Structure–Function Relationships of Soybean Proglycinins at Subunit Levels

Krisna Prak,[†] Kazuyo Nakatani,[†] Tomoyuki Katsube-Tanaka,[‡]
Motoyasu Adachi,[†] Nobuyuki Maruyama,[†] and Shigeru Utsumi*[,†]

Laboratory of Food Quality Design and Development, Graduate School of Agriculture,
Kyoto University, Uji, Kyoto 611-0011, Japan, and Crop Science Laboratory, Faculty of Agriculture,
Tottori University, Tottori, Tottori 680-8553, Japan

Glycinin consists of five kinds of subunits, group I (A1aB1b, A1bB2, and A2B1a) and group II (A3B4 and A5A4B3). cDNAs for individual subunits were cloned by reverse transcription-polymerase chain reaction method and expressed in *Escherichia coli* using pET vector. The recombinant proglycinins were purified by ammonium sulfate fractionation and column chromatography in the form of homotrimers. Physicochemical properties such as molecular dimensions, solubility, surface hydrophobicity, thermal stability, and emulsifying ability of individual proglycinins were studied. Molecular dimensions were proportional to molecular size for all proglycinins except A2B1a. Solubility was intrinsic to each proglycinin. At the ionic strength of 0.5, all proglycinins except A1aB1b showed a very low solubility at acidic pH, but A1aB1b was soluble to higher than 60%. At ionic strength 0.08, all proglycinins exhibited isoelectric precipitation, although A2B1a and A1bB2 were not completely insoluble. The order of emulsifying ability (A1bB2 < A2B1a < A5A4B3 < A3B4 ≤ A1aB1b) was not of the same for surface hydrophobicity (A5A4B3 < A1aB1b ≤ A3B4 < A1bB2 < A2B1a) and thermal stability (A1bB2 ≪ A2B1a ≤ A5A4B3 < A3B4 ≤ A1aB1b).

KEYWORDS: Glycinin; proglycinin; structure–physicochemical function relationship; solubility; thermal stability; soybean

INTRODUCTION

Soybean (*Glycine max* L.) proteins have good nutritional and physicochemical properties among plant proteins. They are composed of two major components, glycinin (11S globulin) and β-conglycinin (7S globulin), accounting for 40 and 30% of the total seed proteins, respectively (*1*, *2*). β-Conglycinin is a trimeric protein composed of three subunits: α (~67 kDa), α' (~71 kDa), and β (~50 kDa). On the other hand, glycinin is a hexameric protein composed of five major subunits (A1aB1b, A1bB2, A2B1a, A3B4, and A5A4B3), each of which consists of an acidic (~30 kDa) and a basic (~20 kDa) polypeptide linked by a single disulfide bond, except A4 of A5A4B3 (*3*). The five subunits have been classified into two groups based on homology in their sequences. Group I consists of A1aB1b (~53.6 kDa), A1bB2 (~52.2 kDa), and A2B1a (~52.4 kDa), and group II consists of A3B4 (~55.4 kDa) and A5A4B3 (~61.2 kDa). The homology of each subunit is more than 84% within a group and 45−49% between groups (*2*, *4*).

In developing seeds, the constituent subunits of 11S globulin are synthesized as a single polypeptide precursor, preproprotein, the signal sequence of which is removed cotranslationally. The resultant proproteins assemble into trimers of ~8S in the endoplasmic reticulum. The proprotein trimers are transported from the endoplasmic reticulum to protein storage vacuoles, where they then are cleaved to form acidic and basic polypeptides that are linked by a disulfide bond (*5*). Finally, the mature proteins assemble into hexamers. The protein reserves are stored in the dormant seed until its germination.

The two major soybean storage proteins apparently play different roles in food and nonfood soy protein products due to their different physicochemical properties such as hydrophobicity, solubility, thermal stability, and emulsification. The use of soy protein products as functional ingredients is gaining increasing acceptance in food manufacturing from the standpoints of human nutrition and health. To further enhance the use of soybean proteins, more studies on the physicochemical properties of each subunit of glycinin and β-conglycinin are required. For this purpose, homogeneous molecular species composed of only one kind of subunit are necessary. However, it is very difficult to obtain a large amount of such homogeneous species from soybean seeds because of their molecular heterogeneity, the presence of many molecular species having different subunit compositions (*1*). In a previous study, our group employed two methods to obtain homogeneous molecular species for β-conglycinin: One is by using mutant soybean cultivars containing β-conglycinin lacking α or α' subunits (*6*,

* To whom correspondence should be addressed. Tel: 81-774-38-3760. Fax: 81-774-38-3761. E-mail: sutsumi@kais.kyoto-u.ac.jp.
† Kyoto University.
‡ Tottori University.

Soybean Proglycinins at Subunit Levels

J. Agric. Food Chem., Vol. 53, No. 9, 2005 **3651**

7), and the other is by constructing *Escherichia coli* expression systems for each subunit of β-conglycinin (*8*, *9*) having the correct conformation (*10*, *11*). We demonstrated the significant contribution of the carbohydrate moieties and core and extension regions to the physicochemical properties of β-conglycinin (*6*, *8*, *9*).

In analogy with our work on β-conglycinin, we studied the physicochemical properties of glycinin composed of only A3B4, only A5A4B3, only group I, and only group II by using mutant soybean cultivars (*12*). Unfortunately, we were not able to obtain individual molecular species of glycinin composed of only one kind of subunit. By employing the *E. coli* expression system, we can prepare such individual molecular species, although they are in the form of proglycinin. Our group has already succeeded in cloning cDNAs encoding A1aB1b (*13*), A2B1a (*14*), and A3B4 (*15*) and also has established the *E. coli* expression systems for A1aB1b (*16*) and A3B4 (*15*) subunit precursors of glycinin. In the present study, we cloned cDNAs for the remaining A1bB2 and A5A4B3 subunits by the reverse transcription-polymerase chain reaction (RT-PCR) method; constructed *E. coli* expression systems for A1bB2, A2B1a, and A5A4B3; and then compared the structural and physicochemical properties of all five subunits.

