High-level Expression, Purification and Functional Properties of Soybean Proglycinin from *Escherichia coli*

Chan-Shick Kim, Seigo Kamiya, Jito Kanamori, Shigeru Utsumi and Makoto Kito

The Research Institute for Food Science, Kyoto University,
Uji, Kyoto 611, Japan
Received January 29, 1990

The proglycinin synthesized in *E. coli* JM105 comprised approximately 20% of the total bacterial proteins, with a yield of 39 mg per liter of culture under the optimum cultivation conditions. The proglycinin was purified to homogeneity by salt precipitation, ion-exchange chromatography, and cryoprecipitation. The purified proglycinin self-assembled to a trimer with a secondary structure similar to that of the glycinin half-molecule from soybean seeds, and had properties of gel formation by heating and precipitation with calcium salt as the native glycinin and glycinin half-molecule do. This indicated that the *E. coli* expression system of glycinin cDNA may be used for the evaluation of the self-assembly and the food qualities of protein-engineered soybean proteins.

One of the major objectives of the food industry is the enrichment of the functional properties and nutritional value of soybean (*Glycine max* L.) storage proteins.\(^1\) Recent progress in gene manipulation and elucidation of the interrelationship between structure and food functionality of soybean proteins have made it possible to adopt protein engineering for the enhancement of soybean quality. To attain this goal, an expression system of cDNAs encoding native and protein-engineered soybean proteins in a microorganism must be developed and then the ability of self-assembly and the functionalities of the expressed proteins should be evaluated before the modified genes are transferred to soybean plants.

Glycinin is one of the predominant storage proteins of soybean and is important in the functional properties of soybean proteins. It is composed of six subunits, each of which consists of an acidic and a basic polypeptide that are linked by a disulfide bridge.\(^2\)\(^\text{-}^6\) Initially, a single polypeptide precursor (pre-proglycinin) consisting of covalently linked acidic and basic polypeptides with a signal sequence is synthesized. The signal sequence is removed cotranslationally in the endoplasmic reticulum.\(^7\) The resultant proglycinin subunits assemble into trimers of about 8S.\(^8\)\(^9\) These complexes move from the endoplasmic reticulum into protein bodies where a specific post-translational cleavage occurs.\(^10\)\(^11\) The cleavage results in subunits that consist of an acidic and a basic polypeptide and assemble into hexamers of about 12S.\(^7\)\(^\text{-}^9\) The following subunits have been identified: A\(_{1A}\)B\(_{1b}\), A\(_{1b}\)B\(_2\), A\(_{2A}\)B\(_{1a}\), A\(_{3B}\) and A\(_{3A}\)B\(_3\).\(^5\)\(^\text{-}^17\) These subunits are classified as the A\(_1\) type (A\(_{1A}\)B\(_{1b}\), A\(_{1b}\)B\(_2\), A\(_{2B}\)A\(_{1A}\)) and the A\(_3\) type (A\(_{3B}\), A\(_{3A}\)B\(_3\)) according to the similarity of N-terminal amino acid sequences.\(^7\)\(^10\) The nutritional quality of the A\(_1\) type is superior to that of the others.\(^10\)

A low level of expression of cDNAs encoding the pre-proglycinins A\(_{1A}\)B\(_{1b}\) and A\(_{2B}\)A\(_{1A}\) in *Escherichia coli* MV1190 has been reported.\(^18\) The expression in *Saccharomyces cerevisiae* of a cDNA encoding the pre-proglycinin A\(_{1A}\)B\(_{1b}\) has been reported.\(^19\) The stepwise deletion of the cDNA region encoding the signal sequence

---

\(^1\) To whom all correspondence should be addressed.
and N-terminal regions of the pre-proglycinin $A_{1a}B_{18}$ results in the increase of the expression level and the stability of the products in E. coli compared with those of the products having its signal sequence. However, the expression level is still low (approximately 6 nmol/l of culture) for application to protein engineering of soybean proteins. The availability of large quantities of expressed proglycinin is crucial to the adoption of protein engineering to improve quality of soybean proteins.

Here we describe the high-level expression in E. coli of glycmin cDNA, purification of the expressed protein, and structures and functional properties of the expressed protein compared with those of glycmin and glycmin half-molecule (trimer, 7S) from soybean seeds.

Materials and Methods

**Bacterial strain, medium and plasmid.** E. coli strain JM105 was used as the host cell. LB medium (pH 7.5) consisted of 1% Bactotryptone, 0.5% yeast extract and 1% NaCl. The expression plasmid used here was pKGA$_{1a}B_{18}$-3, in which the ATG codon in pKK233-2, the expression vector was joined to the fourth codon in the cDNA encoding the proglycinin $A_{1a}B_{18}$. The proglycinin encoded by pKGA$_{1a}B_{18}$-3 was termed $A_{1a}B_{18}$-3.

