGGTTCCAATTGACAAGCTTTGTATTTTAACGACTTTTAACGACAACCTTGAGAAG
ATCAAAAAACAACATAATTATTCGAAGGATCCAAACGATGAGATTTCCTTCAATTTTAC
TGCAGTTTTATTTCGCAGCATCCTCCGCATTAGCTGCTCCAGTCAACACTACAACAGAAG
ATGAAACGGCACAATTCGGGCTGAAGCTGTCATCGGTTACTCAGATTTAGAAGGGGA
TTTCGATGTTGCTGTTTTGCCATTTTCCAACAGCACAAATAACGGGTTATTGTTTATAAA
TACTACTATTGCCAGCATTGCTGCTAAAGAAGAAGGGGTATCT**CTCGA**GAAAAGA**GGT**
TCTAGCTTTTTCGATTCCAAATGTGGAGTCAGATGCAGTAAAGCTGGATACCAAGAAA
GGTGCTTGAAGTACTGTGGTATTTGCTGTGAGAAGTGTCATTGTGTTCATCTGGTACT
TATGGGAACAAGGATGAATGTCCCTGTTATCGTGACCTTAAGAACTCAAAAGGCAATC
CTAAGTGTCCATAA**GCGGCC**GCGAATTAATTCGCCTTAGACATGACTGTTCTCAGTTC
AAGTTGGGCACTTACGAGAAGACCGGTCTTGCTAGATTCTAATCAAGAGGATGTCAGA
ATGCCATTTCG
5' AOX 1 primer site α -factor **Xho I** KEX2 site **Pru p 7** **Not I** **3' AOX 1 primer site**

C

Secretion signal from pPIC9

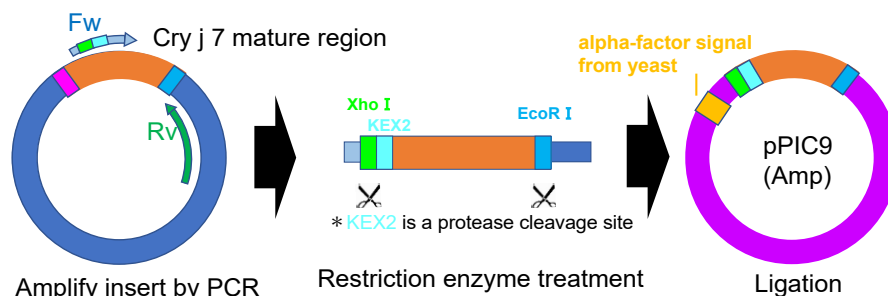


Figure S1. DNA sequences of recombinant plasmids. (A) pPIC9-Cry j 7. (B) pPIC9-Pru p 7. (C) The vector construction of recombinant Cry j 7.