MATERIALS AND METHODS

Preparation of cDNAs Encoding Individual Glycinin Subunits and Construction of their Expression Plasmids. Total RNA was extracted from developing soybean seeds according to Shirzadegan et al. (*17*). For purification of mRNA from total RNA, the mRNA purification Kit (Amersham Pharmacia Biotech) was used according to the manufacturer's instructions. Full-length cDNAs encoding proglycinin were produced by reverse transcription of mRNA using RNA LA PCR Kit (AMV) Ver. 1.1 (TAKARA). The reaction was performed at 42 °C for 15 min, 60 °C for 15 min, and 99 °C for 5 min.

cDNAs encoding proglycinin subunits A5A4B3 were amplified by PCR using full-length cDNAs as a template for individual subunits. The primers used for the A5A4B3 subunit were 5′-ATT AGC TCC AGC AAG CTC AAC GAG TGC-3′ and 5′-C CGC GGA TCC **TTA** TGC GAC TTT AAC ACG GGG TGA GC-3′ for N and C termini, respectively (underlined and bold letters are *Bam*HI site and a stop codon, respectively). Thirty cycles of PCR were performed using LA TAQ (TAKARA) at 94 °C for 30 s, 60 °C for 30 s, and 72 °C for 4 min.

To avoid misannealing caused by the high homology of the nucleotides sequences of A1bB2 to A1aB1b, which is only one nucleotide difference in 42 nucleotides coding the starting 14 amino acids at the N-terminal of the mature forms, two pairs of primers were used. The first pair served for high selectivity in order to get the cDNA encoding A1bB2 lacking 29 base pairs for the N terminus of the mature form. The second pair of primers was used to get the complete cDNA containing the 29 base pairs based on the sequence reported by Cho et al. (*18*). The first pair of primers was 5′-CGA GTGCCA GAT CCA ACG C-3′ for the N-terminal and 5′-CTA CGC ACA CTA ACT AGT GC-3′ for the C-terminal. The reaction was the same to that of A5A4B3. Thirty cycles of PCR were performed using KOD-Plus (TOYOBO) at 94 °C for 30 s and 68 °C for 2 min and 30 s to amplify the cDNA encoding A1bB2 after LA TAQ (TAKARA). The second pair of primers was 5′-TTC AGT TTC AGA GAG CAG CCA CAG CAA AAC GAG TGC CAG ATC CAA CGC CTC AAT GC-3′ and 5′-G CAT GGA TCC AGG GCT **CTA** AGC CAC AAC TCT CCT CTG AGA CTC CTT-3′ for N and C termini, respectively (*Bam*HI site and a stop codon are indicated as underlined and bold letters, respectively).

The cDNA encoding the proglycinin subunit A2B1a was recovered from pGST4-3-1-4 (*14*) using KOD-plus and the primers 5′-CTG AGA GAG CAG GCA CAG CAA AAT-3′ and 5′-GAG TGC GGC CGC AAG CTT **CTA** AGC CAC-3′ for the N and C termini, respectively. Underlined and bold letters indicate *Hin*dIII site and a stop codon, respectively.

The cDNAs encoding the proglycinin subunits A1bB2, A2B1a, and A5A4B3 were blunted using the Blunting Kit (TAKARA) and then digested with *Bam*HI or *Hin*dIII. The resulting large DNA fragments were inserted into *Nco*I (filled-in) and *Bam*HI or *Hin*dIII sites of pET-21d vector (Novagen) to construct the expression plasmids pEA1bB2, pEA2B1a, and pEA5A4B3.

The nucleotide sequences of the cDNAs in the expression plasmids were confirmed by the dideoxy chain termination method of Sanger et al. (*19*), using the Applied Biosystems sequencer model 310 and the ABI dye terminator cycle sequencing kit with *Amplitaq* polymerase FS (Perkin-Elmer, Applied Biosystems).

Expression of Recombinant Proglycinins. The host cells BL21-(DE3), HMS174(DE3), AD494(DE3), and Origami(DE3) containing individual expression plasmids pEA1aB1b (*16*), pEA1bB2, pEA2B1a, pEA3B4 (*15*), and pEA5A4B3 were cultured in normal LB and TB media and in media containing additional NaCl. The cells were grown at 37 °C. Expression was induced when A_{600} reached 0.4 to ∼0.6 with 1 mM isopropyl-1-thio-β-D-galactoside (IPTG), at 18, 20, and 25 °C. After cultivation, cells were harvested by centrifugation at 5000 rpm for 20 min at 4 °C and stored at −20 °C until used. Aliquots of the cells were analyzed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS−PAGE) (*20*) using 11% gel as described previously (*21*). Expressed recombinant proglycinins were identified based on their expected sizes and confirmed by western blotting (*22*) using anti-glycinin antibody followed by goat anti-rabbit IgG-alkaline phosphatase conjugate (Promega).

Purification of Recombinant Proglycinins. All purification steps were carried out at 4 °C and centrifugations were at 6000 rpm for 20 min unless otherwise stated. The basic buffer for all purification steps was buffer A [35 mM potassium phosphate, pH 7.6, 1 mM EDTA, 10 mM 2-mercaptoethanol, 0.1 mM (*p*-amidinophenyl)-methylsulfonyl fluoride, 1 μg/mL pepstatin A, and 1 μg/mL leupeptin]. Ammonium sulfate fractionation followed the procedure of Green et al. (*23*).

Proglycinins A1aB1b, A2B1a, and A3B4. Frozen cells containing proglycinins A1aB1b, A2B1a, and A3B4 were resuspended in buffer B (buffer A containing 0.4 M NaCl) at a density of 25 g/L buffer and lysed by sonication on an ice bath. Insoluble matter was removed by centrifugation. Expressed proteins were initially fractionated using ammonium sulfate: 35% saturation for A1aB1b and A3B4 and 15% saturation for A2B1a. The precipitate was removed by centrifugation, and the soluble fraction containing recombinant proteins was applied on a Toyopearl (Butyl-650M) (TOSOH) column (2.6 cm × 20 cm) equilibrated with buffer B containing 30% ammonium sulfate for A1aB1b and A3B4 and 15% saturation for A2B1a. Elution was carried out with a linear gradient (800 mL) from 30 to 0% ammonium sulfate in buffer B for A1aB1b and A3B4 and from 15 to 0% for A2B1a. Fractionation of A2B1a was further continued with 100 mL of buffer B. Fractions containing proglycinins A1aB1b, A2B1a, and A3B4 were dialyzed against buffer C (buffer A containing 0.15 M NaCl) and clarified by centrifugation. A1aB1b was applied on Mono Q HR 10/10 column (Pharmacia Biotech) equilibrated with buffer C. Elution was performed with a linear gradient from 0.15 to 0.4 M NaCl in buffer A over a period of 120 min at 2 mL/min. A2B1a and A3B4 were applied on HiLoad 26/10 Q-Sepharose high performance column (Pharmacia Biotech) equilibrated with buffer C. Elution was performed with a linear gradient from 0.15 to 0.5 M NaCl in buffer A over a period of 120 min at 2 mL/min.

