

Almond (*Prunus dulcis*) Allergen Pru du 8, the First Member of a New Family of Food Allergens

Huilian Che,^{†,‡} Yuzhu Zhang,^{*,‡,§} Songsong Jiang,^{†,‡} Tengchuan Jin,^{§,||} Shu-Chen Lyu,^{||}
Kari C. Nadeau,^{||} and Tara McHugh[‡]

[†]Beijing Advanced Innovation Center for Food Nutrition and Human Health, College of Food Science and Nutritional Engineering, China Agricultural University, 17 Qinghua Donglu, Haidian District, Beijing 100038, People's Republic of China

[‡]Western Regional Research Center, Pacific West Area, Agricultural Research Service, United States Department of Agriculture, 800 Buchanan Street, Albany, California 94710, United States

[§]Hefei National Laboratory for Physical Sciences at Microscale, Chinese Academy of Sciences Key Laboratory of Innate Immunity and Chronic Disease, School of Life Sciences and Medical Center, University of Science & Technology of China, Hefei, Anhui 230007, People's Republic of China

^{||}Division of Pediatric Immunology, Allergy, and Rheumatology, Department of Pediatrics, Stanford University School of Medicine, 269 Campus Drive, Stanford, California 94305, United States

Supporting Information

ABSTRACT: An almond allergen with two known short peptide sequences was reported as the almond 2S albumin but was later suspected to be almond vicilin. However, this allergen was not designated by the World Health Organization/International Union of Immunological Societies. This study aimed to determine the true identity of this elusive almond allergen. cDNAs were synthesized from total RNA of the Nonpareil almond. The complete sequence of the previously reported almond allergen was determined from its coding sequence. The deduced protein was produced recombinantly and was confirmed to be a food allergen by testing with 18 almond-allergic sera. The allergen is a potential cysteine-rich antimicrobial protein with characteristic C[X]₃C–[X]_{10–12}–C[X]₃C motifs of the hairpinin antimicrobial protein. This first member of a novel family of food allergens was named Pru du 8. The signature motif of the hairpinin antimicrobial protein can be found in the N-terminal region of some vicilin allergens (e.g., Ara h 1). It can also be found in the signal peptide of other vicilin allergens (e.g., Car i 2). In many species, however, vicilins do not contain such a motif, indicating that the presence of the signature motifs of the hairpinin antimicrobial protein in vicilins might be a result of translocation during evolution.

KEYWORDS: food allergen, antimicrobial protein, α -hairpinin, serum IgE, Pru du 8

INTRODUCTION

Food allergies affect millions of patients in all parts of the world,^{1,2} and there is, presently, no cure for it.³ Tree nuts are regarded as healthy, and their consumption is recommended.⁴ However, tree nuts are also a major source of food allergens that can cause life-threatening allergic reactions.^{5,6} A large percentage of the protein weight in a plant seed is made up of a few proteins, such as the 2S, 7S, and 11S seed storage proteins, although there are thousands of different proteins in the seed. Unfortunately, many of these abundant proteins are known to be food allergens in more than one food, including peanuts, soybeans, and tree nuts. Cross-reactivity among different foods (as a result of a high percentage of sequence identity for linear epitopes or structural surface similarity for conformational epitopes) poses additional risks to allergic patients.^{7,8} Food allergy also has a negative impact on the safe use of agricultural products that can result in costly food recalls.⁹

Identification and characterization of food allergens is the first step for understanding the allergenicity of food proteins. Moreover, personalized component-resolved allergy diagnostics may help to better understand food allergies,¹⁰ and such diagnostics require the identification of the allergens and their availability. There are 16 officially designated allergens in

peanuts.¹¹ The prevalence of tree nut allergies in the United States is as high as that of peanut allergy.¹² Almonds are high on the list of the most likely tree nut to trigger allergic reactions.^{13,14} Prior to this study, only four almond allergens^{15–17} have been designated by the World Health Organization (WHO)/International Union of Immunological Societies (IUIS).¹¹ In a review, however, eight almond proteins were referred to as food allergens.¹⁸ One of these proteins was the 2S albumin based on a 2002 study in which almond protein extract was fractionated by anion-exchange chromatography and tested by an immunoblot with five patient sera.¹⁹ A protein band was recognized by at least one serum. After peptidase digestion of the protein, a 6 kDa fragment, which could still be recognized by the serum, and a 2 kDa band, which did not interact with the serum, were isolated. N-terminal sequencing of the protein and one component of the 2 kDa band was successful. On the basis of the alignments of sequences of the N-terminal sequencing results and those of the 2S proteins

Received: May 3, 2019

Revised: July 3, 2019

Accepted: July 9, 2019

Published: July 9, 2019

from walnuts and Brazil nuts, the allergenic protein was believed to be the 2S albumin.¹⁹ A recent study, however, suggested that the immunoglobulin E (IgE) reactive protein from almond might be the cleaved N-terminal region of a vicilin-like protein.²⁰ We isolated an almond vicilin and identified it as a new food allergen.²¹ Nevertheless, it is not the allergen involved in the aforementioned 2002 study because it does not contain the previously reported partial peptide sequences. Here, we report the isolation and characterization of a new almond allergen, which, if cleaved at the right positions, may result in peptide fragments that give N-terminal sequencing data consistent with the 2002 findings of Poltronieri et al.¹⁹

MATERIALS AND METHODS

Total RNA Extraction. Young Nonpareil almonds were collected from Travaille & Phippen, Inc. Total RNA was extracted using an RNeasy Plus Mini Kit (Qiagen, Valencia, CA, U.S.A.) following the protocol of the manufacturer, with minor modifications. Briefly, the inner part of almond, which would later become the kernel and the hard shell of the seed, was ground with a pestle in a mortar containing liquid nitrogen. About 100 mg of ground material was collected in a microcentrifuge tube immediately after liquid nitrogen was evaporated. A total of 500 μ L of buffer RLT was added, and the sample was mixed by inverting 6 times. RLT was one of the proprietary components of the kit whose composition was not provided by Qiagen. The sample was subjected to centrifugation, and the supernatant was loaded onto an RNeasy Mini Spin column after passing through a gDNA Eliminator Mini Spin column and mixed 2:1 with ethanol. The column was washed according to the protocol of the manufacturer, and the total RNA was eluted with 50 μ L RNase-free water. A total of 2 μ L (80 units) of RNase inhibitor (Invitrogen) was added immediately, and the sample was used for reverse transcription (RT) without delay.

