JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY



# Purification and Characterization of the 7S Vicilin from Korean Pine (*Pinus koraiensis*)

TENGCHUAN JIN,[†] SILVIA M. ALBILLOS,[‡] YU-WEI CHEN,[†] MAHENDRA H. KOTHARY,[§] TONG-JEN FU,[||] AND YU-ZHU ZHANG*,[†]

Department of Biological, Chemical, and Physical Sciences, Illinois Institute of Technology, Chicago, Illinois 60616; Illinois Institute of Technology, National Center for Food Safety and Technology, 6502 South Archer Road, Summit-Argo, Illinois 60501; Center for Food Safety and Applied Nutrition, U.S. Food and Drug Administration, Laurel, Maryland 20708; and U.S. Food and Drug Administration, National Center for Food Safety and Technology, 6502 South Archer Road, Summit-Argo, Illinois 60501

Pine nuts are economically important as a source of human food. They are also of medical importance because numerous pine nut allergy cases have been recently reported. However, little is known about the proteins in pine nuts. The purpose of this study was to purify and characterize pine nut storage proteins. Reported here is the first detailed purification protocol of the 7S vicilin-type globulin from Korean pine (*Pinus koraiensis*) by gel filtration, anion exchange, and hydrophobic interaction chromatography. Reducing SDS-PAGE analysis indicated that purified vicilin consists of four major bands, reminiscent of post-translational protease cleavage of storage proteins during protein body packing in other species. The N-terminal ends of vicilin peptides were sequenced by Edman degradation. Circular dichroism (CD) and differential scanning calorimetry (DSC) analyses revealed that pine nut vicilin is stable up to 80 °C and its folding−unfolding equilibrium monitored by intrinsic fluorescence can be interpreted in terms of a two-state model.

**KEYWORDS:** ***Pinus Koraiensis*; pine nut; 7S globulin; vicilin-like seed storage protein; food allergen**

## INTRODUCTION

Pines (*Pinus genus*) are gymnosperms. They are evolutionarily important as they represent an ancient evolutionary branch; they are economically important as they provide a source of timber, wood pulp, and food, and they are medically relevant as pine nut allergy poses a risk to a subpopulation of consumers.

Gymnosperms are evolutionarily ancient, arising as early as 300 million years ago. The conifers, a division to which the pines belong, separated from flowering plants (angiosperms) about 100 million years ago and by most measures have evolved very slowly and conservatively (*1–3*). Pines therefore represent an ancestral branch in the tree of life.

Pines are among the most widely distributed and prominent genera of trees in the world. They extend throughout the Northern Hemisphere from the equator to the Arctic Circle and are the dominant vegetation over large areas (*4*). Pines are the most important forest trees and crop species in the United States and around the globe. One single species, Loblolly pine (*Pinus taeda* L.), alone provides ∼16% of the world's annual timber supply (*5*) and grows on nearly 58 million acres of plantation and natural forest in the southeastern United States (*6*). Pine timber, including softwood and hardwood, is widely used in logs and lumber, as well as in the pulp and paper industry.

Pine nuts are high in nutritional value. For instance, Stone pine (*Pinus pineal*) seeds have a protein content up to 31.6% of the total dry weight of the edible portion (*7, 8*). They also contain a considerable amount of vitamins A, B1, and B2, potassium, magnesium, and other minerals as well as dietary fiber (*7, 9*). Pine nuts are widely used for human consumption in Asia, Europe, and America. They are eaten either raw or roasted and have been used as ingredients in a variety of traditional recipes. Today, pine nuts are harvested in many regions of the northern hemisphere and are marketed globally as a gourmet product. The annual sales of pine nuts in the U.S. market total about 100 million dollars, 80−90% of which is imported, primarily from China where the most widely harvested species is Korean pine (*Pinus koraiensis*) (*10*), which is the most important species in today's international pine nut trade.

Realizing the importance and the lagging status of the conifer genomic study, compared with other agricultural crops, several genomic projects have been launched worldwide (*3*). In the United States, the Loblolly pine genome project started in 1999

* To whom correspondence should be addressed. Phone: 312-567-3484. Fax: 312-567-3494. E-mail: zhangy@iit.edu.

[†] Department of Biological, Chemical, and Physical Sciences, Illinois Institute of Technology.

[‡] Illinois Institute of Technology, National Center for Food Safety and Technology.

[§] U.S. Food and Drug Administration, Center for Food Safety and Applied Nutrition.

[||] U.S. Food and Drug Administration, National Center for Food Safety and Technology.

10.1021/jf801138q CCC: $40.75 
Published on Web 08/09/2008

(*11*). The haploid Loblolly pine genome was estimated to be 22 pg, seven times larger than the human genome (*5*). To date, a complete pine genome has not been reported.

Similar to other tree nuts and peanut, pine nut is also known to be a source of food allergens (*21, 22*). Numerous food allergies are caused by seed storage proteins from nuts. Seed storage proteins account for >5% percent of the total seed mass and accumulate during seed development. They are believed to be rapidly broken down and metabolized during germination and early seedling growth (*21*). On the basis of their solubility and ultracentrifugation sedimentation coefficients (S), common seed storage proteins are divided into 2S albumin, 7S vicilin, and 11−13S globulins. Many vicilin-type seed storage proteins in legume and tree nuts have been identified as allergens (*22–26*) including Ara h 1 in peanut (*27*), Ana o 1 in cashew nut (*23*), Cor a 11 in hazelnut (*28*), and Jug r 2 in English walnut (*26*). Although pine nut allergy cases have been reported (*12–20*), pine nut allergens have not been identified or characterized.