Proglycinin A1bB2 and A5A4B3. Frozen cells for proglycinin A1bB2 and A5A4B3 were resuspended in buffer D (buffer A containing 1.0 M NaCl) at a density of 25 g/L buffer and disrupted by sonication on an ice bath. The insoluble matter was removed by centrifugation. Expressed proteins were initially fractionated by 37.5% ammonium sulfate saturation. After removal of precipitates, soluble fractions were further fractionated by ammonium sulfate saturation to 50 and 60% for A1bB2 and A5A4B3, respectively. The precipitates containing A1bB2 and A5A4B3 were then dissolved in buffer D containing 30% ammonium sulfate and then subjected to Toyopearl (Butyl-650M) column (2.6 cm × 20 cm) following the same procedure for A1aB1b and A3B4. Fractions containing A1bB2 and A5A4B3 were dialyzed against buffer D, concentrated by Centriprep YM-10 (Millipore), and then were applied on a gel filtration column (Hi-Prep 26/60 Sephacryl

3652 *J. Agric. Food Chem.,* Vol. 53, No. 9, 2005

Prak et al.

S-300 HR) using buffer D as a mobile phase. Furthermore, A1bB2 and A5A4B3 were subjected to HiLoad 26/10 Q-Sepharose high performance chromatography (Pharmacia Biotech). Elution was performed similar to A2B1a and A3B4 except for the NaCl concentration of the gradient buffers: 0.15−0.45 M NaCl and 0.25−0.5 M NaCl for A1bB2 and A5A4B3, respectively.

Protein Measurement. The amount of protein in the samples was determined using a Protein Assay Rapid Kit (Wako) with bovine serum albumin as a standard.

Analysis of Self-Assembly into Trimers. Self-assembly of individual proglycinins was analyzed using Hi-Prep 16/60 Sephacryl S-200 HR column as described previously (*8*).

Solubility as a Function of pH. Proglycinins were dialyzed against buffer E [10 mM sodium phosphate, pH 7.6, 0.5 M NaCl, 1 mM EDTA, 0.1 mM (*p*-amidinophenyl)-methylsulfonyl fluoride, 1 μg/mL pepstatin A, 1 μg/mL leupeptin, 0.02% NaN₃, and 10 mM 2-mercaptoethanol]. Ionic strength and pH were adjusted by adding the appropriate amount of 10 and 50 mM of citrate buffer (pH 2.5−6.5), sodium phosphate buffer (pH 5.5−7.8), and ammonium buffer (pH 7.2−9.0) for the ionic strengths 0.08 and 0.5, respectively. The protein samples (0.8 mg/mL) were kept at 4 °C for 18 h and then centrifuged at 15000 rpm at 4 °C for 15 min to separate the soluble and insoluble fractions. The percent solubility (soluble fraction) was determined by comparing the protein content of the resulting solution with the initial protein content of the sample (100% soluble).

Surface Hydrophobicity. Surface hydrophobicities of proglycinins were analyzed using butyl sepharose 4FF and phenyl sepharose 6FF as described previously (*7*) except for the starting concentration of ammonium sulfate [35% saturation instead of 2.3 M (48.75%)]. The proteins were eluted with a linear gradient of ammonium sulfate 35−0% over a period of 100 min at a flow rate of 0.25 mL/min.

Differential Scanning Calorimetry (DSC) Measurement. DSC measurement of proglycinins was carried out as described previously (*8*) using 1 mg/mL of protein. The scan rate was 1 °C/min from 30 to 110 °C under nitrogen gas.

Emulsifying Property. The emulsifying properties of proglycinins were analyzed as described previously (*9*) using 1.5 mL of 0.5 mg/mL of protein at pH 7.6 and ionic strengths 0.5 and 0.08, mixing with 0.25 mL of soybean oil.

RESULTS AND DISCUSSION

Cloning and Sequencing of cDNAs Encoding A1bB2 and A5A4B3 Proglycinins. Because of the high similarity to other group I subunits (*4, 24*) and lower detectable levels of mRNA for proglycinin A1bB2 than those for the other group I subunits in the seed (*18*), it was difficult to get a cDNA encoding proglycinin A1bB2 by the RT-PCR method using a primer corresponding to its mature N-terminal region. To overcome this problem, two pairs of primers were used (see Materials and Methods). The first pair of primers was of high selectivity to produce a portion of the A1bB2 cDNA lacking 29 base pairs from the first codon for the mature form. A second pair of primers was employed to get the complete cDNA by using the incomplete cDNA as a template. The cDNA for proglycinin A5A4B3 subunit was cloned without any problem.

Nucleotide sequences of the cloned cDNAs were determined and were compared with those in the GenBank protein translation database. Codes used for A1bB2 and A5A4B3 are GMGY3 and AB004062, respectively. All coding sequences determined here were identical to those in the reference database except for one base difference in A1bB2 cDNA; the base A at position 460 from the initiation codon of GMGY3 gene was replaced by C, but there was no change in the amino acid.

Expression and Purification of Individual Proglycinins. To achieve an expression of soluble recombinant proglycinins at a high level, we attempted cultivation of various host cells having individual expression plasmids under various conditions.

Table 1. Host Cell and Culture Condition for Proglycinin

| proglycinin | host cell (Novagen) | culture conditions for protein expression | | | |
		medium	salt	temp (°C)	duration (h)
A1aB1b	AD494 (DE3)	LB	0.17 M NaCl	25	40
A2B1a	AD494 (DE3)	LB	0.17 M NaCl	25	40
A1bB2	AD494 (DE3)	LB	1.0 M NaCl	25	40
A3B4	BL21 (DE3)	TB	no	18	40
A5A4B3	Origami (DE3)	TB	0.1 M phosphate	20	40
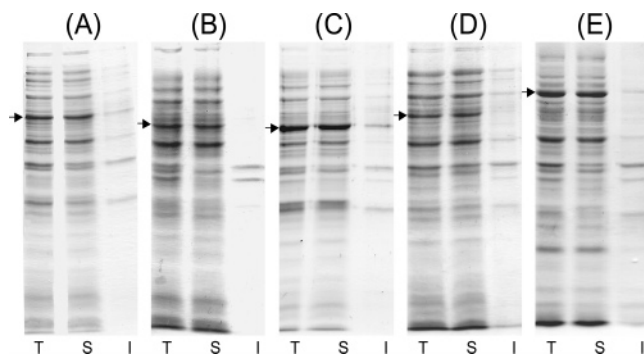
Figure 1. SDS−PAGE (11% acrylamide) profile of the protein expression of recombinant proglycinin subunits. The arrow indicates the position of the expressed proglycinin subunit. T, total protein after cell extraction; S, soluble fraction; and I, insoluble fraction. (**A**) A1aB1b, (**B**) A1bB2, (**C**) A2B1a, (**D**) A3B4, and (**E**) A5A4B3.

