Modification Tolerability of the Hypervariable Region of Soybean Proglycinin

Tomoyuki Katsube,† Andrew Barde Gidamis, Jiro Kanamori, Il Jun Kang, Shigeru Utsumi,* and Makoto Kito

Research Institute for Food Science, Kyoto University, Uji, Kyoto 611, Japan

Glycinin is one of two major storage proteins of soybean seeds. To investigate the modification tolerability of the hypervariable region of proglycinin $A_{1a}B_{1b}$ subunit, chemical properties of the hypervariable region were changed by deleting a polyglutamic acid sequence in the hypervariable region and substituting it with polylysine, polyglutamine, and polymethionine sequences. All modified proglycinins were overproduced in *Escherichia coli* cells, accumulated as soluble proteins, and self-assembled into trimers such as native proglycinin. This indicates that the hypervariable region can tolerate all modifications attempted in this study. Modified proglycinins purified to near homogeneity were attempted for crystallization. Although all of the modified proglycinins were partially limited-proteolyzed in the hypervariable regions, they could form crystals consisting of the two proglycinin molecular species: intact and limited-proteolyzed. Results suggest that the hypervariable region of proglycinin forms a flexible structure. It is concluded that this flexible structure may be responsible for the modification tolerability of the hypervariable region.

Keywords: *Crystallization; glycinin; proglycinin; protein engineering; soybean*

INTRODUCTION

Glycinin is one of the dominant storage proteins of soybean seeds and plays an important role in their utilization in human and animal food systems. However, the nutritional quality and functional properties of glycinin are not wholly adequate for a wider use as a food resource. Therefore, improvement of the nutritional qualities and functional properties of glycinin has been attempted by means of chemical, physical, and biological methods (Utsumi and Kito, 1991). Of these methods, protein engineering appears to be a more promising method in achieving improvement of glycinin qualities because the primary sequence of glycinin can be modified consciously and systematically. To create novel soybean plants that produce modified glycinins having better food functions, it is important to note that the modifications introduced into glycinin should not interfere with any stage of biosynthesis and the accumulation process (Utsumi, 1992). From this viewpoint, it is very important to elucidate which regions of the glycinin molecule can tolerate modification and what kinds of modification it can accept without causing misfolding.

Argos et al. (1985) suggested the presence of a hypervariable region in the glycinin molecule from the comparison of the primary structures of 11S globulins (glycinin-type proteins) and 7S globulins (the other dominant storage proteins in legume seeds). Wright (1987, 1988) aligned the amino acid sequences to maximize the homology among 11S globulins from various legumes and nonlegumes and suggested the existence of five variable regions. The hypervariable region in the glycinin molecule proposed by Argos et al. (1985) corresponds to the fourth variable region from the N terminus as aligned by Wright (1988). Previous investigations into the modification tolerability of the variable regions by our group demonstrated that among the five variable regions at least the first, fourth, and fifth variable regions can tolerate certain modifications (Kim et al., 1990b). Similar observations have also been reported by Dickinson et al. (1990).

The hypervariable regions of 11S globulins are rich in hydrophilic and negatively charged amino acid residues and are suggested to be suitable targets for modifications to improve food qualities of 11S globulins (Argos et al., 1985). The hypervariable region of the proglycinin $A_{1a}B_{1b}$ subunit has a polyglutamic acid sequence composed of eight glutamic acid and one aspartic acid residues (Utsumi et al., 1987). In this study, we investigated the modification tolerability of the hypervariable region by altering the amino acid sequence of the polyglutamic acid sequence.

MATERIALS AND METHODS

Bacterial Strain, Medium, and Plasmids. We used *Escherichia coli* strain JM105 as the host cells (Yanisch-Perron et al., 1985) and LB medium for culture. Plasmids employed here were $pKGA_{1a}B_{1b}$-3 (Utsumi et al., 1988a) and $M13A_{1a}B_{1b}$ [a *Xho*I(filled-in)–*Pvu*II fragment (1.9 kbp) from $pAMA_{1a}B_{1b}2$ was inserted into the *Sma*I site of M13mp18] (Utsumi et al., 1993a). In the expression plasmid $pKGA_{1a}B_{1b}$-3, the ATG codon under the control of the *trc* promoter in the expression vector pKK233-2 (Pharmacia) was joined to the fourth codon in the $A_{1a}B_{1b}$ proglycinin cDNA (Utsumi et al., 1988a). The expressed proglycinin from $pKGA_{1a}B_{1b}$-3 was termed $A_{1a}B_{1b}$-3, in which the N-terminal methionine was retained (Utsumi et al., 1988a).

Site-Directed Mutagenesis and Construction of Expression Plasmids. Single-strand DNA of $M13A_{1a}B_{1b}$ was isolated for mutagenesis. Site-directed mutagenesis using a synthetic oligonucleotide primer containing the desired substitution was performed using a site-directed mutagenesis kit by Takara Shuzo based on the method of Kunkel (1985). The specific primers used were as follows: 5′-CTTGCACTGTGGCTTCTCTTCCTGGGGTCT-3′ for M13IV(ΔGlu), 5′-GCACTGTGGCTTCTTCTTCTTTTTTTTTTTGAGCTCTTCC-

* Author to whom correspondence should be addressed (telephone 81-774-32-3111, ext. 2731; fax 81-774-33-3004).

† Present address: Shimane Women's College, Matsue, Shimane 690, Japan.

0021-8561/94/1442-2639$04.50/0

TGGGGTC-3′ for M13IV(Lys), 5′-GCACTGTGGCTTCTGCTGCTGTTGTTGTTGGAGCTCTTCCTGGGGTC-3′ for M13IV(Gln), and 5′-GCACTGTGGCTTCTCCATCATCATCATCATGAGCTCTTCCTGGGGTC-3′ for M13IV(Met). As a result of mutation, the restriction site *Sac*I was generated in M13IV(Lys), M13IV(Gln), and M13IV(Met) and the restriction site *Fok*I was deleted in all cases. Mutants were selected by using restriction enzymes, and the mutations were verified by the dideoxy nucleotide sequencing method (Sanger et al., 1977).