Seed storage proteins of some pine species have been isolated, mostly by means of buffer extraction. SDS-PAGE analyses of these proteins in mature seed or during development have been reported (*29*). The megagametophytes of *Pinus pinaster* seeds contain two types of oligomeric 7S vicilins of approximately 175 and 190 kDa, respectively. They are composed of 47, 27, and 22 kDa subunits. N-terminal amino acid sequencing showed that the 27 and 22 kDa peptides were originated from partial proteolytic processing of a single storage protein, with the 27 kDa peptide at the N-terminal end (*29*). Here, a 7S vicilin-type seed storage protein from Korean pine nut was purified and identified by SDS-PAGE and N-terminal amino acid sequencing; its thermal as well as chemical stabilities were investigated.

## MATERIAL AND METHODS

**Preparation of Pine Nut Seed Extract.** Pine nut seeds were purchased from a local vender and homogenized in 10 volumes (w/v) of 0.01 M phosphate buffered saline pH 7.4 (PBS), containing protease inhibitors (100 nM aprotinin, 50 $\mu$M antipain, 50 $\mu$M leupeptin, and 0.5 mg/ml pepstatin, all purchased from Sigma, St. Louis, MO) and antibiotics (50 $\mu$g/ml ampicillin and 50 $\mu$g/ml kanamycin, both from Invitrogen, Carlsbad, CA) in a KitchenAid blender. The homogenate was defatted by adding an equal volume of cyclohexane with stirring at 4 °C for 4 h. After centrifugation, the upper layer was discarded, and the aqueous portion was defatted two more times with cyclohexane. The aqueous layer after the last extraction was collected and centrifuged at 12 000*g* at 4 °C for 15 min. The supernatant was filtered with 0.45 $\mu$m syringe filters and was either used for purification by FPLC or stored at −80 °C for further use.

**Purification of Pine Nut Vicilin by FPLC.** The 7S vicilin-type storage protein was purified from defatted pine nut extract at 4 °C by sequential gel filtration, anion exchange, hydrophobic interaction chromatography (HIC), and a second gel filtration step. For the gel filtration step, a crude extract from an equivalent of 5 g of dry pine nut seeds was loaded onto a 300 mL Superdex 200 column (XK 26/70, GE Healthcare, Piscataway, NJ) pre-equilibrated with buffer A (10 mM Tris-HCl, pH 7.9) containing 100 mM NaCl, and the proteins were eluted with the same buffer. The Superdex column was calibrated using thyroglobulin (669 kDa), ferritin (440 kDa), catalase (242 kDa), aldolase (158 kDa), and bovine serum albumin (67 kDa) (GE Healthcare) as molecular weight standards.

Fractions containing pine nut vicilin eluted from the gel filtration column were pooled and dialyzed overnight against buffer A. The sample was then loaded onto an 8 mL Source 15Q anion exchange column (GE Healthcare) pre-equilibrated with buffer A. The column was then washed with buffer A and eluted with buffer A plus a linear gradient of NaCl (0−0.3 M) by mixing buffer A with buffer B (10 mM Tris-HCl, pH 7.9, 1 M NaCl).

The vicilin containing fractions eluted from the anion exchange column were pooled, and ammonium sulfate was added to the pooled sample to a final concentration of 2.5 M. The sample was then loaded onto two interlocked phenyl Sepharose columns (2 × 5 mL, GE Healthcare) pre-equilibrated with HIC binding buffer (2.5 M ammonium sulfate, 10 mM Tris-HCl, pH 7.9). The HIC column was eluted with a 2.5−0 M ammonium sulfate gradient in buffer A. The fractions of the major peak eluted from the column were pooled, and the pooled sample was concentrated and reapplied to the Superdex 200 column as described above to obtain a purified vicilin preparation.

Purified pine nut vicilin was pooled and concentrated in an Amicon ultra-15 centrifugal filter device with 10K nominal molecular weight limit (Millipore, Bedford, MA), and the protein concentration was estimated with a Cary 300 UV−visible spectrophotometer using a theoretical extinction coefficient of $\varepsilon_{280}$ = 34 870 $cm^{-1}$ $mol^{-1}$ calculated according to the vicilin sequence of Loblolly pine.

**Reducing and Nonreducing SDS-PAGE Analysis.** The protein samples were analyzed with 10−20% polyacrylamide gels in Tris-HEPES-SDS running buffer (100 mM Tris, 100 mM HEPES, 3 mM SDS, pH 8.0). Prestained protein molecular weight marker (Precision Plus All Blue, Bio-Rad, Hercules, CA) containing protein standards of 10−250 kDa was used as reference. For reducing SDS-PAGE, 288 mM $\beta$-mercaptoethanol was added to 2× nonreducing sample buffer (20% glycerol, 150 mM Tris-HCl, pH 6.8, 4 mM EDTA, 4% SDS, 0.4% (w/v) bromophenol blue dye). Samples were mixed (1:1, v/v) with the reducing/nonreducing sample buffer, boiled for 5 min, and then cooled on ice before loading onto the gel. The gel was stained in a solution containing 0.1% Coomassie Brilliant Blue R-250, 40% ethanol, and 10% acetic acid for 2 h and destained within a solution with 40% ethanol and 10% acetic acid. Images of the gels were taken by a Nikon digital camera and analyzed with the ImageJ software (*30*).

**Peptide N-Terminal Amino Acid Sequencing.** Purified pine nut vicilin was subjected to SDS-PAGE using 8−25% gradient gels in a PhastSystem (GE Healthcare) and electrophoretically transferred (0.35 A at 25 °C) to a Problott membrane (Applied Biosystems, Foster City, CA) in a transfer buffer (10 mM 3-(cyclohexyl-amino)-1-propane sulfonic acid containing 10% methanol, pH 11.0). The blot was stained with Coomassie Brilliant Blue, and the four main peptide bands were excised and subjected to N-terminal amino acid sequencing by Edman degradation using a Procise Model 491 Protein Sequencer (Applied Biosystems). Cysteine residues are chemically labile during Edman degradation, and no attempts were made to derivatize those to their PE-Cys form so that they could be analyzed.