In the case of A1bB2, no recombinant protein was found using normal LB medium. However, increasing the salt concentration from 0.17 to 0.5, 0.8, and 1 M resulted in a remarkable expression of soluble recombinant protein (data not shown) with 70, 80, and >90% solubilities, respectively. On the other hand, NaCl concentration did not affect the A5A4B3 expression level but affected its solubility when using *E. coli* AD494(DE3). Origami(DE3) and BL21(DE3) were the suitable hosts for the expression of A5A4B3 and A3B4, respectively. The most suitable condition for individual proglycinins is summarized in **Table 1**. In optimizing expression conditions, we found that it was easier to get soluble expression of group I proglycinins than group II ones, suggesting that the folding ability of group I proglycinin is higher than that of group II.

The solubility of recombinant proteins from individual expression plasmids in *E. coli* cells was analyzed by SDS−PAGE after sonication (**Figure 1**). On the basis of band intensity, the solubility of all recombinant proglycinins was estimated to be >80%. These were confirmed by Western blotting (data not shown) using purified antiglycinin antibody.

After purification, the purity of the proteins was assessed by SDS−PAGE (**Figure 2**) and found to be >95%. The number of amino acids and the molecular masses of individual proglycinin subunits calculated based on their cDNA nucleotide sequences are shown in **Figure 2.** We found that the mobility of A2B1a was a bit faster than that of A1bB2 although its molecular mass was slightly bigger than A1bB2. Therefore, it must be the mobility of A1bB2 or A2B1a that was inconsistent with its molecular mass. To further analyze the phenomenon of mobility, we drew a curve Log (molecular mass) of proglycinin subunits vs migration of subunits on the gel (data not shown). We found that the correlation coefficient of the curve was higher when using A1bB2 with the other three subunits. This means that the mobility of proglycinin subunits

Soybean Proglycinins at Subunit Levels

J. Agric. Food Chem., Vol. 53, No. 9, 2005 **3653**
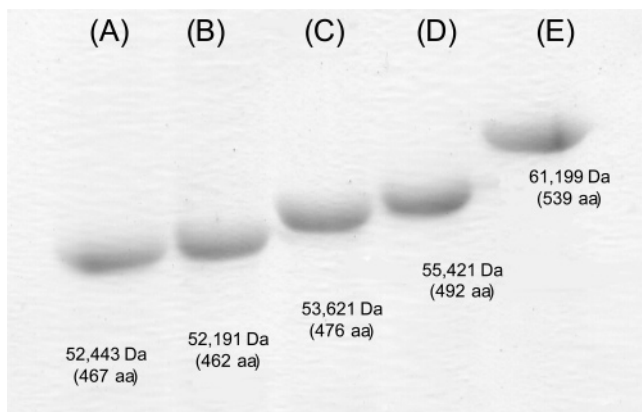
Figure 2. SDS–PAGE analysis of the purified proglycinins. The purified proglycinins were analyzed by SDS–PAGE using 11% acrylamide gel and running at 150 V for 55 min. The number with Da indicates the molecular mass of proglycinin. The number with aa indicates the number of amino acids. (**A**) A2B1a, (**B**) A1bB2, (**C**) A1aB1b, (**D**) A3B4, and (**E**) A5A4B3.

was consistent with the number of their amino acids and their molecular masses except those of A2B1a.

Self-Assembly of Proglycinin Subunits. To assess self-assembly into trimers of individual proglycinin subunits, each purified proglycinin sample was subjected to gel filtration chromatography using Hi-Prep 16/60 Sephacryl S-200 HR column at pH 7.6 and $I = 0.5$. The elution times were 87.3, 93.2, 95.0, 98.2, and 100.7 min for A5A4B3, A3B4, A1aB1b, A1bB2, and A2B1a, respectively (**Figure 3**). Previously, we demonstrated that A1aB1b self-assembles into trimers (*21, 24, 25*) and elutes at 95 min on the same column under the same conditions (*8*). Elutions of A5A4B3 and A3B4 were faster than that of A1aB1b, and those of A1bB2 and A2B1a were slower. These differences were proportional to their differences in mobilities on SDS−PAGE (**Figure 2**). Therefore, all progly-cinins were assessed to be trimers. Our group has proposed three criteria for judging the formation of proper conformation of proglycinin: (i) a high level of expression in *E. coli* (>10% of *E. coli* total proteins), (ii) solubilities comparable with that of globulins, and (iii) self-assembly into trimers (*26, 27*). Because all of the proglycinins constructed here satisfied these criteria, we concluded that they assumed a conformation similar to that of native proglycinins. Therefore, our *E. coli* expression system could be used for the production of individual proglycinins. The

recombinant proteins can then be used to study structural and physicochemical properties of individual proglycinins.

The elution time of each proglycinin did not follow the order of their real molecular masses but followed that of their mobilities on SDS−PAGE. Because A2B1a has five amino acid residues more than A1bB2 (**Figure 2**), it should move slightly slower than A1bB2 on SDS−PAGE, but on the contrary, it moved slightly faster. These indicate that A2B1a has the most compact structure among all proglycinin trimers and monomers. Previously, our group demonstrated that even one amino acid replacement of glycinin A3 polypeptide results in slower mobility on SDS−PAGE than the normal one (*28*). This phenomenon is similar to the unusual behavior of A2B1a on SDS−PAGE. The unusual behavior of A2B1a on gel filtration is a first report for major seed storage proteins. In soybean seeds, many molecular species of different subunit compositions are present (*1*). This indicates that individual proglycinins can form heterotrimers. Because they share the same backbone structure, the difference will therefore be on the molecular surface. To understand the difference in the molecular surface of A2B1a subunit, crystallography of A2B1a homotrimer is desired and is in progress by our group.

Solubility as a Function of pH. Solubility is a fundamental property for physicochemical functions of proteins for food usage. We examined the solubility of proglycinin subunits at different pH values at low ($I = 0.08$) and high ($I = 0.5$) ionic strengths (**Figure 4**).