Expression plasmids for the modified proglycinins IV(ΔGlu), IV(Lys), IV(Gln), and IV(Met) were constructed as follows; each small *Eco*T22I–*Eco*81I fragment (760 bp) from M13IV(ΔGlu), M13IV(Lys), M13IV(Gln), and M13IV(Met) was ligated with a large *Eco*T22I–*Eco*81I fragment (5300 bp) from pKG$A_{1a}B_{1b}$-3 to generate pKGIV(ΔGlu), pKGIV(Lys), pKGIV(Gln), and pKGIV(Met) for the modified proglycinins IV(ΔGlu), IV(Lys), IV(Gln), and IV(Met), respectively.

Expression and Detection of Modified Proglycinins from *E. coli*. JM105 cells harboring individual expression plasmids were cultured for 20 h at 37 °C after expression was induced by the addition of isopropyl β-D-thiogalactopyranoside (Kim et al., 1990a). Cells harvested by centrifugation were disrupted by sonication in 35 mM sodium phosphate buffer (pH 7.6) containing 0.4 M NaCl, 1 mM EDTA, and 2 mM phenylmethanesulfonyl fluoride (PMSF). The cell debris and the supernatant were fractionated by centrifugation (Kim et al., 1990a). The analysis of total cells, cell debris, and supernatant by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE) (Laemmli, 1970) was performed as described previously (Kim et al., 1990a). Amounts of expressed proglycinins were measured by densitometric analysis following Coomassie brilliant blue staining and/or immunoblotting of SDS gels (Utsumi et al., 1988b).

Protein Measurement. Proteins in the samples were determined according to the method of Bradford (1976) with bovine serum albumin as standard.

Analysis of Self-Assembly of the Modified Proglycinins. Soluble extracts (~5 mg/0.5 mL) of JM105 cells harboring individual expression plasmids were dialyzed against 35 mM potassium phosphate buffer (pH 7.6) containing 0.4 M NaCl, 1.5 mM PMSF, and 1 mM EDTA. After dialysis, assembly was analyzed by sucrose density gradient centrifugation followed by immunoblotting (Utsumi et al., 1988a). The 2S, 7S, and 11S fractions purified from soybean seeds according to the method of Thanh and Shibasaki (1976) were run in parallel as size markers.

Isoelectric Focusing of the Modified Proglycinins. The soluble extracts (~10 mg/mL) of JM105 cells harboring individual expression plasmids were diluted with 7 M urea containing 0.2 M 2-mercaptoethanol (2ME) and then electrophoresed using Bio-Lyte 3/10. After electrophoresis, proteins in the gels were blotted onto a nitrocellulose membrane and detected immunologically (Utsumi et al., 1988a).

Purification of the Modified Proglycinins from *E. coli*. Modified proglycinins were purified from JM105 cells from 4.8 L of culture as described in a previous paper (Utsumi et al., 1993b). Briefly, the modified proglycinins were extracted from the cells by sonication and purified by ammonium sulfate fractionation followed by column chromatography using a HiLoad 26/10 Q-Sepharose high-performance column (Pharmacia). Prior to sample loading, the column was equilibrated with 35 mM potassium phosphate buffer (pH 7.6) containing 0.12 M NaCl, 1.5 mM PMSF, 10 mM 2ME, and 0.02% (w/v) NaN_3 (buffer A). Elution was with a 500 mL linear gradient of 0.12–0.36 M NaCl in buffer A.

Crystallization of the Modified Proglycinins. The dialysis equilibrium method used for crystallization of normal proglycinin (Utsumi et al., 1993b) was at first followed for crystallization of the modified proglycinins (10 mg/mL) at 4 °C by varying the concentrations (0.1–0.2 M) and pH (7.0–9.0) of Tris-HCl buffer.

In addition to this method, IV(Met) (10 mg/mL) in 35 mM potassium phosphate buffer (pH 7.6) containing 0.15 M NaCl, 1.5 mM PMSF, 1 mM EDTA, 10 mM 2ME, and 0.02% NaN_3, IV(ΔGlu) and IV(Gln) in 0.125 and 0.15 M Tris-HCl buffer (pH 7.6) containing 1.5 mM PMSF, 1 mM EDTA, 10 mM 2ME, and 0.02% NaN_3, respectively, and IV(Lys) in 0.1 M acetate buffer (pH 6.0) containing 0.5 M NaCl, 1.5 mM PMSF, 1 mM EDTA, 10 mM 2ME, and 0.02% NaN_3 were also subjected to the hanging drop vapor diffusion method using polyethylene glycol 6000 (PEG) as precipitant. Samples (10 μL) placed on a glass plate were mixed with the same amount of the reservoir solution and allowed to equilibrate over 1.5 mL of the reservoir solution in a sealed vessel at 4–8 °C. The reservoir solution was 3–14% (w/v) PEG in the same buffers as were used for the proteins.

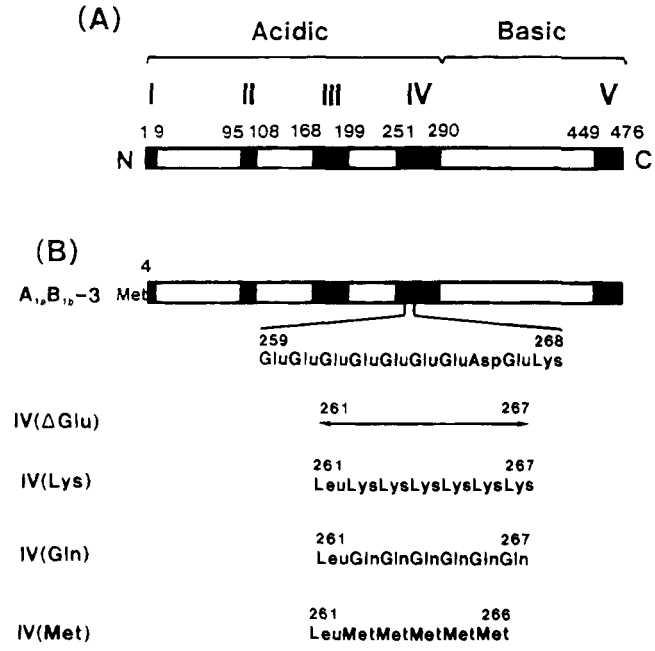

Figure 1. (A) Schematic representation of the proglycinin $A_{1a}B_{1b}$ subunit. The numbers of the residues from the N terminus for the variable regions (I–V) as aligned by Wright (1987, 1988) are shown above the alignment. Black and open areas are variable and conserved regions, respectively. Acidic and basic refer to the acidic and basic polypeptides, respectively. (B) Construction of normal expressed proglycinin $A_{1a}B_{1b}$-3 and modified proglycinins IV(ΔGlu), IV(Lys), IV(Gln), and IV(Met). $A_{1a}B_{1b}$-3 lacks the N-terminal three amino acids and retains the initiation methionine. IV(ΔGlu) lacks glutamic and aspartic acid residues from the 261st to the 267th. The glutamic and aspartic acid residues from the 261st to the 266th or the 267th are substituted with LeuLysLysLysLysLys, LeuGlnGlnGlnGlnGln, and LeuMetMetMetMetMet in IV(Lys), IV(Gln), and IV(Met), respectively.