**Circular Dichroism Spectroscopy.** Circular dichroism (CD) spectra were acquired on a Jasco J715 spectrometer. The spectrometer was calibrated using a camphorsulfonic acid solution according to the manufacturer's instructions. Each spectrum was recorded as an average of 10 accumulations scanning from 250 to 190 nm (50 nm/min), using a 1 mm path length cuvette. Prior to CD spectrum analysis, sample buffer was changed to PBS, and the protein concentration was estimated by its absorbance at 280 nm.

**Differential Scanning Calorimetry (DSC).** DSC measurements were performed using a Mettler Toledo (Columbus, OH) model DSC823$^{e}$/700 differential scanning calorimeter. The protein concentrations were 16.03 mg/mL and 26.16 mg/mL in PBS as determined by BCA protein assays using a BCA kit from Pierce Biotechnology Inc. (Rockford, IL). High sensitivity was achieved by using a MultiSTAR HSS7 sensor with 120 thermocouples in a multilayer arrangement, which amplifies the original signal and reduces the signal-to-noise ratio. Indium was employed as a standard for internal calibration of the instrument. All measurements were performed using 30 $\mu$L HP stainless steel crucibles without pin and membrane gold platted covers as sample pans. A second crucible filled with 30 $\mu$L of buffer was used as a reference. The samples were heated from 25 to 180 °C at a scan rate of 1 °C/min. The onset denaturation temperature $T_o$, the temperature at which denaturation starts, was determined by taking the intercept of the baseline and the extrapolated front edge of the peak. The temperature at the maximum of the peak was regarded as the denaturation temperature $T_d$.

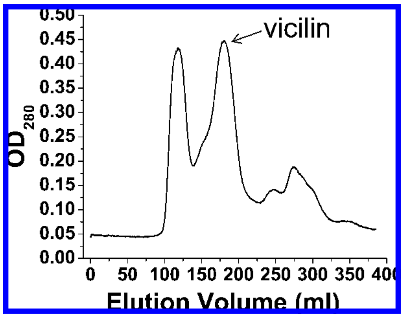


**Figure 1.** Gel filtration of crude pine nut PBS extract. Dialyzed pine nut PBS crude extract (20 mL) was loaded onto a 300 mL Superdex-200 gel filtration column and eluted with a Tris buffer (10 mM, pH 7.9) containing 100 mM NaCl. Vicilin was eluted in the second major peak.

**Fluorescence Spectroscopy.** Fluorescence spectra were obtained using an ISS PCI spectrometer. A series of 25 vicilin samples at folding−unfolding equilibrium were prepared by mixing equal molar stock solutions of native (in PBS) and fully unfolded (in 8.5 M urea) vicilin at different ratios. The samples were allowed to equilibrate at 4 °C overnight before fluorescence measurement. Fluorescence emission spectra were recorded from 320 to 380 nm. The excitation was centered at 280 nm with a bandwidth of 2 nm. All fluorescence measurements were carried out using a 4 mm square quartz cuvette thermostatted at 20 °C. Fluorescence spectra of PBS and 8.5 M urea buffered in PBS were also acquired and were used to construct the baseline of the protein spectra at different urea concentrations.

## RESULTS AND DISCUSSION

**Purification of Pine Vicilin.** In pine nut, vicilin is among the proteins that can be extracted by buffers with low ionic strength (e.g., PBS or 10 mM Tris-HCl, pH 7.9). As shown in **Figure 1**, after gel filtration chromatography, the extracted proteins were separated into four major peaks. As indicated by SDS-PAGE analysis, the largest peak was eluted at ∼188 mL. By using the calibration curve obtained based on the mobility of the molecular weight standards, the apparent molecular weight of vicilin was calculated to be 137 kDa. Further purification, SDS-PAGE analysis, and N-terminal peptide sequencing (see below) indicated that this peak contained vicilin (**Figure 2**, lane 1).

As shown in **Figure 3**, pine nut vicilin was able to bind to the Source 15Q column equilibrated with a buffer of low ionic strength. After the column was washed extensively with buffer A, the protein was eluted with a NaCl gradient, at ∼0.20 M NaCl. However, SDS-PAGE analysis indicated that the vicilin sample collected after anion exchange chromatographic purification still contained a significant amount of impurities (**Figure 2**, lane 2). To further purify pine nut vicilin, two interlocked phenyl Sepharose columns were used. This hydrophobic interaction purification step effectively removed many unwanted proteins as shown in **Figure 4** and **Figure 2**, lane 3.

In order to remove salt and other impurities that were not separated well during the first gel filtration step, the vicilin containing fractions from the phenyl Sepharose columns were purified again using the same gel filtration column. The vicilin sample collected from this additional step consisted of four major bands with molecular masses of ∼45, 34, 25, and 13 kDa, respectively. The purified vicilin has been crystallized for structure determination (*31*) and further SDS-PAGE analysis of a sample redissolved from the vicilin crystals contained the same banding pattern (**Figure 5**), confirming that all four peptides were the result of post-translational processing of the protein. The banding pattern is consistent with those of vicilin-type storage proteins from other species (*32, 33*). No difference

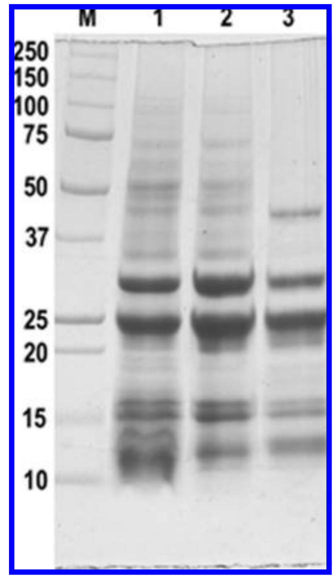


**Figure 2.** SDS-PAGE analysis of pine nut vicilin purification. Pooled vicilin containing fractions from the gel filtration (lane 1), anion exchange (lane 2), and the hydrophobic interaction chromatography (lane 3) were separated with a 4−20% SDS-gel. The molecular masses of the reference bands in the marker (Precision Plus, Bio-Rad, lane M) are given to the left of the gel.