At low ionic strength, A1bB2 and A2B1a exhibited similar solubilities. The lowest solubility was around 20−30% at the pH values 4.5−5.9 and 5.0−5.9 for A2B1a and A1bB2, respectively. Increasing the pH enhanced the solubility of these two subunits but only up to about 60−80%. A1aB1b at the same ionic strength had the narrowest range of insolubility among all proglycinins. In the range of pH 5.7−6.7, its solubility was less than 20%. Group II proglycinins had a wider range of insolubility at $I = 0.08$ than group I ones. The solubilities of A3B4 and A5A4B3 were less than 10% at pH 5.0−6.8 and 6−10% at pH 4.2−6.2, respectively. These phenomena were probably due to the fact that A3B4 and A5A4B3 subunits are richer in acidic and basic amino acids than the others.

At high ionic strength, all proglycinins were 100% soluble at pH $\geq$ 5.8. At lower pH, all proglycinins exhibited low solubility except for A1aB1b, which still had good solubility that never went below 60%. A2B1a, A3B4, and A5A4B3 decreased in solubility to nearly 0, 0−25, and 18−36%, respectively. A1bB2 had very low solubility only at pH $\leq$ 3.2.
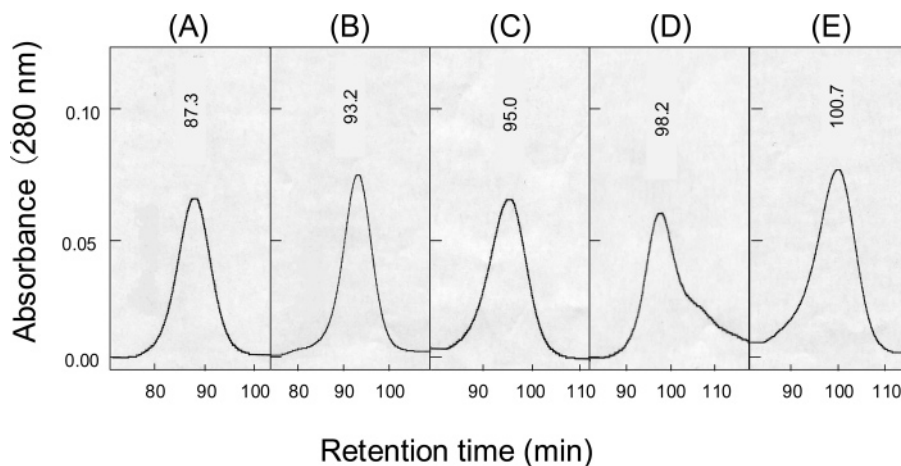
Figure 3. Gel filtration profile of proglycinins. The purified recombinant proglycinins were subjected to chromatography using a Sephacryl S-200 HR column. (**A**) A5A4B3, (**B**) A3B4, (**C**) A1aB1b, (**D**) A1bB2, and (**E**) A2B1a.

3654 *J. Agric. Food Chem.,* Vol. 53, No. 9, 2005

Prak et al.
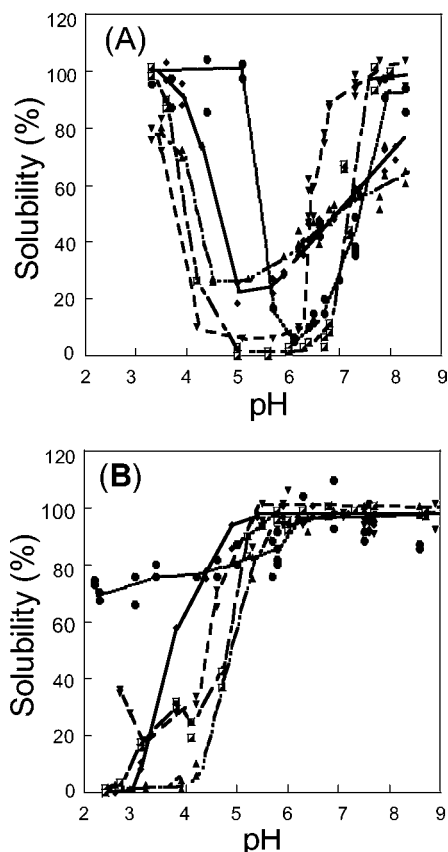
Figure 4. pH dependence of the solubility of proglycinins at ionic strengths of 0.08 (**A**) and 0.5 (**B**). A1aB1b, A2B1a, A1bB2, A3B4, and A5A4B3 are shown by a dashed line with closed cycles, double-dashed and double-dotted lines with closed triangles, a solid line with closed diamonds, single-dashed and single-dotted lines with half-open squares, and a dashed line with inverted triangles, respectively.

The pH dependence of the solubility of proglycinins was intrinsic to each subunit. The numbers of amino acid residues, especially charged residues, in the variable regions are different among the five subunits. This difference may have a big effect on the intrinsic solubilities of individual proglycinins. All proglycinins were completely soluble at pH ≥ 6.0 at high ionic strength. However, at low ionic strength even at pH > 7.6, the solubilities of A1bB2 and A2B1a were less than 70%. We cannot explain these phenomena at this point, but they can be studied using mutational approaches and three-dimensional structures.

Our group isolated glycinin molecular species with restricted subunit compositions: glycinin composed of only group I subunits (group I-glycinin), only group II subunits (group II-glycinin), only A3B4 (A3B4-glycinin), and only A5A4B3 (A5A4B3-glycinin) and studied their solubility at $I = 0.08$ (*29*) and 0.5 (*12*). At $I = 0.08$, the solubility profiles of A3B4- and A5A4B3-glycinins were very similar to those of A3B4 and A5A4B3, and that of group I-glycinin was just like a mixture of those of A1aB1b, A1bB2, and A2B1a. On the other hand, at $I = 0.5$, the solubility profiles of A3B4- and A5A4B3-glycinins were similar to those of A3B4 and A5A4B3 although there were some differences between pH 4 and pH 5 and that of group I-glycinin was just like a mixture of those of A1aB1b, A1bB2, and A2B1a. These suggest that proglycinin and mature glycinin exhibit similar solubility profiles as a function of pH. These results are consistent with the suggestions based on the solubility of various glycinin molecular species that the solubilities of

Table 2. Elution Time of Proglycinin on Hydrophobic Column

hydrophobic column	elution time (min) of proglycinin subunit				
	A1aB1b	A1bB2	A2B1a[a]	A3B4	A5A4B3
butyl sepharose 4FF	48.7	54.3		50.7	43.7
phenyl sepharose 6FF	67.0	74.4		71.1	65.5

[a] A2B1a was not subjected to the columns as it completely precipitated at 30% saturation of ammonium sulfate.

newly synthesized proteins in the endoplasmic reticulum determine the transport pathways of soybean storage protein (*29*).