RESULTS

Construction of Hypervariable Region-Modified Proglycinins. We mutated the nucleotide sequence encoding a polyglutamic acid sequence in the hypervariable region of $A_{1a}B_{1b}$ proglycinin by deleting it for modified proglycinin IV(ΔGlu) and substituting it with nucleotide sequences encoding polylysine for IV(Lys), polyglutamine for IV(Gln), and polymethionine for IV(Met) (Figure 1). As a result of these modifications, the hydropathy profile (Kyte and Doolittle, 1982) of the modified region changed in IV(ΔGlu) (the size of the hydrophilic region became small) and IV(Met) (the hydrophilic region was divided into two areas) but remained unaffected in IV(Lys) and IV(Gln) as shown in Figure 2. The net electric charge in all of the modified proglycinins constructed here changed from the normal state. The secondary structure in the vicinity of the modified region predicted according to the procedure of Chou and Fasman (1974a,b) did not vary,

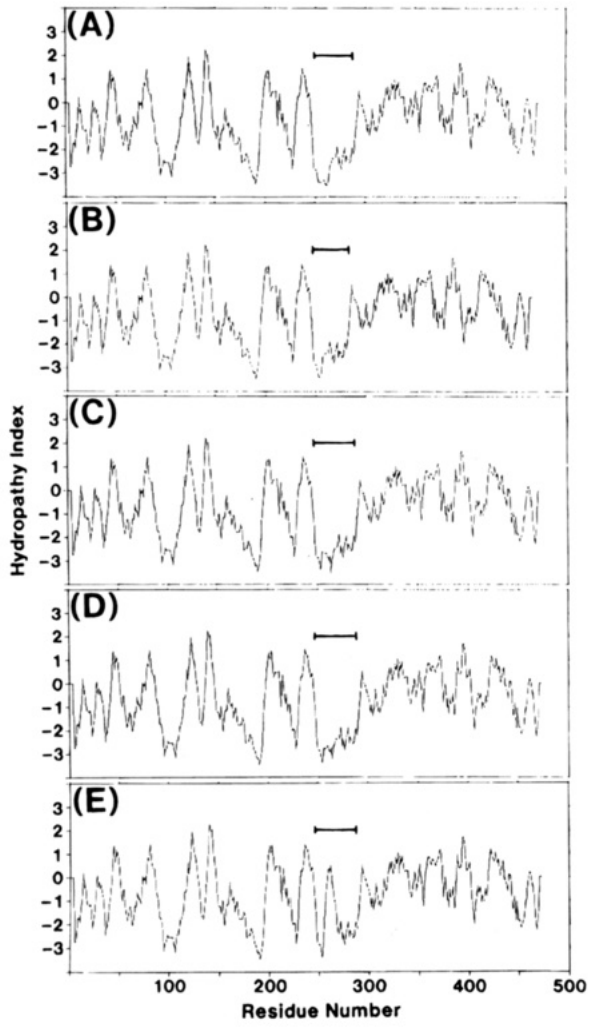

Figure 2. Hydropathy profiles of the normal and modified proglycinins: (A) normal proglycinin $A_{1a}B_{1b}$-3; (B) IV(ΔGlu); (C) IV(Lys); (D) IV(Gln); (E) IV(Met). The solid bar indicates the hypervariable region.

but that which was predicted according to the procedure of Robson's group (Robson and Suzuki, 1976; Garnier et al., 1978) varied from α-helix to β-turn and random in IV(ΔGlu) and from α-helix to α-helix and β-turn in IV(Gln) (data not shown).

Expression Level, Solubility, Self-Assembly, and Isoelectric Point of the Modified Proglycinins. The expression plasmids for the modified proglycinins IV-(ΔGlu), IV(Lys), IV(Gln), and IV(Met) were constructed by using the expression plasmid $pKGA_{1a}B_{1b}$-3 of which the expression level was ~20% of total *E. coli* proteins (Kim et al., 1990a). The nucleotide sequences in the vicinity of the translation initiation site and promoter of each expression plasmid constructed here are the same as that of $pKGA_{1a}B_{1b}$-3. Therefore, the efficiencies of transcription and translation are expected to be similar among the expression plasmids.

Production of the modified proglycinins in JM105 cells harboring individual expression plasmids was assessed by SDS−PAGE (Figure 3). The densitometric scanning of the gels indicates that all of the modified proglycinins accumulated in the cells at the level of ~20% of the total *E. coli* proteins (lanes 2−5) as compared to the scan of the extract from cells harboring the expression vector pKK233-2 alone (lane 1). The expression levels were the same as that of $A_{1a}B_{1b}$-3 (Kim et al., 1990a).

After disruption of the JM105 cells harboring individual expression plasmids, debris and soluble fractions

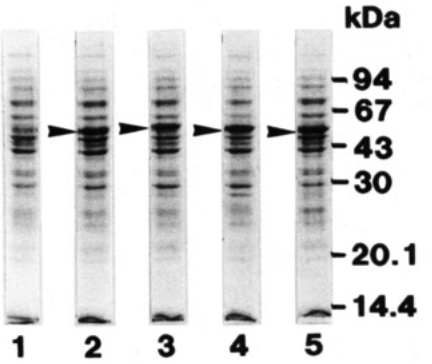

Figure 3. SDS−PAGE analysis of the expressed modified proglycinins from individual expression plasmids. The *E. coli* strain JM105 cells harboring individual expression plasmids were applied to SDS gels (11%, w/v) and stained with Coomassie brilliant blue. (Lane 1) pKK233-2; (lane 2) pKGIV-(ΔGlu); (lane 3) pKGIV(Gln); (lane 4) pKGIV(Lys); (lane 5) pKGIV(Met). Arrowheads indicate the position of the expressed modified proglycinins. The numbers on the right denote molecular mass.