RT and Polymerase Chain Reaction (PCR). Omniscript reverse transcriptase (Qiagen) was used to synthesize cDNA following the protocols of the manufacturer, using 6 μ M of a 20-mer poly-dT primer and 2 μ L of the total RNA as templates in a 20 μ L RT mixture. The RT reaction was incubated for 60 min at 37 °C and then for 5 min at 95 °C to inactivate the reverse transcriptase. A total of 2 μ L of cDNA was used as a template in a PCR using the Q5 high-fidelity DNA polymerase purchased from New England Biolabs (Ipswich, MA, U.S.A.). The forward primer (ZY387, 5'-CACACAGTAT-ATATTGAGCACA-3') used in the PCR was designed on the basis of a mRNA sequence annotated as *Prunus mume* vicilin C72-like mRNA (NCBI Reference Sequence XM_008226058.1). ZY387 and a 35-mer poly-T primer were custom-synthesized by Integrated DNA Technologies, Inc. (Skokie, IL, U.S.A.) and used as primers in the PCR. A PCR product of about 1 kb was obtained and inserted in an in-house cloning vector pBC digested with *EcoR* V. Bacteria dH5 α was transformed with the ligation product, and multiple positive transformants were isolated for plasmid DNA purification. The plasmids were sequenced commercially by Elim Biopharmaceuticals, Inc. (Hayward, CA, U.S.A.) using Elim's primers T3 (AATTA-ACCCTCACTAAAGGG) and M13for-40 (GTTTTCACGATCAGAC). An open reading frame of the inserted sequence was amplified with Q5 DNA polymerase using primers ZY389 (AATCG-CCAGCTGGGTGGTATGGCGACCATGACCAAA, forward) and ZY390 (ATAGACCTCGAGTCAGGCTACCATTCTAAT-TCCTTG, reverse). The PCR product was digested with *EcoR* V and *Xho* I and inserted in the pTMS1 vector²² digested with *Pvu* II and *Xho* I to generate pHMS-Prudu8 for expressing the new protein, which has been designated as Pru du 8 by the WHO/IUIS Allergen Nomenclature Sub-committee. It is also referred to as ssAMP below because it is an antimicrobial protein possibly rich in disulfide bonds. The inserted nucleotide sequence in pHMS-Prudu8 was confirmed by DNA sequencing.

Protein Expression. The pHMS-Prudu8 plasmid was transformed into the *Escherichia coli* strain Shuffle T7 (New England Biolabs). The bacteria started with a positive colony were grown in 1 L batches of lysogeny broth (LB) containing 50 mg/L kanamycin (Gold Biotechnology, St. Louis, MO, U.S.A.) at 37 °C. The temperature of the incubator was set to 18 °C when the optical density at a wavelength of 600 nm (OD₆₀₀) of the culture reached 1.0–1.2, and the protein expression was induced 45 min later by adding isopropyl β -D-1-thiogalactopyranoside (IPTG, Gold Biotechnology) to a final concentration of 1 mM. After 16 h of induction, the bacteria were collected by centrifugation at 2500g and the pellets were resuspended with a His-binding buffer (buffer HB, 20 mM imidazole at pH 8.0 and 500 mM NaCl) and stored at –80 °C in 50 mL centrifuge tubes.

Protein Purification and Tag Removal. Frozen cells were thawed by putting the tube containing the pellet in a beaker with tap water. The thawed pellet was resuspended in 50 mL of buffer HB containing 1 mM benzamidine. The sample was sonicated in a beaker surrounded by ice and water for 6 min and centrifuged at 20000g for 40 min at 18 °C. The recombinant protein was purified from the supernatant with a 5 mL HisTrap FF crude column from GE Healthcare (Piscataway, NJ, U.S.A.). The column was washed with 30 mL of buffer HW (56 mM imidazole at pH 8.0 and 500 mM NaCl), and the protein was eluted with buffer HE (300 mM imidazole at pH 8.0 and 200 mM NaCl). Ubiquitin-like protease (ULP, a sumo protease), which also has a His tag, was expressed and purified similarly in-house. Purified MS-Prudu8 was incubated with an optimized amount of ULP to release recombinant Pru du 8 (rPrudu8) from the expression tags, and the digested product was passed through a new HisTrap column as previously described.^{22,23}

Preparation of the Almond Protein Extract. Nonpareil almond kernels were ground in a KitchenAid blender for 5 min with 8 mL of an extraction buffer (10 mM Tris at pH 8.0) per gram of almond. The supernatant was collected after the sample was subjected to centrifugation for 15 min at 4000g. Much Pru du 6 was cold-precipitated by setting the sample tube on ice for 30 min. The sample was subjected to centrifugation at 20000g, and the supernatant was collected. The pellet was resuspended in the extraction buffer to the same volume of the supernatant. An aliquot of the resuspension was mixed with that of the supernatant at a ratio of 1:10, and the mixture was used as the almond protein extract.

Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis (SDS–PAGE) Analysis. The purified rPrudu8 and Nonpareil almond protein extract sample were incubated at 96 °C in a SDS sample buffer under reducing conditions for 10 min. SDS–PAGE with 4–20% polyacrylamide gels was carried out as described previously.²¹ Pre-stained protein markers (Bio-Rad, Hercules, CA, U.S.A.) with molecular weights of 10, 15, 20, 25, 37, 50, 75, 150, and 250 kDa were used as standards. Gels were stained with Coomassie Brilliant Blue (CBB) and documented with an ImageQuant LAS 400 imager (GE Healthcare).²³

N-Terminal Protein Sequencing. Purified rPrudu8 was subjected to SDS–PAGE and transferred to a polyvinylidene fluoride (PVDF) membrane (GE Healthcare) using a Trans-Blot SD semi-dry transfer cell from Bio-Rad (Hercules, CA, U.S.A.) and a Tris–glycine buffer without methanol following the protocols of the manufacturer. The membrane was stained with CBB using the conventional method. The rPrudu8 band was excised, and N-terminal amino acid sequencing was carried out by the analytical core facility at Tufts Medical School.

Patient Sera. A total of 18 sera were collected from patients and labeled as sera 1–18. The patients were enrolled in the food allergy study at Stanford University under institutional review board approval with informed consent (IRB approval certificate number 8629).

IgE Recognition of Almond Pru du 8. The protein bands in the gels were transferred to PVDF membranes for immunoblot. The membranes were blocked at room temperature in TBST (25 mM Tris at pH 7.4, 137 mM NaCl, 2.7 mM KCl, and 0.1% Tween 20) containing 10% non-fat dry milk. While the membranes were being blocked, individual serum (100 μ L) was diluted 1:10 with TBSTM

(TBST containing 1% non-fat milk) and incubated with 10 μ L of agarose-immobilized protein A purchased from Pierce (Rockford, IL, U.S.A.) for 45 min. The serum samples were then spun at 1000 rpm for 2 min, and supernatants were used as a primary antibody solution to incubate with a blocked membrane at room temperature for 45 min. The membranes were washed 3 times for 5 min each with TBST, processed with a secondary antibody (diluted 1:5000 with TBSTM), and washed again; the secondary antibody was raised against human IgE and conjugated with peroxidase (Sigma). A mixture of Pierce ECL western blotting reagents (Rockford, IL, U.S.A.) was used as a peroxidase substrate, and the ImageQuant LAS4000 imaging system was used for signal detection.

RESULTS

Determination of the Almond Antimicrobial Protein Sequence. cDNAs were synthesized from the RNA isolated from a young Nonpareil almond. Using the cDNA pool as templates, a PCR with primers ZY387 and a poly-T resulted in a 1 kb product. The nucleotide sequence of this product contained an open reading frame that encoded a protein of 264 amino acids. The sequence data has been deposited to the GenBank with accession number MH922028 (Figure S1 of the Supporting Information). Searching the non-redundant protein database at the National Center for Biotechnology Information (NCBI) with the new protein identified six homologous proteins (XP_007223717.1, XP_021811866.1, PQQ11123.1, XP_008224280.1, XP_009335337.1, and XP_008390980.1). These proteins were predicted by automated computational analysis or identified during genome annotation. They are named vicilin-like or vicilin-like antimicrobial peptides 2-2, except one that is named bromodomain-containing protein. Multiple sequence alignment of four of these proteins and Pru du 8 is shown in Figure 1A.

Information about the signal peptide of pecan vicilin was obtained from the N-terminal sequencing results of pecan vicilin purified from the nuts²⁴ and the protein sequence deduced from the pecan vicilin gene. Using the signal peptide of pecan allergen Car i 2 as a query, a BLAST search of the non-redundant protein database returned more antimicrobial proteins (including those that were in the results when searching with Pru du 8) in addition to vicilins. A phylogenetic tree generated with MUSCLE²⁵ based on the sequences of these antimicrobial proteins is shown in Figure 1B.

Purification of Recombinant Pru du 8. The open reading frame was amplified with primers ZY389 and ZY390 and inserted into the vector pTMS1. The expressed protein, His-MBP-sumo-Pru du 8, was the target protein with an N-terminal tag containing poly-His, maltose binding protein (MBP), and sumo. These protein tags were used to enhance the solubility of ssAMP, and sumo was also used as a peptidase recognition site for the removal of the fusion tags. Most of the expressed fusion protein was partitioned in the soluble portion when the lysate was centrifuged. A HisTrap column was used in the immobilized metal affinity chromatography purification of His-MBP-sumo-Pru du 8. After this one-step purification, the sample purity was high enough and the sample was diluted with a Tris buffer containing TDD to make the final condition compatible with ULP digestion. After optimization, ULP was used at a peptidase/substrate molar ratio of 1:800. The reaction was carried out at 20 °C for 4 h and was close to half complete (Figure 2). The recombinant ULP, the cleaved fusion tag, and the undigested fusion protein were removed by passing the digestion product through a Ni²⁺ column. The flow-through contained a strong band at ~31 kDa when