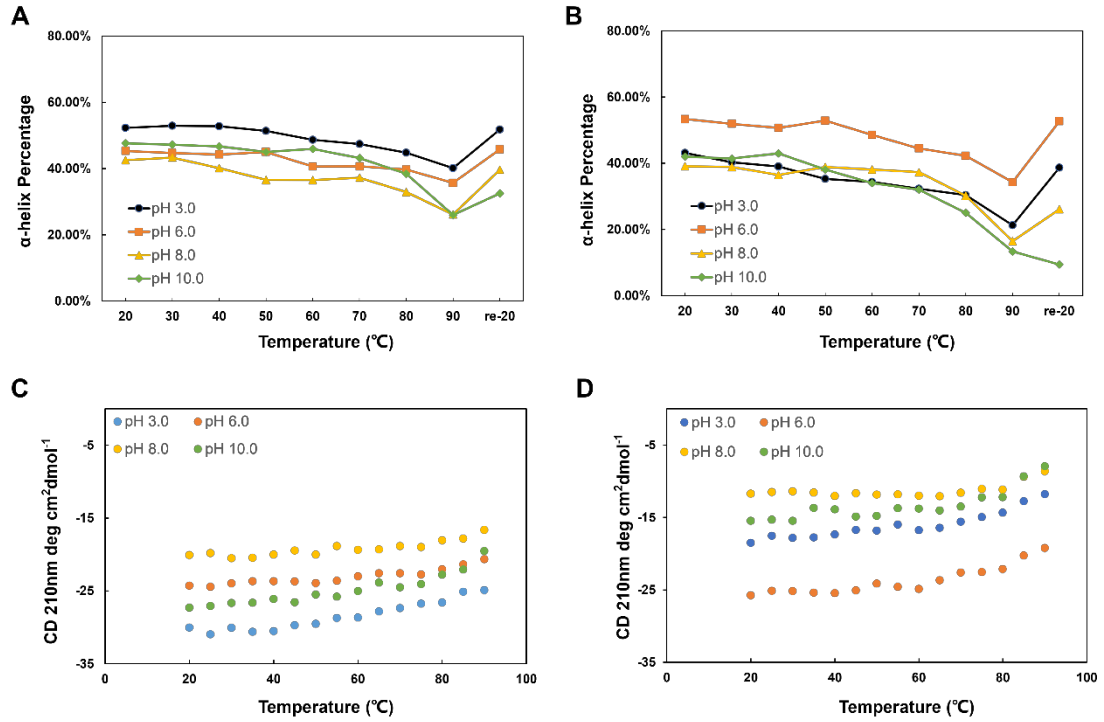


Figure S2. Circular dichroism spectra of Cry j 7 and Pru p 7 analysis at different pH values and temperature. The conformation analysis of Cry j 7 (A) and Pru p 7 (B) spectra through BeStSel website (<https://bestsel.elte.hu/index.php>). X-axis: the temperature ranges from 20 °C to 90 °C and back again to 20 °C. Y-axis: the α -helix structure proportion. The CD spectra values of Cry j 7 (C) and Pru p 7 (D) obtained by scanning at 210 nm for different temperature. Blue dotted: pH 3.0, orange dotted: pH 6.0, yellow dotted: pH 8.0, green dotted: pH 10.0.

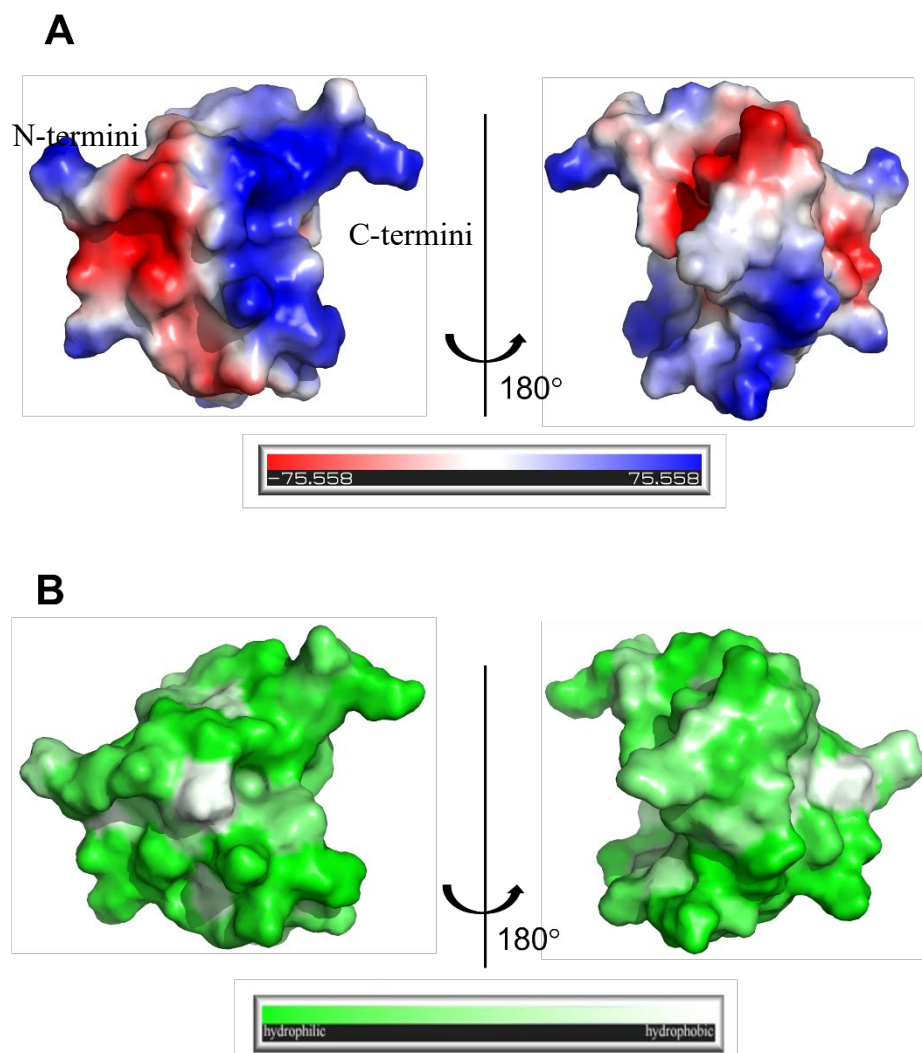


Figure S3. Protein structure characterization of Cry j 7. (A) Electrostatic potential at the surface of the Cry j 7. Hydrophilic and hydrophobic surfaces of the amphiphilic helix are represented. (B) Hydrophilic surface (green) and hydrophobic surface (white) of Cry j 7. This script colors the selection passed in based on the hydrophobicity scale as defined. *

*: Eisenberg D, Schwarz E, Komarony M & Wall R (1984) Amino acid scale: Normalized consensus hydrophobicity scale. J. Mol. Biol. 179:125-142.