Surface Hydrophobicity. Surface hydrophobicity is an important factor that contributes to physicochemical functions such as solubility, emulsifying, and foaming abilities (*30, 31*). At 35% saturation of ammonium sulfate, A2B1a completely precipitated whereas the other proglycinins were completely soluble, indicating that A2B1a has the highest surface hydrophobicity among the five proglycinins. To compare the surface hydrophobicity of the other four proglycinins, they were subjected to hydrophobic columns, butyl and phenyl sepharoses. The higher the surface hydrophobicity, the stronger the surface interaction with the column resulting into slower elution. The order of elution of proglycinin was identical in both columns (**Table 2**). The results indicate that the order of hydrophobicity is as follows: A1bB2 > A3B4 ≥ A1aB1b > A5A4B3. The hydrophobicities of A3B4 and A1aB1b were close to each other but reversed in butyl and phenyl sepharoses. The surface hydrophobicity was measured at pH 7.6. At pH 7.6, all proglycinins were soluble at $I = 0.5$, and the order of solubility at $I = 0.08$ was A5A4B3 > A3B4 > A1aB1b ≥ A1bB2 ≥ A2B1a. Therefore, there is no relationship between surface hydrophobicity and solubility of proglycinins.

The surface hydrophobicities of various glycinin molecular species were assessed by the same hydrophobic columns, and the order of surface hydrophobicities was identical in both columns. The order was A3B4-glycinin > group II-glycinin > A5A4B3-glycinin > group I-glycinin (*12*). This order was completely different from that of proglycinin, indicating that hexamer formation results in drastic change of surface hydrophobicities of glycinin subunits.

DSC Measurement. DSC has been used extensively to study thermodynamic properties of protein denaturation. We performed an experiment on Microcal MC-2 ultrasensitive microcalorimeter (**Figure 5**).

The thermal denaturation midpoint temperatures (T_m) and the denaturation starting temperature of proglycinins ranged from 65 to 78 °C and 52 to 70 °C, respectively. A1aB1b and A3B4 exhibited the highest T_m values while A1bB2 had the lowest. The order of T_m values was A1aB1b ≥ A3B4 > A5A4B3 ≥ A2B1a > A1bB2. It was noted that the shapes of exothermic peaks of A1aB1b and A1bB2 were broad, and those of A2B1a, A3B4, and A5A4B3 were sharp, indicating that the denaturation of A1aB1b and A1bB2 started at much lower temperatures than their T_m values. This resulted in the order of denaturation starting temperature as follows: A3B4 > A1aB1b ≥ A2B1a ≥ A5A4B3 > A1bB2. The T_m values of A1aB1b and A3B4 were 78.1 and 78.0 °C, respectively, being very close to each other, but their denaturation starting temperatures were 66.5 and 69.7 °C, respectively. This indicates that A3B4 is the most stable proglycinin trimer among the five proglycinins. A2B1a and A5A4B3 exhibited very close T_m values (73.3 °C for A2B1a and 73.9 °C for A5A4B3) and denaturation starting temperatures (65.5 °C for A2B1a and 65.0 °C for A5A4B3). Both the T_m

Soybean Proglycinins at Subunit Levels

J. Agric. Food Chem., Vol. 53, No. 9, 2005   **3655**
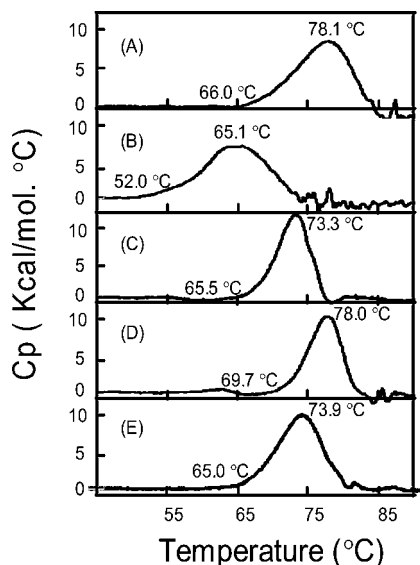
Figure 5. DSC scans of proglycinins. (**A**) A1aB1b, (**B**) A1bB2, (**C**) A2B1a, (**D**) A3B4, and (**E**) A5A4B3.

Figure 6. Particle size distributions of emulsion of proglycinins at ionic strengths of 0.08 (I) and 0.5 (II). (**A**) A1aB1b, (**B**) A1bB2, (**C**) A2B1a, (**D**) A3B4, and (**E**) A5A4B3.

value and the denaturation starting temperature of A1bB2 were the lowest among the five proglycinins. Therefore, the order of the thermal stability of the five proglycinins is A3B4 > A1aB1b > A5A4B3 ≥ A2B1a > A1bB2. In the case of mature glycinins, the thermal stability of A3B4 glycinin was the lowest, 87.2 °C (*12*). These show that the processing of proglycinin to mature form and the resultant hexamer formation result in stabilization of the molecules, but the degree of stabilization is variable among subunits.

Many factors were found to cause the differences in thermal stability. Our group has pointed out five factors to be responsible for the difference in thermal stability between soybean β-conglycinin, β subunit (90.8 °C), and the core region of α′ subunit (82.7 °C) (*11*). Thus, (i) the total cavity volume is larger in α′c; (ii) the cluster of charged residues at the intermonomer interface is smaller in α′c, and α′c lacks intermonomer salt bridge; (iii) the solvent accessible surface is more hydrophobic in α′c; (iv) there are fewer proline residues in α′c; and (v) a loop region between helix 3 and strand J′ in α′ is longer. The combined contributions of these five factors should provide more hydrogen bonds in α′c. To understand the structural reasons why individual proglycinins exhibit different thermal stability, X-ray crystallography of the proglycinins is required.

After heating to 110 °C and cooling to room temperature for the DSC experiment, precipitation of each proglycinin was observed. A3B4 and A1aB1b were found to be the least prone to precipitate, followed by A5A4B3, A2B1a, and A1bB2 subunits; A3B4 ≤ A1aB1b < A5A4B3 < A2B1a < A1bB2. Interestingly, A2B1a was very turbid, but no bulk precipitation was found. The degree of precipitation of proglycinins corresponds to the order of their thermal stability.