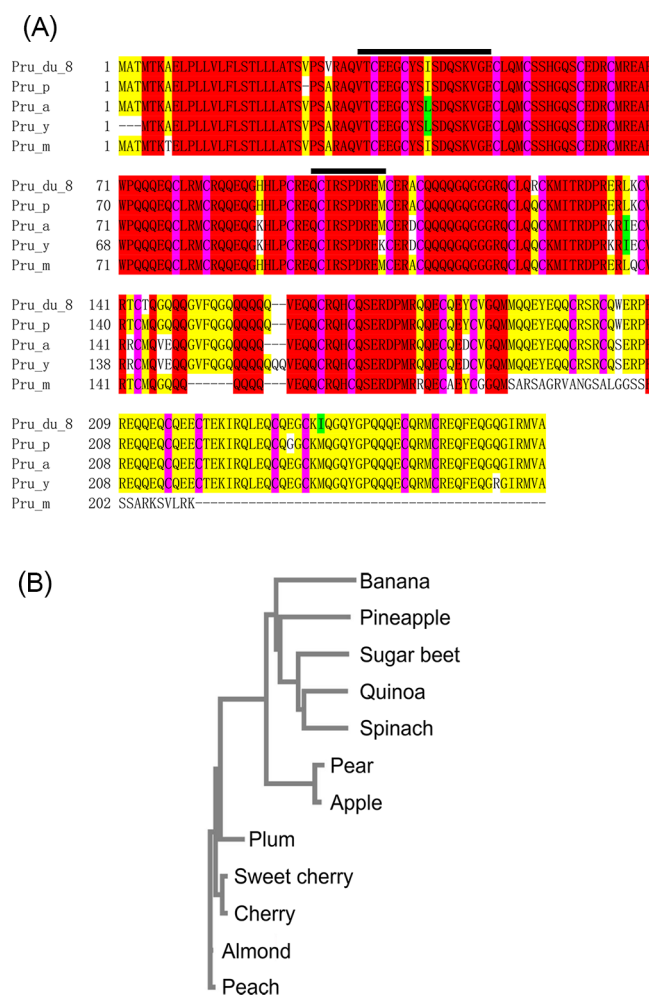


Figure 1. Sequence alignment of ssAMPs. (A) Sequence alignment of ssAMPs from almond (*Prunus dulcis*), peach (*Prunus persica*), sweet cherry (*Prunus avium*), cherry (*Prunus yedoensis* var. *nudiflora*), and Chinese plum (*Prunus mume*). The red background indicates residues conserved in all sequences. Yellow and green backgrounds indicate identical and similar residues in half or more of the sequences at a position, respectively. The Cys residues in the antimicrobial peptide motif are indicated by a magenta background. The black bars on top of the Pru du 8 sequence indicate the peptides that match the N-terminal sequencing results reported by Poltronieri et al.¹⁹ (B) Phylogenetic tree built with ssAMPs from almond (*Prunus dulcis*, MH922028, this study), peach (*Prunus persica*, XP_007223717.1), sweet cherry (*Prunus avium*, XP_021811866.1), cherry (*Prunus yedoensis*, PQQ11123.1), Chinese plum (*Prunus mume*, XP_008224280.1), Chinese white pear (*Pyrus × bretschneideri*, XP_009335337.1), apple (*Malus domestica*, XP_008390980.1), sugar beet (*Beta vulgaris* subsp. *vulgaris*, XP_019107917.1), quinoa (*Chenopodium quinoa*, XP_021716871.1), spinach (*Spinacia oleracea*, XP_021859220.1), banana (*Musa acuminata* subsp. *malaccensis*), and pineapple (*Ananas comosus*, XP_020100752.1).

analyzed by SDS–PAGE (lane 1 in Figure 2), consistent with the theoretical molecular mass of Pru du 8 without an expression tag. The result of 10 cycles of N-terminal amino acid sequencing confirmed the successful removal of the expression tag.

Identification of Almond ssAMP as an Allergen. Sera were collected from 18 patients diagnosed to be allergic to almond by double-blind placebo-controlled food challenges. The patients were enrolled in the food allergy study at Stanford

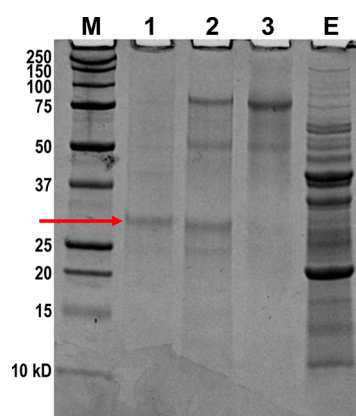


Figure 2. SDS–PAGE analysis of recombinant almond ssAMP. Recombinant ssAMP with a His-MBP-sumo tag after nickel column purification (lane 3), its ULP digestion (lane 2), and ssAMP after tag removal (lane 1) were analyzed with a 4–20% SDS gel, along with almond protein extract (lane E) and protein standards (lane M). The ssAMP band in lane 1 is shown with a red arrow.

University. A 28-year-old adult and 17 pediatric patients ages 3–13 (median of 7) were included.²¹ Some of them were also allergic to walnut.

These sera were individually tested to assess whether they had IgE antibodies specific to almond ssAMP. For the immunoblot experiments, recombinant ssAMP was separated by SDS gels and transferred to PVDF membranes for analysis. Among the 18 sera, 6 were positive for containing IgE antibodies against ssAMP, as shown in the top panel of Figure 3. Sera of non-allergic individuals were used as negative controls in some reported studies on identifying food allergens.²⁶ Such negative controls are especially important when pooled sera are used. However, when sera are individually tested, sera of patients that did not react with the purified allergen can act as negative controls equally well. In the middle and bottom panels of Figure 3, many sera (e.g., sera 17 and 18) were shown not to recognize recombinant ssAMP. These sera could be considered as negative controls for the experiment in detecting positive IgE recognition of recombinant Pru du 8.

Poltronieri and collaborators first reported the serum IgE recognition of a 12 kDa band.¹⁹ In the present study, we could not confidently identify a common band in the almond extract that was recognized by all of the sera in the top panel of Figure 3. A likely reason for this is that the level of ssAMP in almond is low. As such, we did not attempt to identify the natural allergen in the crude extract. To further characterize this allergen, it is desirable to raise an antibody highly specific to recombinant Pru du 8. The antibody can then be used to affinity purify natural Pru du 8 from almond extract for studying the post-translational modifications of the allergen.