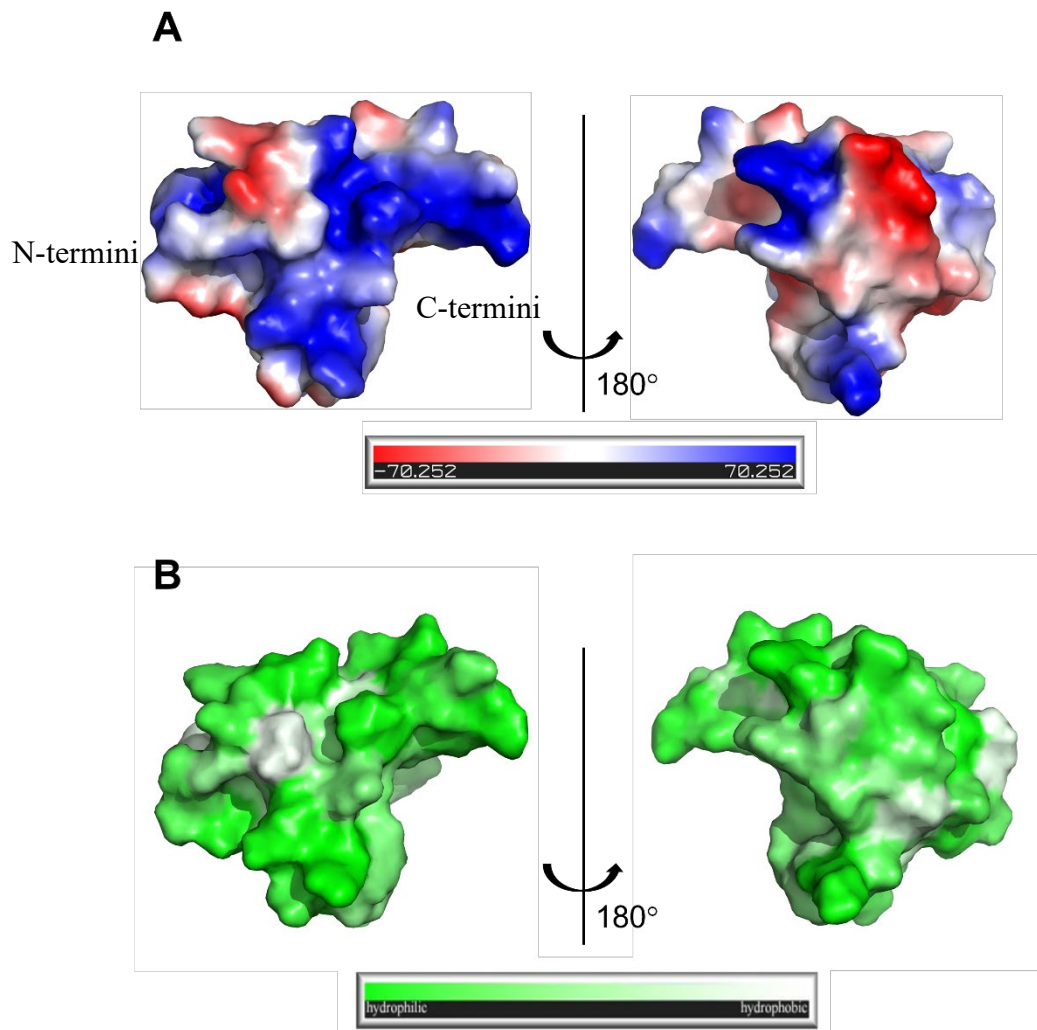


Figure S4. Protein structure characterization of Pru p 7. (A) Electrostatic potential at the surface of the Pru p 7. Hydrophilic and hydrophobic surfaces of the amphiphilic helix are represented. (B) Hydrophilic surface (green) and hydrophobic surface (white) of Pru p 7. This script colors the selection passed in based on the hydrophobicity scale as defined. *

*: Eisenberg D, Schwarz E, Komarony M & Wall R (1984) Amino acid scale: Normalized consensus hydrophobicity scale. J. Mol. Biol.179:125-142.