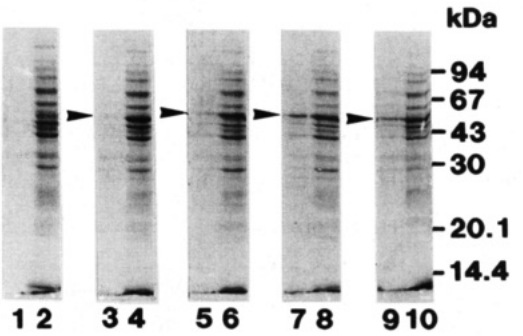

Figure 4. SDS−PAGE analysis of the expressed modified proglycinins in the soluble and insoluble fractions. The insoluble (lanes 1, 3, 5, 7, and 9) and soluble (lanes 2, 4, 6, 8, and 10) fractions of the extracts from the *E. coli* cells harboring pKK233-2 (lanes 1 and 2), pKGIV(ΔGlu) (lanes 3 and 4), pKGIV(Gln) (lanes 5 and 6), pKGIV(Lys) (lanes 7 and 8), and pKGIV(Met) (lanes 9 and 10) were analyzed as described in Figure 3. The amounts of proteins in the soluble and insoluble fractions subjected to analysis correspond to those from the same volume of the extracts. Arrowheads indicate the position of the expressed modified proglycinins. The numbers on the right denote molecular mass.

were subjected to SDS−PAGE and the expressed proteins were detected by Coomassie brilliant blue staining (Figure 4) and immunoblotting (data not shown). More than 80% of all the modified proglycinins were present in soluble fractions.

Soluble fractions of JM105 cells harboring individual expression plasmids were subjected to sucrose density gradient centrifugration. After fractionation, proteins in each fraction were analyzed by SDS−PAGE followed by immunoblotting. As shown in Figure 5, all of the modified proglycinins predominantly sedimented in fractions 9 and 11, which correspond to a size of 7−8S, indicating that they are able to self-assemble into trimers.

The soluble fractions of JM105 cells harboring individual expression plasmids were subjected to isoelectric focusing. The isoelectric points of the normal and modified proglycinins are summarized in Table 1. The observed isoelectric points were parallel to values calculated from their amino acid compositions (Manabe, 1982).

Purification of the Modified Proglycinins. IV-(ΔGlu), IV(Lys), IV(Gln), and IV(Met) expressed in

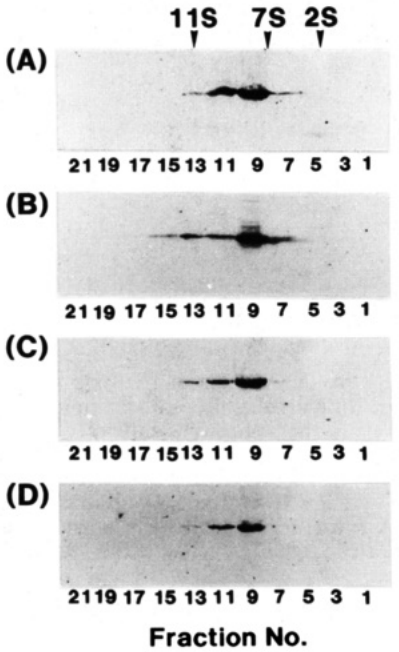

Figure 5. Self-assembly of the modified proglycinins. Assembly was assayed by centrifugation of the extracts from the *E. coli* cells harboring individual expression plasmids on a 12 mL 10–30% (w/v) linear sucrose density gradient. (A) IV(ΔGlu); (B) IV(Gln); (C) IV(Lys); (D) IV(Met). Sedimentation is from right to left. Sedimentation standards are given.

Table 1. Isoelectric Points of the Normal and Modified Proglycinins

protein	obsd[a]	calcd[b]	protein	obsd[a]	calcd[b]
$A_{1a}B_{1b}$-3	5.6	6.4	IV(Gln)	6.5	7.1
IV(ΔGlu)	6.6	7.3	IV(Met)	6.4	7.0
IV(Lys)	7.2	8.0			

[a] The values observed by isoelectric focusing in the presence of 7 M urea. [b] The values calculated from the amino acid compositions.

JM105 cells were purified following the same procedure as for $A_{1a}B_{1b}$-3 (Kim et al., 1990a) using buffer A as the column buffer. IV(ΔGlu), IV(Gln), and IV(Met) behaved similarly to $A_{1a}B_{1b}$-3 during the course of purification. However the concentration of NaCl required for their elution from a Q-Sepharose column was lower than that for $A_{1a}B_{1b}$-3 due to the difference in the isoelectric points of the normal and modified proglycinins (see Table 1). On the other hand, IV(Lys) easily precipitated during the dialysis against the column buffer. Since the precipitate contained IV(Lys) at the purity of ~80% (see Figure 6) and could be solubilized by a high ionic strength buffer ($\mu > 0.4$), it was used, as such, as the purified modified proglycinin IV(Lys).

The purified modified proglycinins were subjected to SDS–PAGE (Figure 6). In the absence of 2ME, two or three bands with molecular sizes of ~55 000 corresponding to a proglycinin subunit were predominantly observed (lanes 3–6): two bands in IV(ΔGlu), IV(Gln), and IV(Met) (lanes 3–5) and three bands in IV(Lys) (lane 6). On the other hand, in the presence of 2ME, all of the modified proglycinins gave two bands with molecular sizes of 30 000–34 000 and two bands with 20 000–24 000 in addition to one band corresponding to an intact proglycinin subunit (lanes 8–11). These results indicate that all of the modified proglycinins were partly limited-proteolyzed during purification, resulting in two species of protein with molecular sizes of 30 000–34 000 and 20 000–24 000, linked by a disulfide bond. The mature glycinin subunit, which is composed of the acidic and basic polypeptides linked by a disulfide bond, moved faster than the native proglycinin (lanes 1 and 2). Therefore, it is likely that the bands with molecular sizes of ~55 000 having slower and faster mobilities in each modified proglycinin (lanes 3–6) correspond to the intact proglycinin and the limited-proteolyzed proglycinin, respectively. The ratio of the intact proglycinin to the limited-proteolyzed proglycinin was >70% in IV(ΔGlu), >60% in IV(Gln) and IV(Met), and <30% in IV(Lys) (lanes 3–6).