In comparison of the signal difference in the extract lanes of the images detected by the sera (e.g., sera 17 and 18), it could be noticed that the total signal strengths from all allergens in the protein extract recognized by each serum were not proportional to the almond-specific IgE determined by ImmunoCap.²¹ One major reason for this discrepancy might be that ImmunoCap measures IgEs that recognize the native proteins, while an immunoblot mainly detects IgEs that recognize linear epitopes. These data may point to a need to develop methods for comparing the prevalence of linear and conformational epitopes recognized by IgEs for studying the

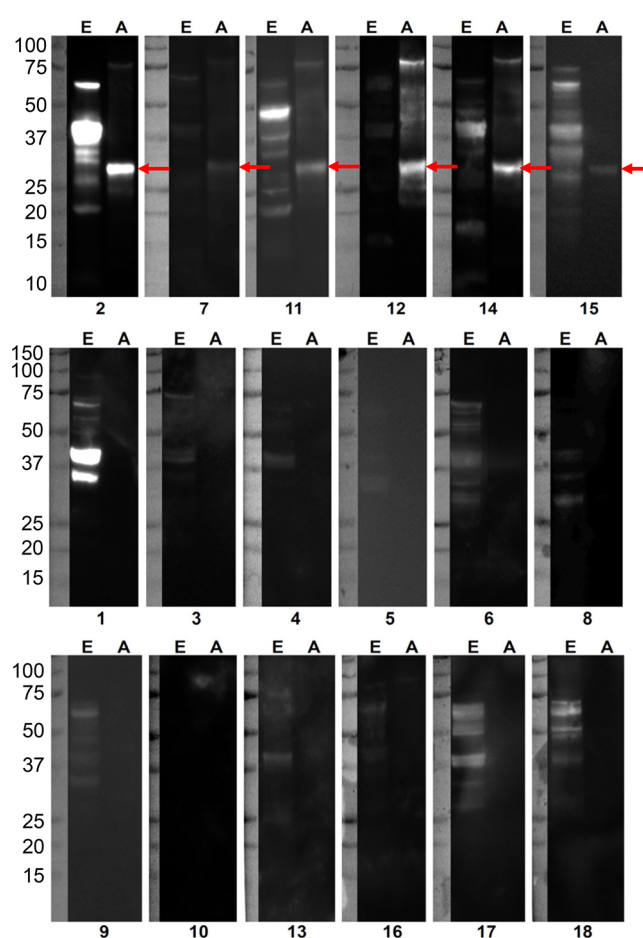


Figure 3. Sera IgE recognition of recombinant Pru du 8. For immunoblot analysis, PVDF membranes with protein bands from almond protein extract (lane E) and ssAMP sample (lane A) were blotted with individual serum from 18 almond-allergic patients. The serum number is labeled at the bottom of the blot. (Top panel) Sera containing IgE that recognize rPru du 8 (as indicated by the red arrows). (Middle and bottom panels) Sera without rPru du 8-specific IgE.

relative importance of different types of IgE epitopes in childhood and persistent food allergies.

DISCUSSION

The literature reported the partial sequences (VTXEEGXYSISDQSKVGE and QXIRSPDREM) of two short peptides from a protein band that reacted with almond-allergic patient serum IgE.¹⁹ This allergen was thought to be the 2S albumin because the sequences of the short peptides share high amino acid identities with those in the 2S albumin of other species.¹⁹ Recently, however, it was debated that this allergen is likely to be the 7S vicilin of almond for the following reasons: (1) The reported short sequences matched 100% with two peptide stretches of the predicted translation of a partial mRNA in the GenBank (BU574985.1) if the uncertain residues in the reported partial peptide sequences were cysteines. (2) The peptide sequence translated from the partial mRNA had homology to the N-terminal region of the 7S vicilin seed storage protein deduced from its coding sequence in some species.²⁰

Using the peptide translated from BU574985.1 as a query, a tBLASTn search of the non-redundant nucleotide database

resulted in three hits. Among them, the vicilin C72-like mRNA (XM_008226058.1) from *Prunus mume* (commonly known as Chinese plum or Japanese apricot) was the only one that also contained 5' and 3' untranslated sequences. We used a primer based on the 5' prime end untranslated sequence and a poly-T primer in the PCR and isolated the cDNA that encoded Pru du 8 shown in Figure 1. The results showed that this protein has sequence homology with the N-terminal domain of vicilins from some species. However, Pru du 8 is not a vicilin because it lacks the two cupin domains that are the main feature of the vicilin protein, although genome annotation with automated computational analysis named the Pru du 8 orthologous proteins in other species as vicilin-like proteins as a result of their sequence similarity to the signal peptide or N-terminal region of vicilins from some species. Some vicilins, such as vicilin of macadamia (*Macadamia integrifolia*), have a N-terminal region that has antimicrobial activity,²⁷ while vicilins from many species do not have this N-terminal region.²⁸

On the other hand, using the signal peptide of pecan vicilin Car i 2 as a query, a BLAST search of the non-redundant protein database not only returned vicilins from different species as expected but also identified many predicted proteins from various species that consist of only the antimicrobial domain. The relationship of these species as suggested by the phylogenetic tree built with these antimicrobial proteins seemed to agree with that of the species classifications, indicating that the antimicrobial proteins are likely in a separate family and not in the cupin protein superfamily. Thus, Pru du 8 is the first member of a new allergen family whose functions require further investigation.

Plant antimicrobial proteins are divided into different families.²⁹ One of them is the α -hairpinin AMP, which assumes a helix–loop–helix structure containing two disulfide bonds between the C3C (short for C[X]₃C) pairs of a characteristic C[X]₃C–[X]_{10–12}–C[X]₃C motif in their peptide sequence.^{29,30} The almond ssAMP isolated in this study contained seven repeats of the motif (Figure 1A and Figure 4A). Pru du 8 orthologous proteins included in Figure 1B contained 5 (Chinese plum) to 12 (apple and chickweed) coupled C3C motifs.