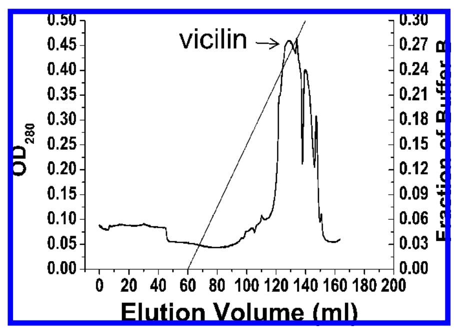


**Figure 3.** Anion exchange purification of Korean pine vicilin. The vicilin containing peak from the gel filtration step was pooled and loaded onto an 8 mL Source 15Q anion exchange column pre-equilibrated with the 10 mM Tris buffer. The column was eluted using the Tris buffer with a linear NaCl gradient (0−0.3 M over 75 mL) by mixing the Tris buffer with buffer B (10 mM Tris, pH 7.9, 1 M NaCl). The vicilin peak is indicated.

was observed between the migration patterns of the bands in reduced and nonreduced SDS gels, indicating that there was no disulfide bond connection between any of the peptides revealed in the SDS gel. These results are also consistent with reported data for vicilins from other species (*32, 34*). Vicilins are trimeric proteins of MW ranging from ∼150 to 190 kDa, with subunits of ∼50 kDa. No disulfide bond was reported, but proteolytic processing and glycosylation were believed to occur (*32, 35*).

**Identification of the N-Terminal Amino Acid of the Peptides.** In order to determine whether the multiple bands in the purified vicilin sample resulted from a single gene product, we sought to identify the peptides represented by the four major bands by N-terminal amino acid sequencing. The N-terminal amino acid sequences of the different bands determined by Edman degradation are shown in **Table 1**. The 45 and 34 kDa bands had the same N-terminal sequence, GRREEEREEN-PYVFHSD, indicating that these two bands resulted from the same protein. **Table 1** also indicates that the 25 kDa band represents two peptides that migrated together and one of these

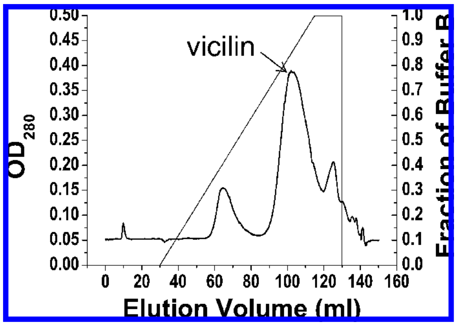


**Figure 4.** Hydrophobic interaction chromatographic purification of pine vicilin. A vicilin sample in 2.5 M ammonium sulfate was loaded onto a 10 mL phenyl-sepharose column. After washing with the binding buffer (10 mM Tris-HCl, pH 7.9, 2.5 M ammonium sulfate), the column was eluted with an 80 mL linear ammonium sulfate gradient of 2.5−0 M by mixing the binding buffer with buffer A (10 mM Tris-HCl, pH 7.9). The vicilin containing peak is indicated in the graph.

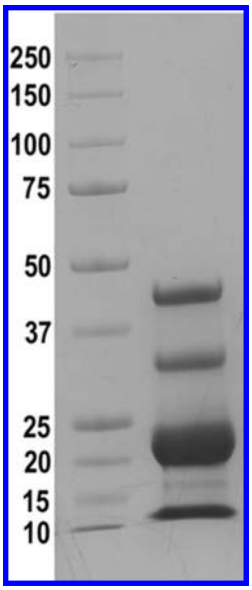


**Figure 5.** SDS-PAGE analysis of the pine nut vicilin protein crystals. This sample was used to determine the N-terminal amino acid sequences of the major peptides in pine nut vicilin. The molecular masses of the reference bands in the protein marker are shown on the left side of the gel.

shared the same N-terminal amino acid sequence with the 45 and 34 kDa peptides, again indicating that these peptides are from the same protein. The 13 kDa band may contain more than two peptides as judged by the N-terminal sequencing results. However, the N-terminal sequencing results as a whole suggested that all four major bands resulted from the processing of the vicilin protein when compared with the known sequence of vicilin from Loblolly pine. As shown in **Figure 6**, by searching the EST databases at the NCBI and the contig sequences (*36*) of Loblolly pine, the N-terminal sequences of the 45, 34, and one of the 25 kDa peptides could be matched to that of aa56-aa73 of the Loblolly pine vicilin with one deletion. The mismatch between the sequence of the vicilin we purified and that translated from the contig sequence of the Loblolly pine vicilin could be due to a sequencing error in the EST sequence or due to errors in our N-terminal amino acid sequencing. However, both of these are unlikely because the N-terminal amino acid sequencing was performed multiple times from each of the three bands and the contig also covered multiple entries of EST sequences. Surprisingly, this sequence

**Table 1.** N-Terminal Amino Acid Sequences of Korean Pine Vicilin

| Resi # | 45 kDa | 34 kDa | 25 kDa | 13 kDa |
|---|---|---|---|---|
| 1 | G | G | GA | GASTE |
| 2 | R | R | RS | HP |
| 3 | R | R | RA | QGTYA |
| 4 | E | E | EQ | EHQDVT |
| 5 | E | E | EP | IP |
| 6 | E | E | EK | NDAI |
| 7 | R | R | RP | YIPN |
| 8 | E | E | EF | NLYFE |
| 9 | E | E | EN | RNY |
| 10 | N | N | NL | VLR |
| 11 | P | P | PN | RRVL |
| 12 | Y | Y | Y | A |
| 13 | V | V | V | R |
| 14 | F | F | F | L |
| 15 | H | H | H | R |
| 16 | S | S | S | T |
| 17 | D | D | D | G |
| 18 | | | | T |

of 17 amino acids is identical to a sequence in the vicilin-like storage protein of white spruce (*Picea glauca*). These results raised the question whether the Korean pine vicilin is more closely related to that in the white spruce than to that in Loblolly pine.