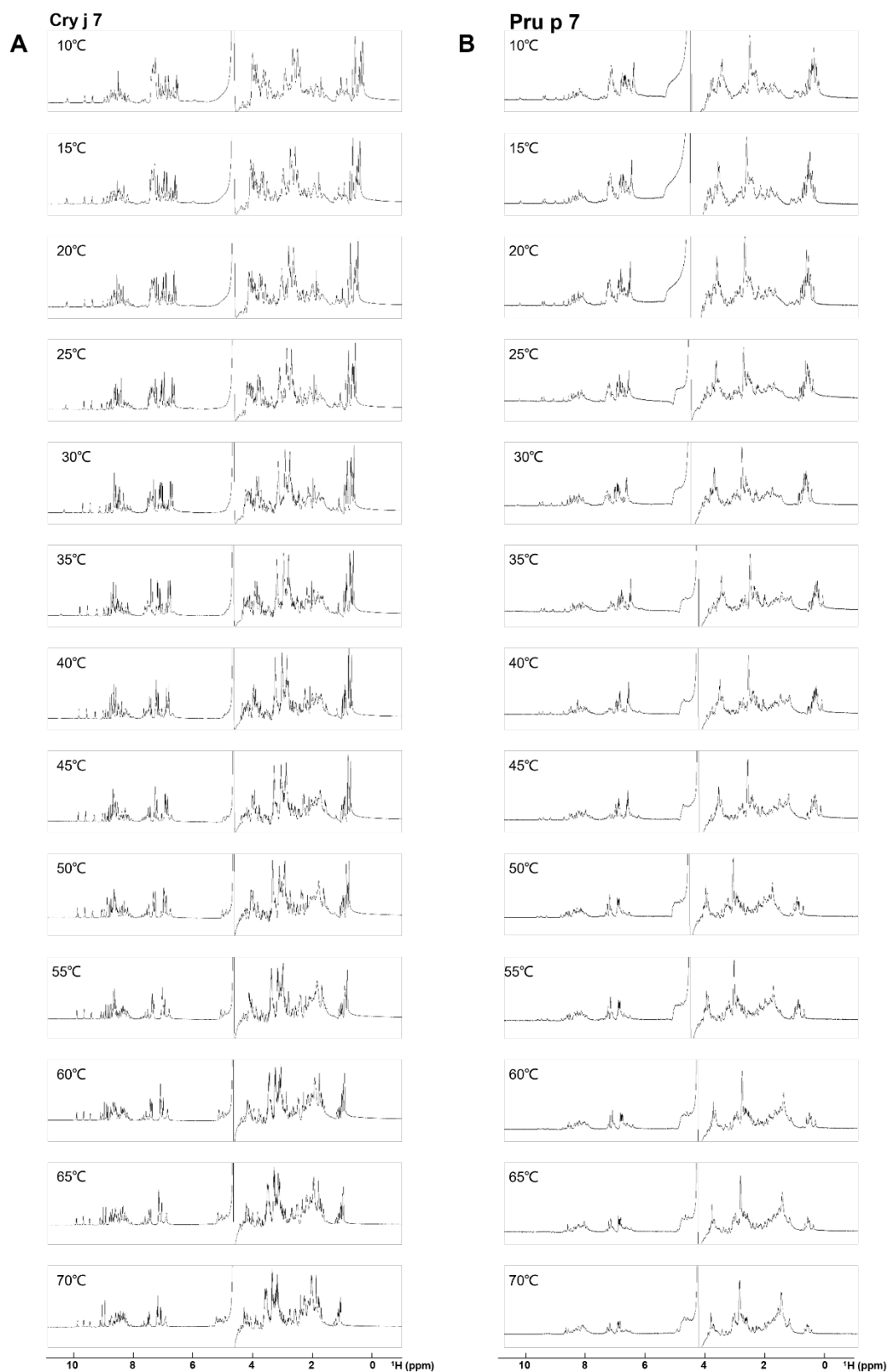


Figure S5. ^1H NMR spectra of Cry j 7 (A) and Pru p 7 (B) at temperatures ranging from 10 °C to 70 °C in 10% $\text{D}_2\text{O}/\text{H}_2\text{O}$ solution, pH 3.0.