For the vicilins that contain coupled C3C motif(s), their N-terminal antimicrobial peptide regions also contain different numbers of α -hairpinin AMP repeats. The peanut vicilin (Ara h 1)³¹ and almond vicilin have only one α -hairpinin AMP motif in their N-terminal region. While the pecan vicilin Car i 2, which has a 311AA signal peptide,²⁴ consists of five α -hairpinin AMP repeats (Figure 4A) in its signal peptide and one hairpinin AMP motif in the N-terminal of the mature protein. Pine nut vicilin (Pin k 2) and convicilin do not contain the AMP repeat.^{32–34} These results indicate that this coupled C3C motif may be able to expand by element duplication and move to different genes by translocation. On the other hand, the repeats in Pru du 8 and those in Car i 2 do not have high sequence identity (Figure 4B), suggesting that the α -hairpinin AMP repeat might exist very early in the plant evolution and continued to diverge in different species. Apple AMP seemed to contain six repeats of an [X]₁₁–C[X]₃C–[X]₆–C[X]₃C–[X]₉–C[X]₃C–[X]₉–C[X]₃C–[X]₁₂ motif (Figure 4C). This pattern of the separations of the C3C elements does not match that of any of the well-studied AMP families.²⁹ Each of the repeats contained C[X]₃C–[X]₉–C[X]₃C sub-elements that are similar to the coupled C3C motif of the α -hairpinin antimicrobial protein. Whether these repeats or their C3C

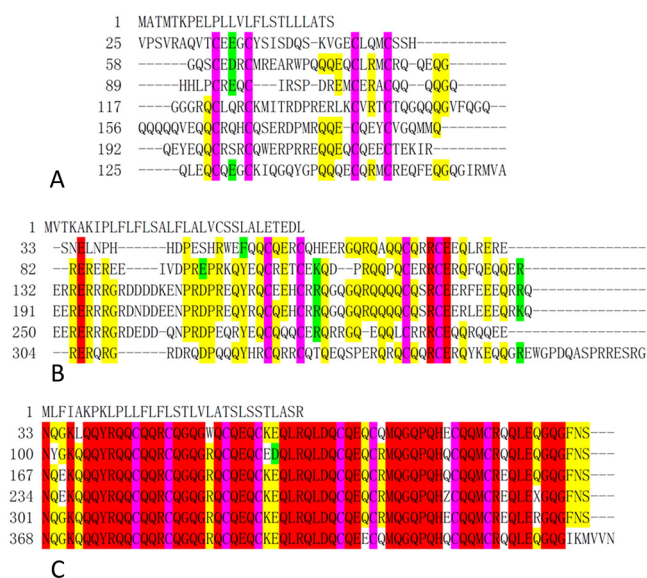


Figure 4. Potential α -hairpinin repeats in ssAMPs and vicilins. (A) α -Hairpinin repeats in almond ssAMP. (B) α -Hairpinin repeats in the signal peptide and the N-terminal region of Car i 2 (see the text). (C) Potential α -hairpinin repeats in apple ssAMP, for which two of the codons in the database showed degenerations likely as a result of low confidence in the genome sequence. Background coloring is the same as in Figure 1A.

elements form similar α -hairpinin structures awaits future investigation.

In conclusion, an almond allergen was previously reported in the literature, but its identity was not clear. It was initially considered to be 2S albumin but was recently debated to be 7S vicilin. Recently, we have isolated an almond vicilin gene and identified it as a food allergen.²¹ However, almond vicilin was shown to be different from the allergen previously reported in 2002. To further characterize almond allergens, the sequence of a mRNA encoding a potentially disulfide-rich antimicrobial peptide/protein was determined. This protein encompasses the peptides identified in the allergen with debated identity. IgE recognition of recombinant ssAMP was tested with 18 sera. The results indicated that almond ssAMP is a food allergen that belongs to a novel family of allergenic food proteins.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jafc.9b02781.

Almond disulfide-bond-rich antimicrobial protein ssAMP (Figure S1) (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*Telephone: 510-559-5981. Fax: 510-559-5818. E-mail: yuzhu.zhang@usda.gov.

ORCID

Yuzhu Zhang: 0000-0001-7882-5692

Tengchuan Jin: 0000-0002-1395-188X

Notes

Disclaimer: Mention of trade names or commercial products in this publication is solely for the purpose of providing specific

information and does not imply recommendation or endorsement by the United States Department of Agriculture. The authors declare no competing financial interest.

■ ABBREVIATIONS USED

IPTG, isopropyl β -D-1-thiogalactopyranoside; ULP, ubiquitin-like protease; PVDF, polyvinylidene fluoride