The 25 kDa band also appears to contain a peptide with an N-terminal amino acid sequence (ASAQPKPFNLN) that closely matches (7 out of 11 of) that of the *Pinus taeda* vicilin (*36*) starting from aa262. Furthermore, this sequence also appears to be at the N-terminal end of a peptide contained in the 13 kDa band. The 13 kDa band also appears to contain a different peptide with an N-terminal sequence (G[HP]QEINY[NLY-E]RVRARLRTGT) that closely matches a peptide in the Loblolly vicilin starting from aa355. Considering the challenge of N-terminal amino acid sequencing when a band consists of more than one peptide, these data strongly indicated that the four bands in the SDS-gel of the purified vicilin sample were all part of the same protein. Although the exact sequence of Korean pine vicilin is yet to be determined, these data indicated that vicilin was post-translationally cleaved at the peptide bond before G56 and partially cleaved before A262 and G355, in reference to the Loblolly pine sequence (**Figure 6**). It was unlikely that the cleavages occurred during purification because protease inhibitor cocktails were added, these vicilin bands were also present in the crude protein extract, and they accounted for a large percentage of the whole protein content.

**Urea Induced Denaturation.** The urea-induced folding−unfolding transitions of pine nut vicilin were measured by monitoring the intrinsic fluorescence signals of vicilin as a function of urea concentration (**Figure 7A**). For the spectrum at each urea concentration, the "center of mass" wavelength of the emission was calculated with the following equation:

$$<\lambda> = (\sum\lambda_i \times I_i)/\sum I_i$$

where $\lambda_i$ is the wavelength of the $i$-th data point and $I_i$ the fluorescence intensity at $\lambda_i$. The "center of mass" wavelength is closely linked to the "Red shift" induced during urea-induced unfolding. For each sample, the fractions of native and unfolded protein were calculated by assuming that $<\lambda>$ was a weighted average of that for the native and unfolded protein. The fractional populations of the native and unfolded protein (denoted $f_n$ and $f_u$, respectively) were then used to calculate the free energy of unfolding, $\Delta G = kT\ln(f_u/f_n)$. In analyzing the data, a two-state model with a linear dependence of the

```
  1 MAFVSLLTIL LAISSCSVAL TEPLATVANQ GVFPEDHGRG HHGVFPEEHG RGHRRGRREE EEREENPYVF
 71 HSDRFRMRAS SEAGEIRALP NFGEASELLE GISKYRVTCI EMRPNTVMLP HYLDATWILY VTGGRGYIAY
141 VHQNELVKRK LEEGDVFGVP RGHTFYLVNN DDHNSLRITS LLRTVSTMRG EYEPYYVAGG RNPETVYSAF
211 SDDVLEAAFN TNVEKLEHIF GAHRRGVIFH ANEEQIREMM RREGFSAESM SASKHPKPLN LRNQKPDFEN
281 DNGRFTRAGP NENPLLDAVD VTVGFGVLNP GTMTAPSHNT KATSIAIVTQ GEGRIEMACP HLGQQHGWSS
351 RREKGDQEIN YQRVRARLRT GTVYVVPAGH PITEIACTEG HLEILWFDIN TSGNERQFLA GKNNVLQTLE
421 KEVRQISFNI PRGEELDEVL RRQKDQVILR GPQMQRRDEP RSSS
```

**Figure 6.** N-terminal amino acid sequencing of pine nut vicilin. The amino acid sequence of the open reading frame (ORF) of *Pinus taeda* viclilin (*36*) is shown. Except for one insertion, 17 cycles of Edman degradation of more than one band in *Pinus koraiensis* vicilin resulted in a peptide sequence match that of a peptide in *Pinus taeda* vicilin (underlined residues). Additional N-terminal sequences of peptides in the vicilin bands (see **Table 1**) are also closely matched to amino acid sequences (shaded or double underlined residues) in the ORF (see text).

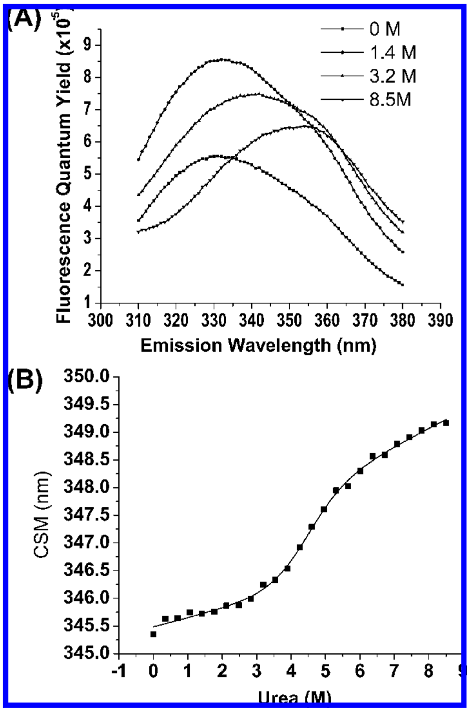


**Figure 7.** Urea induced pine nut vicilin denaturation. (**A**) Typical fluorescence emission spectra of pine vicilin at different urea concentrations. The excitation beam was fixed at 280 nm with a 2 nm bandwidth. (**B**) The "center-of-mass" wavelength of the vicilin fluorescence emission spectrum as a function of urea. The solid line is the two-state fit of the folding-unfolding equilibrium.