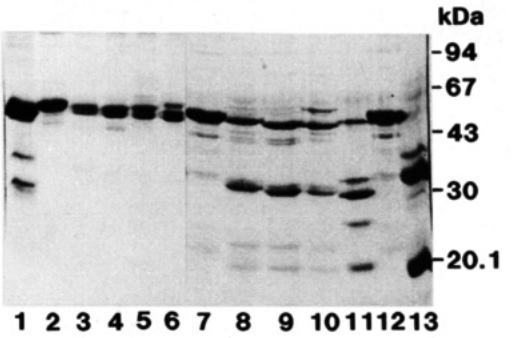

Figure 6. SDS–PAGE analysis of the purified modified proglycinins. The purified modified proglycinins (~3 μg) were subjected to SDS–PAGE in the absence (lanes 1–6) and presence (lanes 7–13) of 2ME and stained with Coomassie brilliant blue. (Lanes 1 and 13) Soybean glycinin; (lanes 2, 7, and 12) normal proglycinin; (lanes 3 and 8) IV(ΔGlu); (lanes 4 and 9) IV(Gln); (lanes 5 and 10) IV(Met); (lanes 6 and 11) IV(Lys). The numbers on the right denote molecular mass.

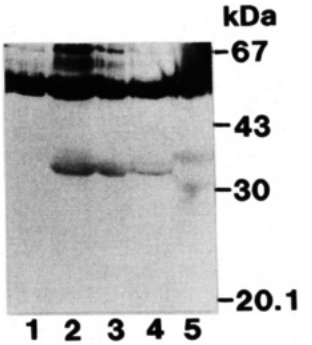

Figure 7. Immunological detection of the modified proglycinins. The purified modified proglycinins (~3 μg) were subjected to SDS–PAGE in the presence of 2ME. (Lane 1) Normal proglycinin; (lane 2) IV(Gln); (lane 3) IV(ΔGlu); (lane 4) IV(Met); (lane 5) IV(Lys). The numbers on the right denote molecular mass.

Proglycinin-related proteins separated by SDS–PAGE in the presence of 2ME were detected immunologically by using anti-glycinin serum (Figure 7). The anti-glycinin serum prepared against the native glycinin reacts weakly with the acidic polypeptides but not with the basic polypeptides (Moreira, 1980; Ellis et al., 1988; Utsumi et al., 1993c). The bands having the sizes of ~55 000 and 30 000–34 000 were observed, but not the bands having the sizes of 20 000–24 000 (Figure 7). This indicates that the bands with molecular sizes of 30 000–34 000 and 20 000–24 000 were derived from the acidic and basic polypeptide regions of the modified proglycinins, respectively. Thus, it is most likely that the limited proteolysis of the modified proglycinins occurred in their hypervariable regions as shown in Figure 8.

The purified modified proglycinins were subjected to sucrose density gradient centrifugation (data not shown). The results showed that the purified modified progly-

Table 2. Crystallization of the Modified Proglycinins at Various Conditions of Tris-HCl Buffer

	concentration and pH[a] of Tris-HCl buffer									
	pH 7.0	pH 7.6		pH 8.0		pH 8.25		pH 8.5		pH 9.0
protein	0.125 M	0.125 M	0.15 M	0.125 M	0.15 M	0.125 M	0.15 M	0.125 M	0.15 M	0.15 M
IV(ΔGlu)	–[b]	ppt[c]	yes[d]	no[d]	–	yes	–	yes	yes	–
IV(Lys)	–	ppt	ppt	ppt	ppt	ppt	ppt	ppt	ppt	ppt
IV(Gln)	no	no	no	yes	no	ppt	no	ppt	no	–
IV(Met)	–	yes	yes	yes	–	no	–	ppt	–	–

[a] Adjusted at room temperature. [b] Not done. [c] Precipitate formation. [d] Yes/no; refer to crystal formation.

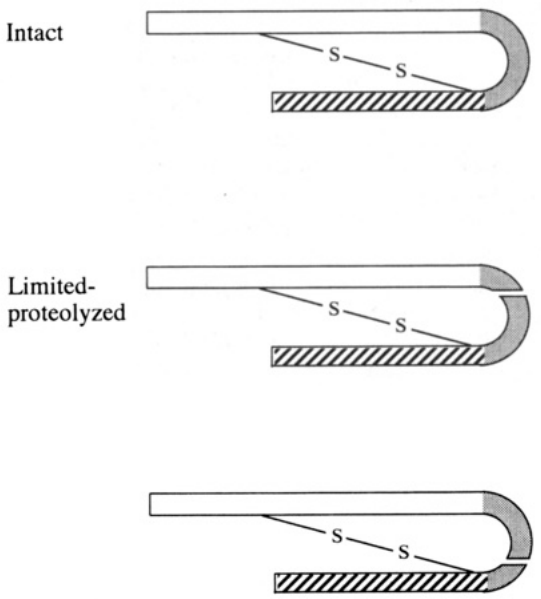

Figure 8. Schematic representation of intact and limited proteolyzed proglycinins. Open, lightly shaded, and hatched areas indicate acidic polypeptide, hypervariable, and basic polypeptide regions, respectively.

cinins sedimented in fractions corresponding to trimers. This indicates that the trimer structure of the modified proglycinins was stable during the course of purification and was not affected by limited proteolysis.