■ REFERENCES

- (1) Geiselhart, S.; Hoffmann-Sommergruber, K.; Bublin, M. Tree nut allergens. *Mol. Immunol.* **2018**, *100*, 71–81.
- (2) Sicherer, S. H.; Sampson, H. A. Food allergy: Epidemiology, pathogenesis, diagnosis, and treatment. *J. Allergy Clin. Immunol.* **2014**, *133*, 291–307.
- (3) Moore, L. E.; Stewart, P. H.; deShazo, R. D. Food Allergy: What We Know Now. *Am. J. Med. Sci.* **2017**, *353*, 353–366.
- (4) Rehm, C. D.; Drewnowski, A. Replacing American snacks with tree nuts increases consumption of key nutrients among US children and adults: Results of an NHANES modeling study. *Nutr. J.* **2017**, *16*, 17.
- (5) Sampson, H. A.; Mendelson, L.; Rosen, J. P. Fatal and near-fatal anaphylactic reactions to food in children and adolescents. *N. Engl. J. Med.* **1992**, *327*, 380–384.
- (6) Yunginger, J. W.; Sweeney, K. G.; Sturner, W. Q.; Giannandrea, L. A.; Teigland, J. D.; Bray, M.; Benson, P. A.; York, J. A.; Biedrzycki, L.; Squillace, D. L.; et al. Fatal food-induced anaphylaxis. *JAMA* **1988**, *260*, 1450–1452.
- (7) Beardslee, T. A.; Zeece, M. G.; Sarath, G.; Markwell, J. P. Soybean glycinin G1 acidic chain shares IgE epitopes with peanut allergen Ara h 3. *Int. Arch. Allergy Immunol.* **2000**, *123*, 299–307.
- (8) de Leon, M. P.; Drew, A. C.; Glaspole, I. N.; Suphioglu, C.; O'Hehir, R. E.; Rolland, J. M. IgE cross-reactivity between the major peanut allergen Ara h 2 and tree nut allergens. *Mol. Immunol.* **2007**, *44*, 463–471.
- (9) Sathe, S. K.; Liu, C.; Zaffran, V. D. Food Allergy. *Annu. Rev. Food Sci. Technol.* **2016**, *7*, 191–220.
- (10) Flores Kim, J.; McCleary, N.; Nwaru, B. I.; Stoddart, A.; Sheikh, A. Diagnostic accuracy, risk assessment, and cost-effectiveness of component-resolved diagnostics for food allergy: A systematic review. *Allergy* **2018**, *73*, 1609–1621.
- (11) World Health Organization (WHO)/International Union of Immunological Societies (IUIS) Allergen Nomenclature Subcommittee. *Allergen Nomenclature*; <http://www.allergen.org/> (accessed Sept 19, 2018).
- (12) Cox, A.; Sicherer, S. H. Peanut and tree nut allergy. *Chem. Immunol. Allergy* **2015**, *101*, 131–144.
- (13) Sicherer, S. H.; Furlong, T. J.; Munoz-Furlong, A.; Burks, A. W.; Sampson, H. A. A voluntary registry for peanut and tree nut allergy: Characteristics of the first 5149 registrants. *J. Allergy Clin. Immunol.* **2001**, *108*, 128–132.
- (14) Sicherer, S. H.; Burks, A. W.; Sampson, H. A. Clinical features of acute allergic reactions to peanut and tree nuts in children. *Pediatrics* **1998**, *102*, No. e6.
- (15) Tawde, P.; Venkatesh, Y. P.; Wang, F.; Teuber, S. S.; Sathe, S. K.; Roux, K. H. Cloning and characterization of profilin (Pru du 4), a cross-reactive almond (*Prunus dulcis*) allergen. *J. Allergy Clin. Immunol.* **2006**, *118*, 915–922.
- (16) Willison, L. N.; Tripathi, P.; Sharma, G.; Teuber, S. S.; Sathe, S. K.; Roux, K. H. Cloning, expression and patient IgE reactivity of recombinant Pru du 6, an 11S globulin from almond. *Int. Arch. Allergy Immunol.* **2011**, *156*, 267–281.
- (17) Albillos, S. M.; Jin, T.; Howard, A.; Zhang, Y.; Kothary, M. H.; Fu, T. J. Purification, crystallization and preliminary X-ray characterization of prunin-1, a major component of the almond (*Prunus dulcis*) allergen amandin. *J. Agric. Food Chem.* **2008**, *56*, 5352–5358.
- (18) Costa, J.; Mafra, I.; Carrapatoso, I.; Oliveira, M. B. Almond allergens: Molecular characterization, detection, and clinical relevance. *J. Agric. Food Chem.* **2012**, *60*, 1337–1349.
- (19) Poltronieri, P.; Cappello, M. S.; Dohmae, N.; Conti, A.; Fortunato, D.; Pastorello, E. A.; Ortolani, C.; Zacheo, G. Identification and characterisation of the IgE-binding proteins 2S albumin and conglutin gamma in almond (*Prunus dulcis*) seeds. *Int. Arch. Allergy Immunol.* **2002**, *128*, 97–104.
- (20) Garino, C.; De Paolis, A.; Coisson, J. D.; Arlorio, M. Pru du 2S albumin or Pru du vicilin? *Comput. Biol. Chem.* **2015**, *56*, 30–32.
- (21) Che, H. L.; Zhang, Y. Z.; Lyu, S. C.; Nadeau, K. C.; McHugh, T. Identification of almond (*Prunus dulcis*) vicilin as a food allergen. *J. Agric. Food Chem.* **2019**, *67*, 425–432.
- (22) Zhang, Y.; Fan, Y. A Mutant Sumo Facilitates Quick Plasmid Construction for Expressing Proteins with Native N-termini After Tag Removal. *Mol. Biotechnol.* **2017**, *59*, 159–167.
- (23) Zhang, Y.; Du, W. X.; Fregevu, C.; Kothary, M. H.; Harden, L.; McHugh, T. H. Expression, Purification, and Characterization of Almond (*Prunus dulcis*) Allergen Pru du 4. *J. Agric. Food Chem.* **2014**, *62*, 12695–12700.
- (24) 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. Identification and Characterization of a New Pecan [*Carya illinoensis* (Wangenh.) K. Koch] Allergen, Car i 2. *J. Agric. Food Chem.* **2016**, *64*, 4146–4151.
- (25) Edgar, R. C. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **2004**, *32*, 1792–1797.