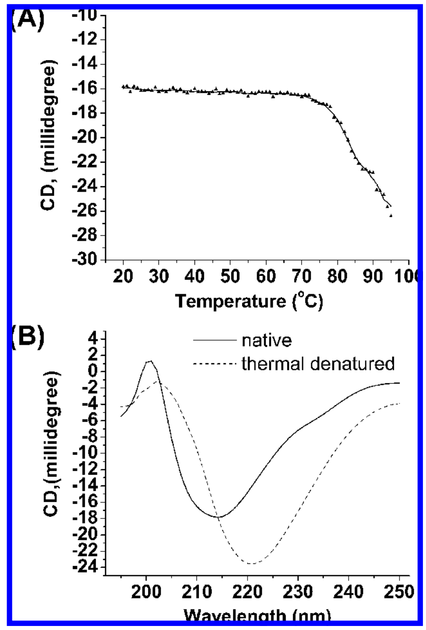


**Figure 8.** Thermal denaturation of pine nut vicilin monitored by CD. (**A**) Irreversible conformation change of pine nut vicilin induced by heat. The CD signal at 220 nm of a vicilin sample with an $OD_{280}$ of 0.325 was used to monitor the conformation change of pine vicilin. (**B**) CD spectra of native and thermal denatured pine vicilin. The spectra of a Korean pine vicilin sample before (solid line) and after (dotted line) thermal denaturation are shown. Both spectra were recorded at 20 °C.

unfolding free energy ($\Delta G$) on the concentration of the denaturant was assumed. This model is characterized by a midpoint concentration ($C_m$), the denaturant concentration at which $\Delta G = 0$, and the slope of the $\Delta G$ dependence on the denaturant concentration ($m$):

$$\Delta G = m(C_m - c) = \Delta G^0 - mc$$

where $\Delta G^0$ represents the free energy of unfolding in the absence of a denaturant and can be derived by a linear extrapolation. As shown in **Figure 7B**, the folding−unfolding process of pine nut vicilin is mainly an equilibrium process between the native and the denatured states. Fitting the data with two transitions slightly improved the fitting statistics but did not significantly change the fitting parameters for the main transition. The midurea-concentration for the two-state model was 4.40 ± 0.13 M, and the extrapolated $\Delta G^0$ was 19.04 ± 3.29 kJ/mol. These data indicated that, as judged by urea denaturation, this 7S protein is not unusually stable compared to other globular proteins.

**Circular Dichroism Spectroscopy.** In order to assess the thermal stability of the pine nut vicilin, far UV CD was used to monitor its secondary structure content as a function of temperature. As shown in **Figure 8**, the circular dichroism spectrum indicated that vicilin is a well-structured protein. Although the content of each of the secondary structure types is often difficult to calculate accurately from CD data, it can be seen from **Figure 8** that pine nut vicilin is composed mainly of $\beta$-sheets. This is consistent with the known secondary structures of a number of 7S storage proteins from different types of beans (soybean (*Glycine max*) (*37, 38*), Jack bean (*Canavalia pubescens*) (*39, 40*), mung bean (*Vigna radiata* L. Wilczek), and French bean (*Phaseolus vulgaris*) *41, 42*). However, monitoring the CD signal as a function of temperature did not detect a transition from regular secondary structures to random coil up to 94 °C. On the contrary, the CD signal at 220 nm started to increase sharply at ∼75 °C. There was no observable precipitation up to 94 °C, and this transition was not a reversible process. This phenomenon has also been reported for vicilins from other species (*43*) when the CD signal at 208 nm was monitored and the increase in optical activity was attributed to increased $\beta$-structures upon thermal denaturation. However, our data indicated that the largest CD signal change due to the thermal transition occurred at ∼220 nm. Further investigation is needed in order to understand the mechanism of this transition at the molecular level.

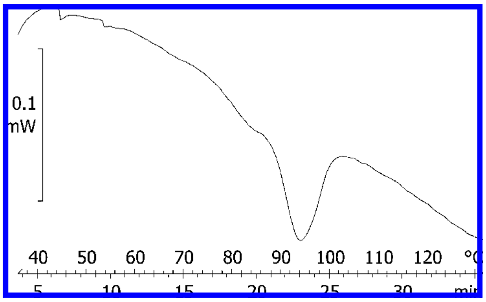


**Figure 9.** Thermal denaturation of pine nut vicilin monitored by DSC. The temperature was scanned at a rate of 1 °C/min.

**Thermal Denaturation of 7S Vicilin Monitored by DSC.** As shown in **Figure 9**, DSC analysis showed an endotherm peak for the 7S pine nut vicilin with an onset temperature of $T_o$ = 84.13 ± 2.25 °C and denaturation temperature of 93.23 ± 1.36 °C. A gradual increase in the peak temperature of an oligomeric protein with the increase in protein concentration has been reported for the rabbit muscle glyceraldehyde-3-phosphate dehydrogenase (*44*). In the present study, however, neither the endotherm peak temperature nor the calculated enthalpy of the pine nut vicilin thermal transition was affected by the protein concentration between 0.5 and 3.0 mg/mL. Thus, the dependence of the denaturation temperature on protein concentration observed previously (*44*) does not seem to be a general property for oligomeric proteins. Rescanning of the sample for a second time after thermal denaturation showed a DSC curve without thermal transitions that involved heat exchange, indicating that the process of heat-induced denaturation was irreversible. A gel-like disk of protein aggregates formed inside the pressure sealed sample pan after heat-induced denaturation. No further thermal transition was observed up to 180 °C, although the CD data indicated that the vicilin protein was not in a random coil state after the transition at ∼85 °C. Further studies are required to characterize the thermally denatured vicilin protein in order to assess whether it consists of highly stable structural components.

In summary, we have purified a 7S vicilin from Korean pine. Sequence analysis suggested that this vicilin is related to the vicilin of the Loblolly pine. SDS-PAGE analyses revealed four bands resulting from post-translational processing of a single gene product. The thermal stability of the purified vicilin was accessed by CD and DSC. It was found that heating to 94 °C resulted in irreversible changes in the conformation of the protein. However, a significant portion of the protein is not in random coil after thermal denaturation. Urea denaturation study indicated that the purified pine nut vicilin was not unusually stable. Molecular cloning of Korean pine vicilin and structural studies of the protein in the future may advance the understanding of the physical and chemical properties of the cupin family of proteins. Such knowledge is of great interest since it may provide insight into the structural property and thermal stability of food allergens, many of which are members of the cupin family.