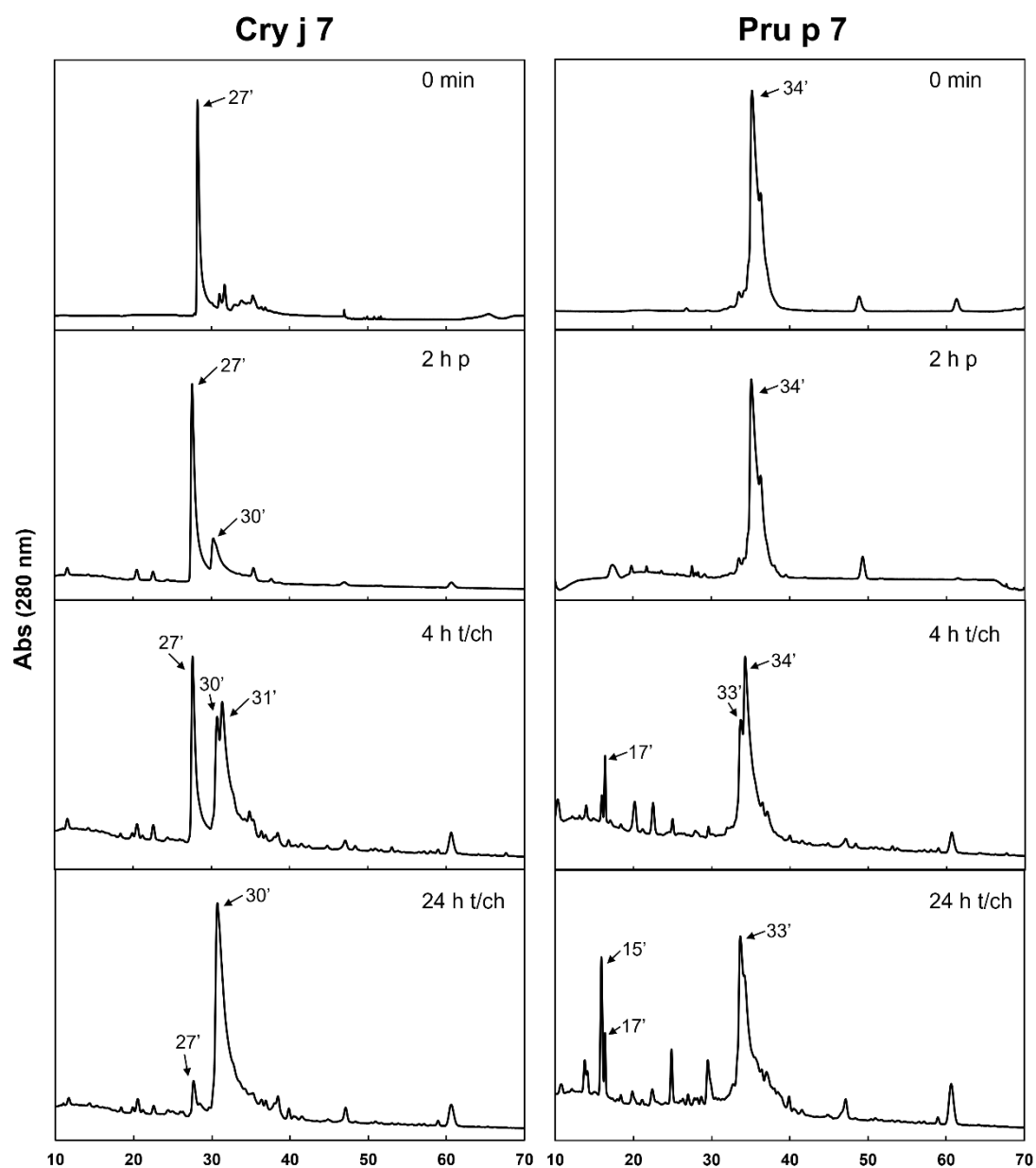


Figure S6. RP-HPLC analysis on stimulated gastrointestinal digestion of Cry j 7 and Pru p 7. RP-HPLC on C18 column with a linear gradient of acetonitrile concentration from 10% to 40% for 60 minutes (0min, 2 h p: pepsin digestion for 2 hours, 4 h t/ch and 24 h t/ch: subsequent trypsin/chymotrypsin digestion for 4 hours and 24 hours).

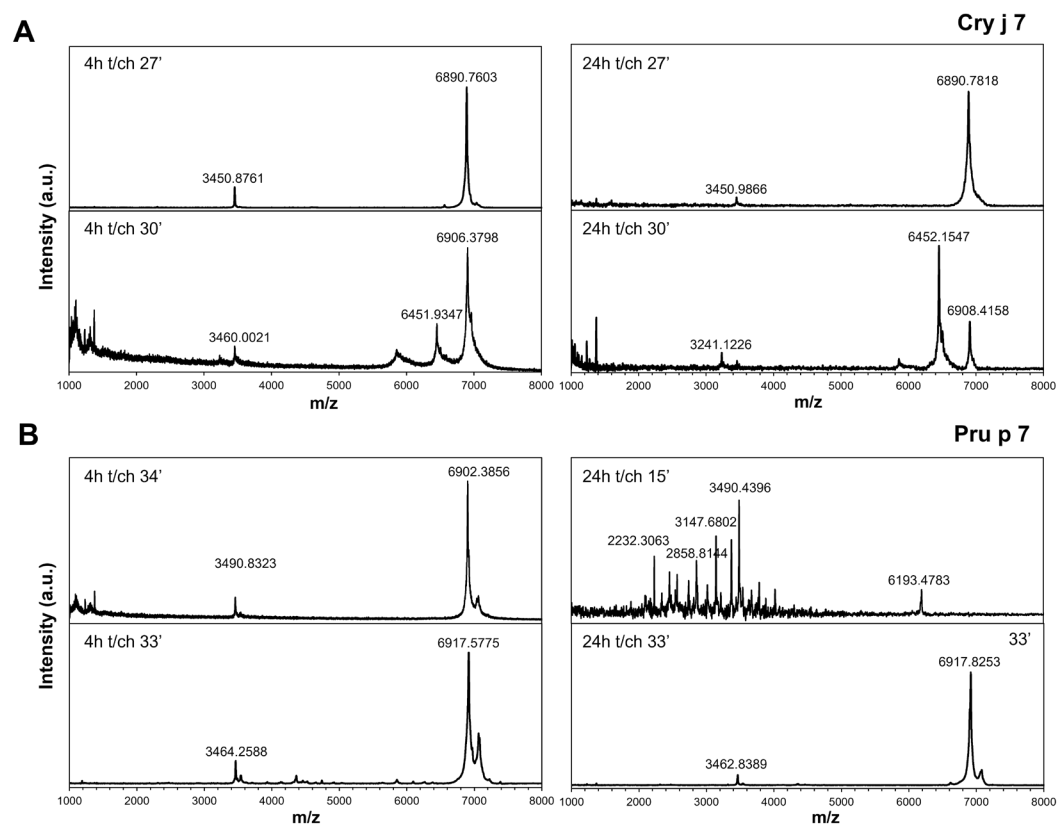


Figure S7. MALDI-TOF mass spectra of the RP-HPLC fractions of gastrointestinal digestion of Cry j 7 (A) and Pru p 7 (B) for 4 hours and 24 hours. The fractions of Cry j 7: 27 min and 30 min. Fractions of Pru p 7: 15 min / 34 min and 33 min.