Crystallization of the Modified Proglycinins. As summarized in Table 2, IV(ΔGlu), IV(Gln), and IV(Met) formed crystals during dialysis against Tris-HCl buffer within 1 week. IV(Lys) has a tendency to form precipitate in all conditions listed in Table 2 and also at pH 5.5 and 6.0 using acetate buffer. The crystals obtained by the dialysis method did not, however, grow to a size suitable for X-ray analysis. All of the modified proglycinins were then subjected to the hanging drop vapor diffusion method using PEG as a precipitant. Within 2–3 days, IV(ΔGlu), IV(Gln), and IV(Met) formed crystals in the protein–PEG solutions containing 3–8%, 4–6%, and 6–8% PEG, respectively. The largest crystals of IV(ΔGlu), IV(Gln), and IV(Met) being more than 0.3 mm in length were formed in 4%, 6%, and 8% PEG, respectively, within 2 weeks. IV(Lys) formed small crystals in 13% PEG after 1 month and grew to 0.5 mm in length after 3 months. Figure 9 shows the crystals of the modified proglycinins obtained by the hanging drop vapor diffusion method. The shape of the crystals of IV(ΔGlu), IV(Lys), IV(Gln), and IV(Met) is different from that of the normal proglycinin crystals (Gidamis et al., 1994).

The protein samples used for crystallization consisted of two kinds of molecular species; an intact proglycinin and a limited-proteolyzed proglycinin. To examine whether the crystals of the modified proglycinins consisted of only the intact species or not, the crystals were

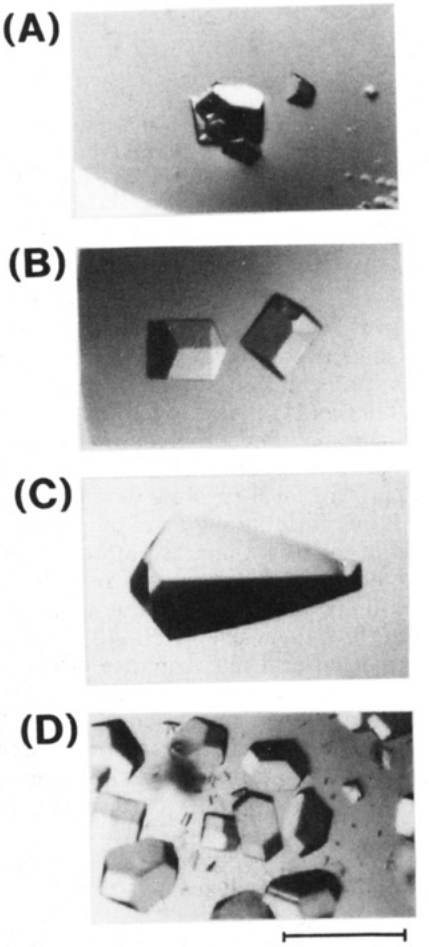

Figure 9. Photomicrographs of crystals of the modified proglycinins. (A) Crystals of IV(ΔGlu) formed at 4 °C in 0.125 M Tris-HCl buffer (pH 7.6) containing 4% PEG; (B) crystals of IV(Lys) formed at 8 °C in 0.1 M acetate buffer (pH 6.0) containing 0.5 M NaCl and 13% PEG; (C) crystals of IV(Gln) formed at 4 °C in 0.15 M Tris-HCl buffer (pH 7.6) containing 6% PEG; (D) crystals of IV(Met) formed at 4 °C in 35 mM potassium phosphate buffer (pH 7.6) containing 8% PEG. The bar represents 1 mm.

subjected to SDS–PAGE (Figure 10). The SDS–PAGE profiles of the individual crystals were very similar to those of the samples used for crystallization (compare Figures 6 and 10). Thus, these crystals are composed of both molecular species.

DISCUSSION

Modification tolerability can be assessed from the ability of the modified proglycinins to form proper conformation similar to that of the normal proglycinin. We proposed previously the following three criteria for judging the formation of proper conformation: (1) a high level of expression in *E. coli* (10–20% of *E. coli* total proteins), (2) solubilities comparable with that of globulins, and (3) self-assembly into trimers (Kim et al.,

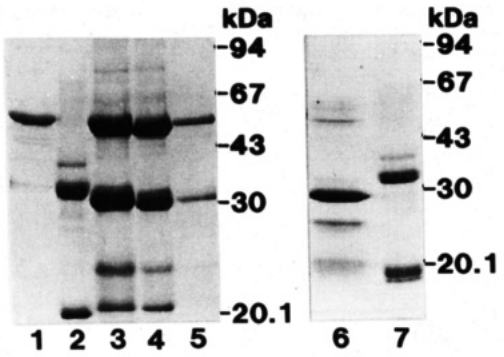

Figure 10. SDS–PAGE analysis of the crystals of modified proglycinins. The crystals were washed with the buffer used for crystallization and then subjected to SDS–PAGE in the presence of 2ME. (Lane 1) Normal proglycinin; (lane 2) soybean glycinin; (lane 3) crystals of IV(Gln); (lane 4) crystals of IV(Met); (lane 5) crystals of IV(ΔGlu); (lane 6) crystals of IV(Lys); (lane 7) soybean glycinin. The numbers on the right denote molecular mass.

1990b). Since all of the modified proglycinins constructed here, i.e., IV(ΔGlu), IV(Lys), IV(Gln), and IV(Met), satisfied all three criteria, we concluded that they assumed a conformation similar to that of native proglycinin. This conclusion was confirmed by the finding that all of the modified proglycinins constructed here and in the previous studies (Kim et al., 1990b; Utsumi et al., 1993a) can form crystals (Gidamis et al., 1994). Thus, the hypervariable region of proglycinin can tolerate all of the modifications attempted here. In other words, the negative charge, the strong hydrophilicity, and the α-helix structure of the polyglutamic acid sequence in the hypervariable region are not important for the formation and maintenance of the proglycinin trimer structure. Given the results of these modifications, it is likely that each modified proglycinin has different structural characteristics as suggested from the observation that suitable crystallization conditions and the shape of the crystals varied among the modified proglycinins.