## LITERATURE CITED

(1) Pryer, K. M.; Schneider, H.; Smith, A. R.; Cranfill, R.; Wolf, P. G.; Hunt, J. S.; Sipes, S. D. Horsetails and ferns are a monophyletic group and the closest living relatives to seed plants. *Nature* **2001**, *409* (6820), 618–622.
(2) Soltis, P. S.; Soltis, D. E.; Savolainen, V.; Crane, P. R.; Barraclough, T. G. Rate heterogeneity among lineages of tracheophytes: integration of molecular and fossil data and evidence for molecular living fossils. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99* (7), 4430–4435.
(3) Conifer genome network. http://www.pinegenome.org/ (accessed August1, 2007).
(4) Straus, S. H.; Doerksen, A. H. Restriction fragment analysis of pine phylogeny. *Evolution* **1990**, *44* (4), 1081–1096.
(5) Liang, C.; Wang, G.; Liu, L.; Ji, G.; Fang, L.; Liu, Y.; Carter, K.; Webb, J. S.; Dean, J. F. ConiferEST: an integrated bioinformatics system for data reprocessing and mining of conifer expressed sequence tags (ESTs). *BMC Genomics* **2007**, *8*, 134.
(6) Baker, J. B.; Langdon, O. G. *Pinus taeda* L. Loblolly pine. In *Silvics of North America*, Vol. 1 Conifers; U.S.D.A. Forest Service Agriculture Handbook 654: Washington, DC, 1990; pp 1018−1051.
(7) Nergiz, C.; Dönmez, I. Chemical composition and nutritive value of *Pinus pinea* L. seeds. *Food Chem.* **2004**, *86* (3), 365–368.
(8) Venkatachalam, M.; Sathe, S. K. Chemical composition of selected edible nut seeds. *J. Agric. Food Chem.* **2006**, *54* (13), 4705–4714.
(9) USDA. USDA National Nutrient Database for Standard Reference. http://www.nal.usda.gov/fnic/foodcomp/search/ (accessed February 1, 2008.
(10) Sharashkin, L.; Gold, M. Pine nut: species, products, markets, and potential for U.S. production. In Northern Nut Growers Association 95th Annual Report, Proceedings for the 95th annual meeting, Columbia, MO, August 16−19, 2004.
(11) Johnson, A. F.; Baucom, C.; Brawner, A.; Golden, K.; Hubbard, K. M.; Peter, G.; Retzel, E. F.; Stambaugh, A.; Whetten, R. W.; Sederoff, R. R. The *Pinus taeda* L. (Loblolly Pine) EST Sequencing Project. In Plant, Animal & Microbe Genomes X Conference, San Diego, CA, January 12−16, 2002.
(12) Añó, M. A.; Maselli, J. P.; Sanz Mf, M. L.; Fernandez-Benitez, M. Allergy to pine nut. *Allergol. Immunopathol.* **2002**, *30* (2), 104–108.
(13) Garcia-Menaya, J. M.; Gonzalo-Garijo, M. A.; Moneo, I.; Fernandez, B.; Garcia-Gonzalez, F.; Moreno, F. A 17-kDa allergen detected in pine nuts. *Allergy* **2000**, *55* (3), 291–293.
(14) Fine, A. J. Hypersensitivity reaction to pine nuts (pinon nuts-pignolia). *Ann. Allergy* **1987**, *59* (3), 183–184.
(15) Roux, N.; Hogendijk, S.; Hauser, C. Severe anaphylaxis to pine nuts. *Allergy* **1998**, *53* (2), 213–214.
(16) Senna, G.; Roncarolo, D.; Dama, A.; Mistrello, G. Anaphylaxis to pine nuts and immunological cross-reactivity with pine pollen proteins. *J. Invest. Allergol. Clin. Immunol.* **2000**, *10* (1), 44–46.
(17) Ibanez, M. D.; Lombardero, M.; San Ireneo, M. M.; Munoz, M. C. Anaphylaxis induced by pine nuts in two young girls. *Pediatr. Allergy Immunol.* **2003**, *14* (4), 317–319.
(18) van de Scheur, M. R.; Bruynzeel, D. P. Acute anaphylaxis after pine nut skin testing. *Ann. Allergy, Asthma, Immunol.* **2004**, *92* (1), 93.
(19) Rubira, N.; Botey, J.; Eseverri, J. L.; Marin, A. Allergy to pine nuts in children. *Allerg. Immunol. (Paris)* **1998**, *30* (7), 212–216.
(20) Nielsen, N. H. Systemic allergic reaction after ingestion of pine nuts, *Pinus pinea*. *Ugeskr. Laeg.* **1990**, *152* (48), 3619–3620.
(21) Muntz, K.; Belozersky, M. A.; Dunaevsky, Y. E.; Schlereth, A.; Tiedemann, J. Stored proteinases and the initiation of storage protein mobilization in seeds during germination and seedling growth. *J. Exp. Bot.* **2001**, *52* (362), 1741–1752.
(22) Sanchez-Monge, R.; Lopez-Torrejon, G.; Pascual, C. Y.; Varela, J.; Martin-Esteban, M.; Salcedo, G. Vicilin and convicilin are potential major allergens from pea. *Clin. Exp. Allergy* **2004**, *34* (11), 1747–1753.
(23) Wang, F.; Robotham, J. M.; Teuber, S. S.; Tawde, P.; Sathe, S. K.; Roux, K. H. Ana o 1, a cashew (*Anacardium occidental*) allergen of the vicilin seed storage protein family. *J. Allergy Clin. Immunol.* **2002**, *110* (1), 160–166.
(24) Mills, E. N.; Jenkins, J.; Marigheto, N.; Belton, P. S.; Gunning, A. P.; Morris, V. J. Allergens of the cupin superfamily. *Biochem. Soc. Trans.* **2002**, *30* (Pt 6), 925–929.
(25) Shin, D. S.; Compadre, C. M.; Maleki, S. J.; Kopper, R. A.; Sampson, H.; Huang, S. K.; Burks, A. W.; Bannon, G. A. Biochemical and structural analysis of the IgE binding sites on ara h1, an abundant and highly allergenic peanut protein. *J. Biol. Chem.* **1998**, *273* (22), 13753–13759.