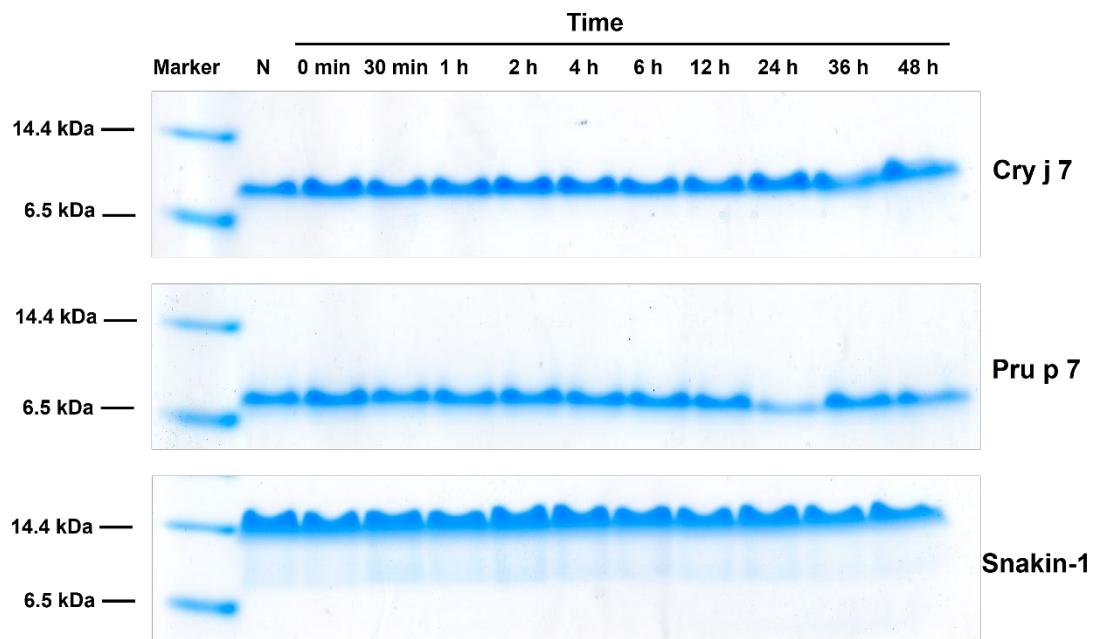


Figure S8. Degradation analysis of Cry j 7 and Pru p 7 *in vitro*. Endosomal degradation assay by Cathepsin S enzyme in sodium acetate buffer, pH 5.0 without DTT, and the degradation profile was analyzed by tricine SDS PAGE. Lane Marker: protein molecular weight marker. Lane N: sample without DTT and Cathepsin S enzyme.