Emulsification. The emulsifying ability of proglycinins was analyzed by measuring particle size distributions after homogenization and sonication of samples with soybean oil at pH 7.6 and two ionic strengths ($I = 0.08$ and $I = 0.5$) (**Figure 6**). Proteins with a good emulsifying ability exhibit a small average particle size. Results obtained here indicate that all subunits exhibited a better emulsifying ability at $I = 0.5$ than at $I = 0.08$. The emulsifying ability of proglycinins can be ranked as follows: A1aB1b (3.9 and 10.5 μm) ≥ A3B4 (4.0 and 11.7 μm) > A5A4B3 (6.7 and 18.8 μm) > A2B1a (9.9 and 19.4
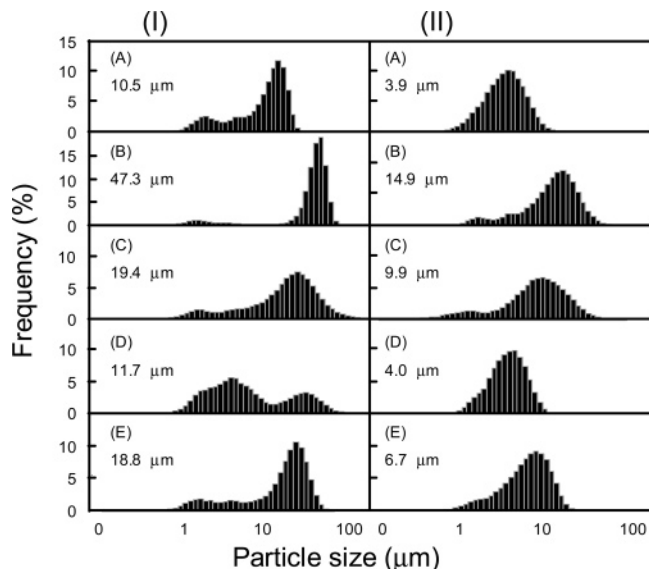
μm) > A1bB2 (14.9 and 47.3 μm). Values in parentheses indicate the average particle sizes at high and low ionic strengths. Surface hydrophobicity was reported to be an important factor for physicochemical function such as the emulsifying ability, because it is related to the binding of a protein to oil (*30*, *31*). However, the order of emulsifying abilities of proglycinins did not follow their order of surface hydrophobicities. A1bB2 had higher surface hydrophobicity among the five subunits but exhibited the poorest emulsifying ability. A2B1a had the highest surface hydrophobicity, but its emulsifying ability was only slightly better than that of A1bB2. In addition, suitable conformational change at the interface between oil and water, structural flexibility, is also one important factor (*32*). Thermal stability is related to structural flexibility, but the order of the emulsifying abilities of proglycinins did not follow the order of their thermal stabilities.

Generally, a balance of hydrophobicity and hydrophilicity of a protein is a factor for its emulsifying ability (*32*). Each glycinin subunit has 4−6 variable regions (*25*, *33*, *34*) that are rich in acidic residues. The longest one of the variable regions is located at the C terminus of the acidic polypeptide and is called a hypervariable region that consists of 42, 29, 35, 70, and 103 residues corresponding to 8.8, 6.3, 7.5, 14.2, and 19.1% of the total amino acid residues (*2*) in A1aB1b, A1bB2, A2B1a, A3B4, and A5A4B3, respectively. Similarly, the N-terminal extension region of β-conglycinin α and α′ consists of 125 and 141 residues corresponding to 23 and 24.4% of the total amino acid residues of α and α′, respectively. In previous studies, we demonstrated that the hypervariable region of mature glycinin (*12*) and the extension region of β-conglycinin (*9*) played a critical role in emulsifying ability. However, the order of contribution of variable region size as compared to the molecular size of proglycinin was A5A4B3 > A3B4 > A1aB1b > A2B1a > A1bB2, being not all consistent with the order of emulsifying ability of proglycinins. The effect on the emulsifying ability of the hypervariable region of proglycinins is different from that of mature glycinins and the extension region of β-conglycinin subunits. This might be due to the fact that the hypervariable region of proglycinin is located at the inside of the molecule, whereas the extension region and the hypervariable region are

3656 *J. Agric. Food Chem.,* Vol. 53, No. 9, 2005

Prak et al.

located at the N terminus of β-conglycinin and C terminus of acidic polypeptide of mature glycinin, respectively. Therefore, these regions can be more flexible than the hypervariable region in the proglycinins.

Thus, the emulsifying ability of proglycinin results from a combination of many factors such as surface hydrophobicity, solubility, flexible region (variable region), and structural stability that contribute to the balance of hydrophobicity and hydrophilicity and its suitable conformational change at the interface between oil and water.

Proglycinins have intrinsic solubility and have different physicochemical properties from each other. These characteristics are determined by many factors one of which might be the variable region. Recently, we succeeded in the X-ray crystallography of proglycinin A1aB1b (*25*) and mature glycinin A3B4 (*34*). X-ray crystallography of the other subunits will further elucidate the molecular basis for these properties.

LITERATURE CITED

(1) Utsumi, S. Plant food protein engineering. In *Advances in Food and Nutrition Research*; Kinsella, J. E., Ed.; Academic Press: SanDiego, CA, 1992; pp 89−208.

(2) Utsumi, S.; Matsumura, Y.; Mori, T. Structure−function relationships of soy proteins. In *Food Proteins and Their Applications*; Damodaran, S., Paraf, A., Eds.; Marcel Dekker: New York, 1997; pp 257−291.

(3) Staswick, P. E.; Hermodson, M. A.; Nielsen, N. C. The amino acid sequence of the A2B1a subunit of glycinin. *J. Biol. Chem.* **1984**, *259*, 13424−13430.

(4) Nielsen, N. C.; Dickinson, C. D.; Cho, T. J.; Thanh, V. H.; Scallon, B. J.; Fischer, R. L.; Sims, T. L.; Drews, G. N.; Goldberg, R. B. Characterization of the glycinin gene family in soybean. *Plant Cell* **1989**, *1*, 313−328.

(5) Dickinson, C. D.; Hussein, E. H. A.; Nielsen, N. C. Role of posttranslational cleavage in glycinin assembly. *Plant Cell* **1989**, *1*, 459−469.

(6) Maruyama, N.; Mohamad Ramlan, M. S.; Takahashi, K.; Yagasaki, K.; Goto, H.; Hontani, N.; Nakagawa, S.; Utsumi, S. The effect of the N-linked glycans on structural features and physicochemical functions of soybean β-conglycinin homotrimers. *J. Am. Oil Chem. Soc.* **2002**, *79*, 139−144.

(7) Maruyama, N.; Mohamed Salleh, M. R.; Tkahashi, K.; Yagasaki, K.; Goto, H.; Hontani, N.; Nakagawa, S.; Utsumi, S. Structure-physicochemical function relationships of soybean β-conglycinin heterotrimers. *J. Agric. Food Chem.* **2002**, *50*, 4323−4326.