(26) Teuber, S. S.; Jarvis, K. C.; Dandekar, A. M.; Peterson, W. R.; Ansari, A. A. Identification and cloning of a complementary DNA encoding a vicilin-like proprotein, jug r 2, from english walnut kernel (*Juglans regia*), a major food allergen. *J. Allergy Clin. Immunol.* **1999**, *104* (6), 1311–1320.

(27) Viquez, O. M.; Konan, K. N.; Dodo, H. W. Structure and organization of the genomic clone of a major peanut allergen gene, Ara h 1. *Mol. Immunol.* **2003**, *40* (9), 565–571.

(28) Lauer, I.; Foetisch, K.; Kolarich, D.; Ballmer-Weber, B. K.; Conti, A.; Altmann, F.; Vieths, S.; Scheurer, S. Hazelnut (*Corylus avellana*) vicilin Cor a 11: molecular characterization of a glycoprotein and its allergenic activity. *Biochem. J.* **2004**, *383* (Pt 2), 327–334.

(29) Allona, I.; Casado, R.; Aragoncillo, C. Biochemical genetics of a 7S globulin-like protein from *Pinus pinaster* seed. *Theor. Appl. Genet.* **1994**, *88*, 454–459.

(30) Abramoff, M.; Magalhaes, P. S.ram, image processing with ImageJ. *Biophotonics Int.* **2004**, *11* (7), 36–42.

(31) Jin, T.; Fu, T.-J.; Kothary, M. H.; Howard, A.; Zhang, Y.-Z. Crystallization and initial crystallographic characterization of a vicilin-type seed storage protein from *Pinus koraiensis*. *Acta Crystallogr., Sect. F: Struct. Biol. Cryst. Commun.* **2007**, *63* (Pt 12), 1041–1043.

(32) Lopez-Torrejon, G.; Salcedo, G.; Martin-Esteban, M.; Diaz-Perales, A.; Pascual, C. Y.; Sanchez-Monge, R. Len c 1, a major allergen and vicilin from lentil seeds: protein isolation and cDNA cloning. *J. Allergy Clin. Immunol.* **2003**, *112* (6), 1208–1215.

(33) O'Kane, F. E.; Happe, R. P.; Vereijken, J. M.; Gruppen, H.; van Boekel, M. A. Characterization of pea vicilin. 1. Denoting convicilin as the α-subunit of the Pisum vicilin family. *J. Agric. Food Chem.* **2004**, *52* (10), 3141–3148.

(34) Mendoza, E. M.; Adachi, M.; Bernardo, A. E.; Utsumi, S. Mungbean [*Vigna radiata* (L.) Wilczek] globulins: purification and characterization. *J. Agric. Food Chem.* **2001**, *49* (3), 1552–1558.

(35) Gatehouse, J. A.; Lycett, G. W.; Croy, R. R.; Boulter, D. The post-translational proteolysis of the subunits of vicilin from pea (*Pisum sativum* L.). *Biochem. J.* **1982**, *207* (3), 629–632.

(36) The DFCI pinus gene index (PGI). http://compbio.dfci.harvard.edu/tgi/cgi-bin/tgi/tc_report.pl?tc=TC57029&species=Pine (accessed May 26, 2007).

(37) Maruyama, N.; Adachi, M.; Takahashi, K.; Yagasaki, K.; Kohno, M.; Takenaka, Y.; Okuda, E.; Nakagawa, S.; Mikami, B.; Utsumi, S. Crystal structures of recombinant and native soybean β-conglycinin β homotrimers. *Eur. J. Biochem.* **2001**, *268* (12), 3595–3604.

(38) Maruyama, Y.; Maruyama, N.; Mikami, B.; Utsumi, S. Structure of the core region of the soybean β-conglycinin α′ subunit. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004**, *60* (Pt 2), 289–297.

(39) Ko, T. P.; Day, J.; McPherson, A. The refined structure of canavalin from jack bean in two crystal forms at 2.1 and 2.0 A resolution. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2000**, *56* (Pt 4), 411–420.

(40) Ko, T. P.; Ng, J. D.; Greenwood, A.; McPherson, A. Determination of three crystal structures of canavalin by molecular replacement. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1993**, *49* (Pt 5), 478–489.

(41) Lawrence, M. C.; Izard, T.; Beuchat, M.; Blagrove, R. J.; Colman, P. M. Structure of phaseolin at 2.2 A resolution. Implications for a common vicilin/legumin structure and the genetic engineering of seed storage proteins. *J. Mol. Biol.* **1994**, *238* (5), 748–776.

(42) Lawrence, M. C.; Suzuki, E.; Varghese, J. N.; Davis, P. C.; Van Donkelaar, A.; Tulloch, P. A.; Colman, P. M. The three-dimensional structure of the seed storage protein phaseolin at 3 A resolution. *EMBO J.* **1990**, *9* (1), 9–15.

(43) Koppelman, S. J.; Knol, E. F.; Vlooswijk, R. A.; Wensing, M.; Knulst, A. C.; Hefle, S. L.; Gruppen, H.; Piersma, S. Peanut allergen Ara h 3: isolation from peanuts and biochemical characterization. *Allergy* **2003**, *58* (11), 1144–1151.

(44) Markossian, K. A.; Khanova, H. A.; Kleimenov, S. Y.; Levitsky, D. I.; Chebotareva, N. A.; Asryants, R. A.; Muronetz, V. I.; Saso, L.; Yudin, I. K.; Kurganov, B. I. Mechanism of thermal aggregation of rabbit muscle glyceraldehyde-3-phosphate dehydrogenase. *Biochemistry* **2006**, *45* (44), 13375–13384.

**Received for review April 9, 2008. Revised manuscript received June 22, 2008. Accepted June 22, 2008.**

JF801138Q