- (26) Rabjohn, P.; Helm, E. M.; Stanley, J. S.; West, C. M.; Sampson, H. A.; Burks, A. W.; Bannon, G. A. Molecular cloning and epitope analysis of the peanut allergen Ara h 3. *J. Clin. Invest.* **1999**, *103*, 535–542.
- (27) Marcus, J. P.; Green, J. L.; Goulter, K. C.; Manners, J. M. A family of antimicrobial peptides is produced by processing of a 7S globulin protein in *Macadamia integrifolia* kernels. *Plant J.* **1999**, *19*, 699–710.
- (28) Dure, L., III An unstable domain in the vicilin genes of higher plants. *New Biol.* **1990**, *2*, 487–493.
- (29) Tam, J. P.; Wang, S.; Wong, K. H.; Tan, W. L. Antimicrobial Peptides from Plants. *Pharmaceuticals* **2015**, *8*, 711–757.
- (30) Slavokhotova, A. A.; Rogozhin, E. A.; Musolyamov, A. K.; Andreev, Y. A.; Oparin, P. B.; Berkut, A. A.; Vassilevski, A. A.; Egorov, T. A.; Grishin, E. V.; Odintsova, T. I. Novel antifungal alpha-hairpinin peptide from *Stellaria media* seeds: Structure, biosynthesis, gene structure and evolution. *Plant Mol. Biol.* **2014**, *84*, 189–202.
- (31) Burks, A. W.; Cockrell, G.; Stanley, J. S.; Helm, R. M.; Bannon, G. A. Recombinant peanut allergen Ara h 1 expression and IgE binding in patients with peanut hypersensitivity. *J. Clin. Invest.* **1995**, *96*, 1715–1721.
- (32) Zhang, Y. Z.; Du, W. X.; Fan, Y. T.; Yi, J.; Lyu, S. C.; Nadeau, K. C.; McHugh, T. H. Identification, characterization, and initial epitope mapping of pine nut allergen Pin k 2. *Food Res. Int.* **2016**, *90*, 268–274.
- (33) Jin, T.; Wang, Y.; Chen, Y. W.; Fu, T. J.; Kothary, M. H.; McHugh, T. H.; Zhang, Y. Crystal structure of the Korean pine (*Pinus koraiensis*) 7S seed storage protein with copper ligands. *J. Agric. Food Chem.* **2014**, *62*, 222–228.
- (34) Jin, T.; Wang, Y.; Chen, Y. W.; Albillos, S. M.; Kothary, M. H.; Fu, T. J.; Tankersley, B.; McHugh, T. H.; Zhang, Y. Z. Isolation and characterization of Korean pine (*Pinus koraiensis*) convicilin. *Plant Physiol. Biochem.* **2014**, *80C*, 97–104.

```

1  CACACAGTATATATTTGAGCACAAGC ATG GCG ACC ATG ACC AAA GCT GAG CTT CCT
    1▶ M  A  T  M  T  K  A  E  L  P
57  TTG TTG GTT TTG TTC CTA TCC ACT CTT CTC CTT GCT ACC TCT GTG CCT TCT
11▶ L  L  V  L  F  L  S  T  L  L  L  A  T  S  V  P  S
108 GTT CGT GCC CAG GTC ACC TGC GAG GAA GGA TGC TAT TCG ATA TCT GAT CAG
28▶ V  R  A  Q  V  T  C  E  E  G  C  Y  S  I  S  D  Q
    ───────────────────────────────────────────────────────────────────────────
159 TCG AAA GTG GGG GAG TGC TTG CAA ATG TGT TCA TCC CAT GGC CAG TCA TGT
45▶ S  K  V  G  E  C  L  Q  M  C  S  S  H  G  Q  S  C
    ───────────────────────────────────────────────────────────────────────────
210 GAA GAT CGT TGC ATG AGG GAA GCA AGG TGG CCG CAG CAG CAA GAA CAG TGC
62▶ E  D  R  C  M  R  E  A  R  W  P  Q  Q  Q  E  Q  C
261 CTG AGA ATG TGC AGG CAG CAG GAG CAG GGA CAT CAC TTG CCG TGC CGT GAG
79▶ L  R  M  C  R  Q  Q  E  Q  G  H  H  L  P  C  R  E
312 CAA TGC ATC CGG AGT CCA GAC CGA GAG ATG TGC GAA AGA GCT TGC CAG CAG
96▶ Q  C  I  R  S  P  D  R  E  M  C  E  R  A  C  Q  Q
    ───────────────────────────────────────────────────────────────────────────
363 CAG CAA GGT CAA GGA GGA GGG CGC CAG TGC CTG CAA CGA TGT AAG ATG ATA
113▶ Q  Q  G  Q  G  G  G  R  Q  C  L  Q  R  C  K  M  I
414 ACC AGA GAT CCG CGG GAG CGG CTA AAG TGT GTA AGA ACG TGC ACG CAG GGG
130▶ T  R  D  P  R  E  R  L  K  C  V  R  T  C  T  Q  G
465 CAG CAG CAG GGT GTT TTT CAA GGG CAG CAG CAG CAG CAG CAA GTT GAG CAG
147▶ Q  Q  Q  G  V  F  Q  G  Q  Q  Q  Q  Q  Q  V  E  Q
516 CAG TGC CGG CAA CAT TGC CAA TCG GAG CGC GAC CCA ATG AGA CAG CAA GAG
164▶ Q  C  R  Q  H  C  Q  S  E  R  D  P  M  R  Q  Q  E
567 TGT CAG GAA TAC TGT GTT GGG CAA ATG ATG CAG CAG GAA TAC GAG CAG CAG
181▶ C  Q  E  Y  C  V  G  Q  M  M  Q  Q  E  Y  E  Q  Q
618 TGC CGG TCG CGT TGC CAA TGG GAG CGC CCC AGG AGG GAG CAG CAG GAG CAG
198▶ C  R  S  R  C  Q  W  E  R  P  R  R  E  Q  Q  E  Q
669 TGC CAG GAA GAG TGT ACT GAG AAA ATA AGG CAG CTA GAG CAA TGC CAG GAA
215▶ C  Q  E  E  C  T  E  K  I  R  Q  L  E  Q  C  Q  E
720 GGG TGC AAG ATC CAG GGG CAA TAT GGT CCA CAG CAA CAG GAG TGT CAA AGG
232▶ G  C  K  I  Q  G  Q  Y  G  P  Q  Q  Q  E  C  Q  R
771 ATG TGT AGG GAG CAG TTT GAA CAA GGA CAA GGA ATT AGA ATG GTA GCC TGA
249▶ M  C  R  E  Q  F  E  Q  G  Q  G  I  R  M  V  A
822 TGCAATAGAAGTTGTATGCATGTAAGTAGTTGGCTGTGCTAGAGTTTAAGAGAGTATTATAGCTCATG
890 GAGAGTGAGCTTTAACTTAAATTGTATGTGTGAAATAAAGAGCTCTAGCTTAATAAACTACAATGAT
958 GACAAGTTCTATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

```

Fig. S1. Almond disulfide bond rich antimicrobial protein ssAMP. The mRNA sequence is shown in red letters. The translation of the open reading frame is shown in black with one-letter code for the amino acids. The underlines mark the regions that contain the peptide sequences partially determined in a previous report of an allergenic almond protein by Poltronieri et al. The mRNA sequence is deposited in the Genbank with accession number MH922028.