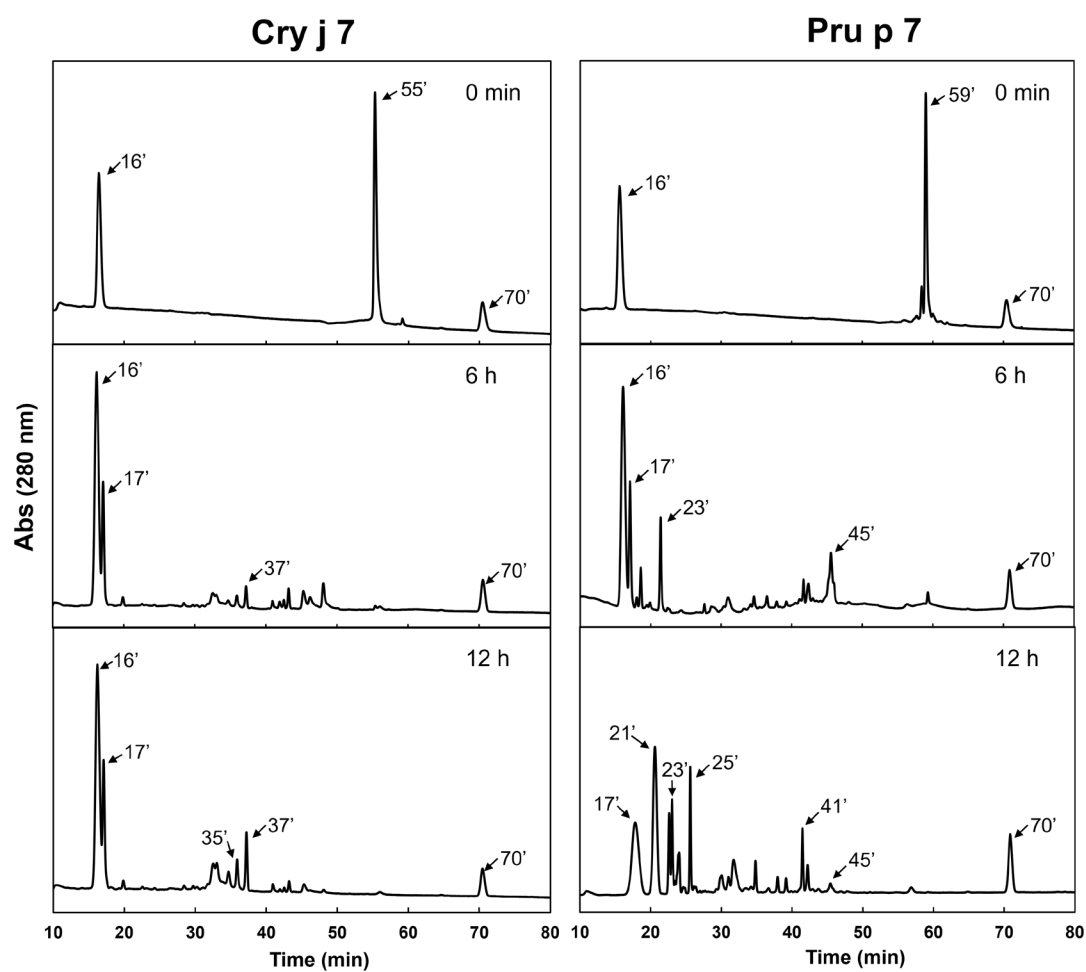


Figure S9. RP-HPLC analysis on Cathepsin S enzyme digestion of Cry j 7 and Pru p 7.

RP-HPLC on C18 column with a linear gradient of acetonitrile concentration from 5% to 35% for 70 minutes (0 min, 6 hours, and 12 hours).