All of the modified proglycinins constructed here were partially limited-proteolyzed to give two species of protein with molecular masses of 30 000–34 000 and 20 000–24 000 which are linked by a disulfide bond during the dialysis against buffer A ($\mu = 0.22$). These modified proglycinins were fairly resistant to proteinase digestion during extraction from the cells and the analysis by sucrose density gradient centrifugation (see Figures 4 and 5). Previously we constructed many modified proglycinins ΔI, ΔV8, Gly12, Ser88, IV+4Met, and V+4Met which can form proper conformation similar to that of the normal proglycinin (Kim et al., 1990b; Utsumi et al., 1993a). Among these, IV+4Met, in which an oligopeptide of four contiguous methionine residues was inserted into the hypervariable region, was also limited-proteolyzed like the modified proglycinins constructed here. However, the other modified proglycinins ΔI, ΔV8, Gly12, Ser88, and V+4Met, in which modifications were not introduced into the hypervariable region, were stable at low ionic strength condition (Kim et al., 1990b; Utsumi et al., 1993a). These observations indicate that the conformational integrity of the proglycinins modified in the hypervariable region may have been partly lost at low ionic strength conditions ($\mu < 0.22$) and therefore became susceptible to *E. coli* proteinase attack. As a result of partial limited proteolysis, the purified proteins contained two kinds of proglycinin molecular species: intact and limited-proteolyzed. However, all of the modified proglycinins could form crystals that are also composed of the two molecular species. These facts indicate that the intact and limited-proteolyzed molecular species share a similar conformation (see Figure 8) and that the hypervariable region of proglycinin may form a very flexible structure. In other words, this region would be invisible on an electron density map upon X-ray crystallographic analysis of proglycinin. A similar phenomenon was observed in the three-dimensional structure of glutathione synthetase of *E. coli* B (Tanaka et al., 1992; Yamaguchi et al., 1993). Only the peptide bond Arg233–Gly234 in the region between Ile226 and Arg241, which is invisible on the electron density map, is cleaved by arginylendopeptidase, even though the glutathione synthetase contains 19 residues of arginine, most of which are located at the surface of the molecule. Such a flexibility may be responsible for the hypervariable region's ability to tolerate many kinds of modification.

As described above, the proglycinins modified in the hypervariable region are susceptible to *E. coli* proteinase attack. This does not mean that such modifications are not suitable for improvement of food qualities of glycinin, because we observed that IV+Met can be synthesized, accumulated, processed, and self-assembled similarly to normal glycinin in tobacco seeds (Utsumi et al., 1993c). Still, proteinase attack susceptibility should be considered when such modified proteins are extracted from plants.

Soybean proteins do not fully exhibit their functional properties in the acidic range (pH 3–6) because their solubility decreases due to the effect of their isoelectric points, which also lie within the acidic range (Kinsella, 1979). Soybean proteins are therefore rarely used for acidic foods such as mayonnaise and yogurt. The isoelectric points of the modified proglycinins constructed here are higher than that of the normal proglycinin (Table 1). Therefore, these modified proglycinins are expected to exhibit good functional properties in the acidic range.

Recently, modifications based on gene manipulation to improve nutritional qualities of seed proteins (Hoffman et al., 1988; Wallace et al., 1988; Ohtani et al., 1991; De Clercq et al., 1990; Saalbach et al., 1990) and to produce the neuropeptide Leu-enkephalin in plants (Vandekerckhove et al., 1989) were attempted. Successful attempts were restricted to the modifications that were introduced into variable regions of the target proteins. The variable regions were deduced from comparison of the amino acid sequences of similar type proteins to that of the target protein. These results coupled with our findings suggest that the hypervariable region can be selected as a suitable target for modifications to improve the food qualities of the target proteins even if the three-dimensional structure of the target protein is not known.

ACKNOWLEDGMENT

This work was supported in part by grants from the Ministry of Education, Science and Culture of Japan and Research Committee of Soy Protein (to S.U.).

LITERATURE CITED

Argos, P.; Narayana, S. V. L.; Nielsen, N. C. Structural Similarity between Legumin and Vicilin Storage Proteins from Legumes. *EMBO J.* **1985**, *4*, 1111–1117.

Bradford, M. M. A Rapid and Sensitive Method for the Quantitation of Microgram Quantities of Protein Utilizing the Principle of Protein-Dye Binding. *Anal. Biochem.* **1976**, *72*, 248–254.

Chou, P. Y.; Fasman, G. D. Conformational Parameters for Amino Acids in Helical, β-sheet, and Random Coil Regions Calculated from Proteins. *Biochemistry* **1974a**, *13*, 211–222.

Chou, P. Y.; Fasman, G. D. Prediction of Protein Conformation. *Biochemistry* **1974b**, *13*, 222–245.

De Clearcq, A.; Vandewiele, M.; Van Damme, J.; Guerche, P.; Van Montagu, M.; Vandekerckhove, J.; Krebbers, E. Stable Accumulation of Modified 2S Albumin Seed Storage Proteins with Higher Methionine Contents in Transgenic Plants. *Plant Physiol.* **1990**, *94*, 970–979.

Dickinson, C. D.; Scott, M. P.; Hussein, E. H. A.; Argos, P.; Nielsen, N. C. Effect of Structural Modifications on the Assembly of a Glycinin Subunit. *Plant Cell* **1990**, *2*, 403–413.

Ellis, J. R.; Shirsat, A. H.; Hepher, A.; Yarwood, J. N.; Gatehouse, J. A.; Croy, R. R. D.; Boulter, D. Tissue-specific Expression of a Pea Legumin Gene in Seeds of *Nicotiana plumbaginifolia*. *Plant Mol. Biol.* **1988**, *10*, 203–214.

Garnier, J.; Osguthorpe, D. J.; Robson, B. Analysis of the Accuracy and Implications of Simple Methods for Predicting the Secondary Structure of Globular Proteins. *J. Mol. Biol.* **1978**, *120*, 97–120.

Gidamis, A. B.; Mikami, B.; Katsube, T.; Utsumi, S.; Kito, M. Crystallization and Preliminary X-ray Analysis of Soybean Proglycinins Modified by Protein Engineering. *Biosci., Biotechnol., Biochem.* **1994**, *58*, 703–706.

Hoffman, L. M.; Donaldson, D. D.; Herman, E. M. A Modified Storage Protein Is Synthesized, Processed, and Degraded in the Seeds of Transgenic Plant. *Plant Mol. Biol.* **1988**, *11*, 717–729.

Kim, C.-S.; Kamiya, S.; Kanamori, J.; Utsumi, S.; Kito, M. High-level Expression, Purification and Functional Properties of Soybean Proglycinin from *Escherichia coli*. *Agric. Biol. Chem.* **1990a**, *54*, 1543–1550.