(8) Maruyama, N.; Katsube, T.; Wada, Y.; Oh, M. H.; Barba de la Rosa, A. P.; Okuda, E.; Nakagawa, S.; Utsumi, S. The roles of the N-linked glycans and extension regions of soybean β-conglycinin in folding, assembly and structural features. *Eur. J. Biochem.* **1998**, *258*, 854−862.

(9) Maruyama, N.; Sato, R.; Wada, Y.; Matsumura, Y.; Goto, H.; Okuda, E.; Nakagawa, S.; Utsumi, S. Structure-physicochemical function relationships of soybean β-conglycinin constituent subunits. *J. Agric. Food Chem.* **1999**, *47*, 5279−5284.

(10) 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*, 3595−3604.

(11) Maruyama, Y.; Maruyama, N.; Mikami, B.; Utsumi, S. Structure of the core region of the soybean β-conglycinin α' subunit. *Acta Crystallogr.* **2004**, *D60*, 289−297.

(12) Maruyama, N.; Prak, K.; Motoyama, S.; Choi, S.-K.; Yagasaki, K.; Ishimoto, M.; Utsumi, S. Structure-physicochemical function relationships of soybean glycinin at subunit level assessed by using mutant lines. *J. Agric. Food Chem.* **2004**, *52*, 8197−8201.

(13) Utsumi, S.; Kohno, M.; Mori, T.; Kito, M. An alternate cDNA encoding glycinin A1aBx subunit. *J. Agric. Food Chem.* **1987**, *35*, 210−214.

(14) Utsumi, S.; Kim, C.-S.; Kohno, M.; Kito, M. Polymorphism and expression of cDNAs encoding glycinin subunits. *Agric. Biol. Chem.* **1987**, *51*, 3267−3273.

(15) Choi, S.-K.; Adachi, M.; Utsumi, S. Identification of the bile acid-binding region in the soy glycinin A1aB1b subunit. *Biosci., Biotechnol., Biochem.* **2002**, *66*, 2395−2401.

(16) Katsube, T.; Kurisaka, N.; Ogawa, M.; Maruyama, N.; Ohtsuka, R.; Utsumi, S.; Takaiwa, F. Accumulation of soybean glycinin and its assembly with the glutelins in rice. *Plant Physiol.* **1999**, *120*, 1063−1073.

(17) Shirzadegan, M.; Christie, P.; Seemann, J. R. An efficient method for isolation of RNA from tissue cultured plant cells. *Nucleic Acids Res.* **1991**, *60*, 6055.

(18) Cho, T.-J.; Davies, C. S.; Fisher, R. L.; Turner, N. E.; Goldberg, R. B.; Nielsen, N. C. Molecular characterization of an aberrant allele for the *Gy3* glycinin gene: A chromosomal rearrangement. *Plant Cell* **1989**, *1*, 339−350.

(19) Sanger, F.; Kicklen, S.; Coulson, A. R. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. U.S.A.* **1977**, *74*, 5463−5467.

(20) Laemmli, U. K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **1970**, *227*, 680−685.

(21) Kim, C.-S.; Kamiya, S.; Kanamori, J.; Utsumi, S.; Kito, M. High-level expression, purification and functional properties of soybean proglycinin from *Echerichia coli*. *Agric. Biol. Chem.* **1990**, *54*, 1543−1550.

(22) Burnette, W. H. Western blotting: electrophoretic transfer of proteins from SDS-polyacrylamide gels to unmodified nitrocellulose and radiographic detection with antibody and radionated protein. *Anal. Biochem.* **1981**, *112*, 95−203.

(23) Green, A. A.; Hughes, W. L. *Mothods in Enzymology*; Colowick, S. P., Kaplan, N. O., Eds; Academic Press: New York, 1955; Vol. 1, p 67.

(24) Utsumi, S.; Gidamis, A. B.; Mikami, B.; Kito, M. Crystallization and preliminary X-ray crystallographic analysis of the soybean proglycinin expressed in *Escherichia coli*. *J. Mol. Biol.* **1993**, *233*, 177−178.

(25) Adachi, M.; Takenaka, Y.; Gidamis, A. B.; Mikami, B.; Utsumi, S. Crystal structure of soybean proglycinin A1aB1b homotrimer. *J. Mol. Biol.* **2001**, *305*, 291−305.

(26) Kim, C.-S.; Kamiya, S.; Sato, T.; Utsumi, S.; Kito, M. Improvement of nutritional value and functional properties of soybean glycinin by protein engineering. *Protein Eng.* **1990**, *3*, 725−731.

(27) Katsube, T.; Gidamis, A. B.; Kanamori, J.; Kang, I. J.; Utsumi, S.; Kito, M. Modification tolerability of the hypervariable region of soybean proglycinin. *J. Agric. Food Chem.* **1994**, *42*, 2639−2645.

(28) Maruyama, N.; Fukuda, T.; Saka, S.; Inui, N.; Kotoh, J.; Miyagawa, M.; Hayashi, M.; Sawada, M.; Moriyama, T.; Utsumi, S. Molecular and structural analysis of electrophoretic variants of soybean seed storage proteins. *Phytochemistry* **2003**, *64*, 701−708.

(29) Mori, T.; Maruyama, N.; Nishizawa, K.; Higasa, T.; Yagasaki, K.; Ishimoto, M.; Utsumi, S. The composition of newly synthesized proteins in the endoplasmic reticulum determines the transport pathways of soybean seed storage proteins. *Plant J.* **2004**, *40*, 238−349.

(30) Nakai, S. Structure−function relationships of food proteins with an emphasis on the importance of protein hydrophobicity. *J. Agric. Food Chem.* **1983**, *31*, 676−683.

(31) Nakai, S.; Li-Chan, E. *Hydrophobic Interaction in Food Systems*; CRC Press: Boca Raton, FL, 1988.

Soybean Proglycinins at Subunit Levels

J. Agric. Food Chem., Vol. 53, No. 9, 2005 **3657**

(32) Damodaran, S. Protein-stabilized forms and emulsions. In *Food Proteins and Their Applications*; Damodaran, S., Paraf, A., Eds.; Marcel Dekker: New York, 1997; pp 57−110.

(33) Lawrence, M. C.; Izard, T.; Beuchat, M.; Blagrove, R. J.; Colman, P. M. Structure of phaseolin at 2.2 Å resolution implications for a common vicilin/legumin structure and the genetic engineering of seed storage proteins. *J. Mol. Biol*. **1994**, *238*, 748−778.

(34) Adachi, M.; Kanamori, J.; Masuda T.; Yagasaki K.; Kitamura, K.; Mikami, B.; Utsumi, S. Crystal structure of soybean 11S globulin: Glycinin A3B4 homohexamer. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 7395−7400.