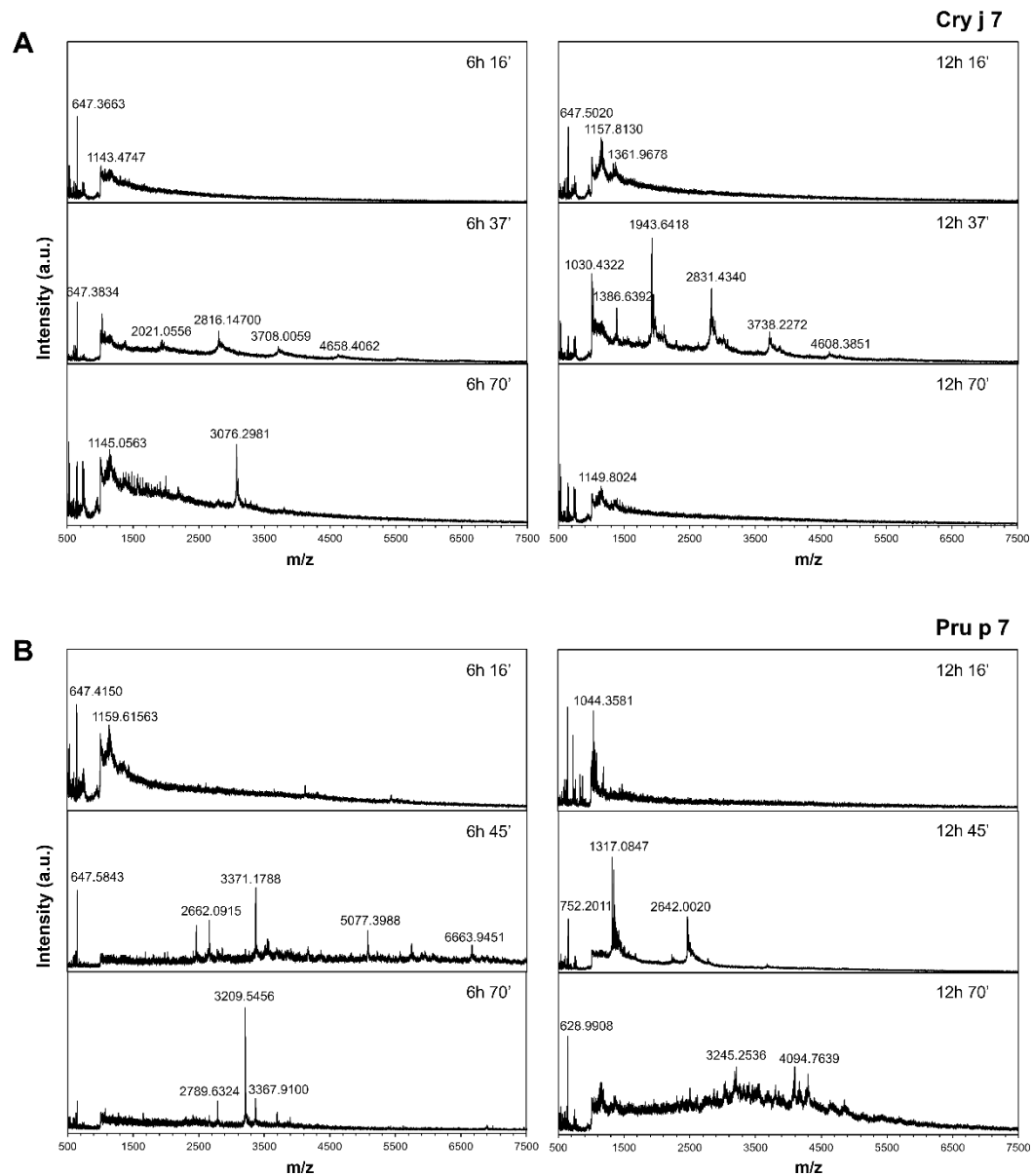


Figure S10. MALDI-TOF mass spectra of the RP-HPLC fractions of Cathepsin S enzyme digestion. (A) The HPLC fractions (16 min, 37 min, and 70 min) of Cry j 7 after Cathepsin S enzyme digestion for 6 hours or 12 hours. (B) The HPLC fractions (16 min, 45 min, and 70 min) of Pru p 7 after Cathepsin S enzyme digestion for 6 hours or 12 hours.

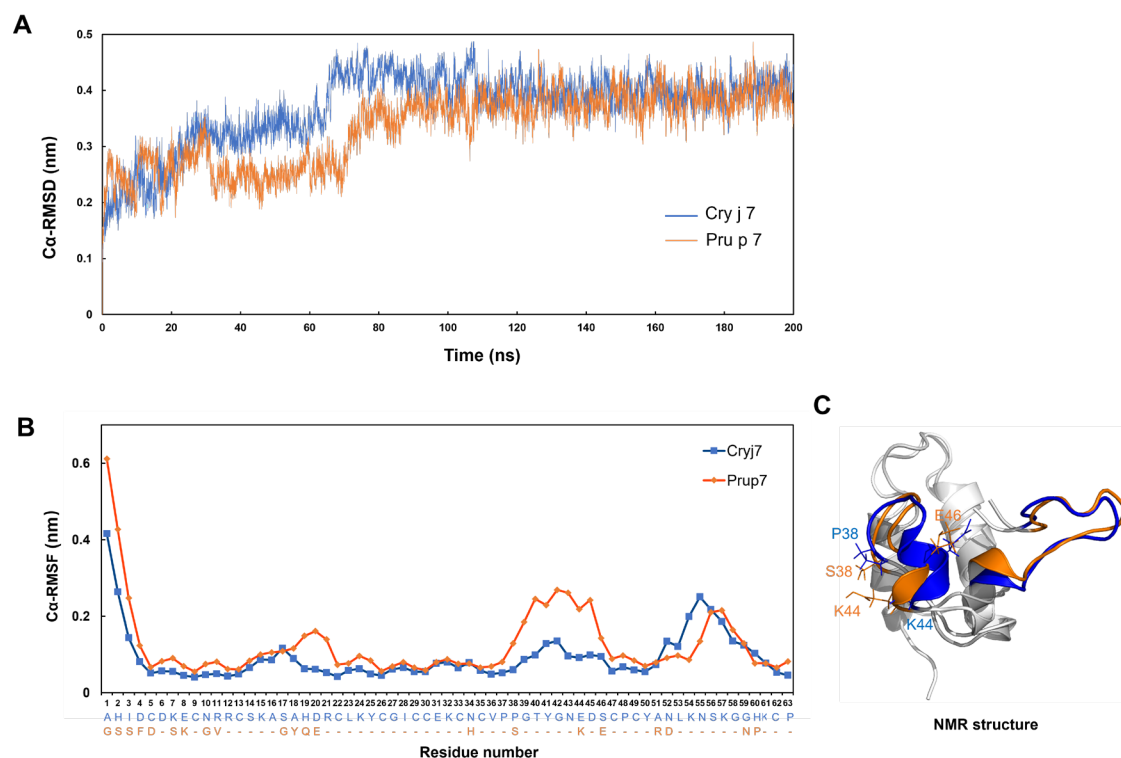


Figure S11. Molecular dynamics simulation results of Cry j 7 and Pru p 7 at 300 k. (A) RMSD plots for C-alpha after lsq fit to C-alpha of Cry j 7 and Pru p 7 throughout 200 ns simulation. (B) RMSF plots of Cry j 7 and Pru p 7 throughout the last 50 ns simulation. X-axis: amino acid sequence of Cry j 7 and Pru p 7; Y-axis: C α -RMSF data. (C) NMR structural alignment of Cry j 7 and Pru p 7. Blue color: P38-S46, and A51 -K61 residues of Cry j 7. Orange color: S38-E46, and E51-K61 residues of Pru p 7. Blue color: Cry j 7. Red color: Pru p 7.

Acknowledgments

I am truly grateful to all the professors on our course committee who have made this work possible, with special thanks to the professor Tomoyasu Aizawa for his invaluable suggestions and support. I would also like to express my gratitude to Dr. Hiroyuki Kumeta and Dr. Yasuhiro Kumaki for their helpful support and guidance in the use of NMR instrumentation and data analysis. I also deeply thank my fellow students for their advice, discussion, support and kindness in my experiments as well as in my daily life. Finally, I would like to express grateful appreciation to my family for the support and encouragement during my campus life in Hokkaido University.