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.* **1990b**, *3*, 725–731.

Kinsella, J. E. Functional Properties of Soy Proteins. *J. Am. Oil Chem. Soc.* **1979**, *56*, 242–258.

Kunkel, T. A. Rapid and Efficient Site-specific Mutagenesis without Phenotypic Selection. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 488–492.

Kyte, J.; Doolittle, R. F. A Simple Method for Displaying the Character of a Protein. *J. Mol. Biol.* **1982**, *157*, 105–132.

Laemmli, U. K. Cleavage of Structural Proteins During Assembly of the Head of Bacteriophage T4. *Nature* **1970**, *227*, 680–685.

Manabe, T. Utilization of Microcomputer in Protein Chemistry. *Kagaku No Ryoiki* **1982**, *36*, 470–486.

Moreira, M. A. Purification and Characterization of the Acidic and Basic Polypeptides of Glycinin. *Diss. Abst. Int. B.* **1980**, *41*, 2007B.

Ohtani, T.; Galili, G.; Wallace, J. C.; Thompson, G. A.; Larkins, B. A. Normal and Lysine-containing Zeins Are Unstable in Transgenic Tobacco Seeds. *Plant Mol. Biol.* **1991**, *16*, 117–128.

Robson, B.; Suzuki, E. Conformational Properties of Amino Acid Residues in Globular Proteins. *J. Mol. Biol.* **1976**, *107*, 327–356.

Saalbach, G.; Jung, R.; Kunze, G.; Manteuffel, R.; Saalbach, I.; Müntz, K. Expression of Modified Legume Storage Protein Genes in Different Systems and Studies on Intracellular Targeting of *Vicia faba* Legumin in Yeast. In *Genetic Engineering of Crop Plants*; Lycett, G. W., Grierson, D., Eds.; Butterworth: London, 1990; pp 151–158.

Sanger, F.; Nicklen, S.; Coulson, A. R. DNA Sequencing with Chain-Terminating Inhibitors. *Proc. Natl. Acad. Sci. U.S.A.* **1977**, *74*, 5463–5467.

Tanaka, T.; Kato, H.; Nishioka, T.; Oda, J. Mutational and Proteolytic Studies on a Flexible Loop in Glutathione Synthetase from *Escherichia coli* B: The Loop and Arginine 233 Are Critical for the Catalytic Reaction. *Biochemistry* **1992**, *31*, 2259–2265.

Thanh, V. H.; Shibasaki, K. Major Proteins of Soybean Seeds. A Straightforward Fractionation and Their Characterization. *J. Agric. Food Chem.* **1976**, *24*, 1117–1121.

Utsumi, S. Plant Food Protein Engineering. In *Advance in Food and Nutrition Research*; Kinsella, J. E., Ed.; Academic: San Diego, CA, 1992; Vol. 36, pp 89–208.

Utsumi, S.; Kito, M. Improvement of Food Protein Functions by Chemical, Physical, and Biological Modifications. *Comments Agric. Food Chem.* **1991**, *2*, 261–278.

Utsumi, S.; Kohno, M.; Mori, T.; Kito, M. An Alternate cDNA Encoding Glycinin $A_{1a}B_x$ Subunit. *J. Agric. Food Chem.* **1987**, *35*, 210–214.

Utsumi, S.; Kim, C.-S.; Sato, T.; Kito, M. Signal Sequence of Preproglycinin Affects Production of the Expressed Protein in *Escherichia coli*. *Gene* **1988a**, *71*, 349–358.

Utsumi, S.; Sato, T.; Kim, C.-S.; Kito, M. Processing of Preproglycinin Expressed from cDNA-encoding $A_{1a}B_{1b}$ Subunit in *Saccharomyces cerevisiae*. *FEBS Lett.* **1988b**, *233*, 273–276.

Utsumi, S.; Gidamis, A. B.; Kanamori, J.; Kang, I. J.; Kito, M. Effects of Deletion of Disulfide Bonds by Protein Engineering on the Conformation and Functional Properties of Soybean Proglycinin. *J. Agric. Food Chem.* **1993a**, *41*, 687–691.

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.* **1993b**, *233*, 177–178.

Utsumi, S.; Kitagawa, S.; Katsube, T.; Kang, I. J.; Gidamis, A. B.; Takaiwa, F.; Kito, M. Synthesis, Processing and Accumulation of Modified Glycinins of Soybean in the Seeds, Leaves and Stems of Transgenic Tobacco. *Plant Sci.* **1993c**, *92*, 191–202.

Vandekerckhove, J.; Van Damme, J.; Van Lijsebettens, M.; Botterman, J.; De Block, M.; Vandewiele, M.; De Clercq, A.; Leemans, J.; Van Montagu, M.; Krebbers, E. Enkephalins Produced in Transgenic Plants Using Modified 2S Seed Storage Proteins. *Bio/Technology* **1989**, *7*, 929–932.

Wallace, J. C.; Galili, G.; Kawata, E. E.; Cuellar, R. E.; Shotwell, M. A.; Larkins, B. A. Aggregation of Lysine-containing Zeins into Protein Bodies in *Xenopus* Oocytes. *Science* **1988**, *240*, 662–664.

Wright, D. J. The Seed Globulins. In *Developments in Food Proteins*; Hudson, B. J. F., Ed.; Elsevier: London, 1987; Vol. 5, pp 81–157.

Wright, D. J. The Seed Globulins. In *Developments in Food Proteins*; Hudson, B. J. F., Ed.; Elsevier: London, 1988; Vol. 6, pp 119–178.

Yamaguchi, H.; Kato, H.; Hata, Y.; Nishioka, T.; Kimura, A.; Oda, J.; Katsube, Y. Three-dimensional Structure of the Glutathione Synthetase from *Escherichia coli* B at 2.0 Å Resolution. *J. Mol. Biol.* **1993**, *229*, 1083–1100.

Yanisch-Perron, C.; Vieira, J.; Messing, J. Improved M13 Phage Cloning Vectors and Host Stains: Nucleotide Sequences of the M13mp18 and pUC19 Vectors. *Gene* **1985**, *33*, 103–119.

Received for review April 11, 1994. Accepted August 22, 1994.[⊗]

[⊗] Abstract published in *Advance ACS Abstracts*, October 1, 1994.