**Expression and detection of $A_{1a}B_{18}$-3 in E. coli.** Three hundred ml of LB medium containing 25 mg/ml ampicillin was inoculated with 3 ml of a full-grown culture of JM105 harboring the expression plasmid pKGA$_{1a}B_{18}$-3 and cultured at 37°C with a shaking speed of around 90 strokes/min. At Ab$_{600}$ = 0.3, isopropyl-$\beta$-D-thiogalactopyranoside (IPTG) was added to a final concentration of 1 mm. After incubation for 20 hr at 37°C, the induced cells were harvested by centrifugation. A sample was boiled in SDS-sample buffer (62.5 mM Tris-HCl, pH 6.8, 2% SDS, 20% glycerol and 0.2 mM 2-mercaptoethanol) directly, and analyzed by SDS-polyacrylamide gel electrophoresis (SDS-PAGE).

**Purification of $A_{1a}B_{18}$-3 from E. coli.** All steps were done at 4°C.

**Step 1: Extraction and centrifugation.** Forty grams of E. coli (strain JM105/pKGA$_{1a}B_{18}$-3) cells (wet weight) from 12 l of culture were suspended in 240 ml of 50 mM Tris-HCl buffer (pH 8.0) containing 1 mM EDTA and 1.5 mM phenylmethylsulfonyl fluoride (PMSF) (buffer A), and lysozyme was added to a final concentration of 1 mg/ml. After the suspension was incubated on ice for 30 min, the cells were disrupted by sonication. The cell debris and unbroken cells were removed by centrifugation at 10,000×g for 20 min. $A_{1a}B_{18}$-3 was purified from the resulting supernatant.

**Step 2: Fractionation with (NH$_4$)$_2$SO$_4$.** The crude extract was adjusted carefully to 40% saturation with solid (NH$_4$)$_2$SO$_4$. After this was stirred for 30 min, the precipitate was removed by centrifugation at 10,000×g for 20 min. The supernatant was adjusted to 65% saturation and stirred for 30 min. The precipitated protein was collected by centrifugation at 10,000×g for 20 min, and resuspended in 35 mM potassium phosphate buffer (pH 7.6) containing 0.15 mM NaCl, 10 mM 2-mercaptoethanol, and 0.02% NaN$_3$ (buffer B). The suspension was dialysed overnight against 3×21 of buffer B.

**Step 3: Q-Sepharose column chromatography.** The dialyzed material was put onto a column of Q-Sepharose (1.7×24 cm) (Pharmacia) equilibrated with buffer B. The column was washed with buffer B until the A$_{280}$ of the eluate had fallen below 0.1. The column was then eluted with a 500-mM linear gradient of 0.15-0.5 M NaCl in buffer B. Fractions (21 ml/fraction) containing $A_{1a}B_{18}$-3 were monitored by SDS-PAGE and pooled.

**Step 4: Cryoprecipitation.** The pooled material was dialyzed against 6 mM Tris-HCl buffer (pH 6.3) containing 10 mM 2-mercaptoethanol and 0.02% NaN$_3$ (buffer C) for 2 days with several changes of the dialysis buffer. The dialyzed material was centrifuged at 10,000×g for 20 min. The precipitate was suspended in 3.5 mM potassium phosphate buffer (pH 7.6) containing 0.02% NaN$_3$ (buffer D), and dialyzed overnight against buffer D. The purified preparation thus obtained was judged to be homogeneous by SDS-PAGE under reducing conditions (Fig. 2, lane 4).

**SDS-PAGE.** SDS-PAGE was done by the method of Laemmli using 11% polyacrylamide gels. Proteins were visualized with Coomassie Brilliant Blue R-250. The relative amount of $A_{1a}B_{18}$-3 in each sample was estimated from a densitometric scan using a Shimadzu Dual-Wavelength TLC Scanner CS-910 densitometer.

**Protein measurement.** Protein was measured by the method of Bradford with bovine serum albumin as a standard.

**Preparation of glycmin and glycmin half-molecule.** The purified glycmin fraction was prepared from soybean seeds (Glycine max L. var. Tsuru-no-ko) as described previously. The glycmin half-molecule (trimer, 7S) was prepared from the purified glycmin solution by dialysis against buffer D and sucrose density gradient centrifugation by the method of Utsumi et al.

**Analysis of self-assembly of $A_{1a}B_{18}$-3 expressed in E. coli.** The purified $A_{1a}B_{18}$-3 was dialyzed against 3.5 mM potassium phosphate buffer (pH 7.6) (buffer E). Assembly was assayed by sucrose density gradient centrifugation as described previously except for 3.5 mM potassium phos-
Expression and Functionality of Proglycinin from *E. coli*

The purified A1aB1b-3, glycinin and glycinin half-molecule (7S) proteins were thoroughly dialyzed against buffer E just before use. After dialysis, the protein solution was concentrated to approximately 10% protein by ultrafiltration and then diluted to the desired protein concentration with the dialyzing buffer. The protein gel was prepared by heating for 30 min at 100°C by the micro-method of Utsumi et al.27

**Calcium precipitation behavior.** Protein solutions (0.5 mg/ml of buffer E) were added with calcium chloride to adjust the final concentration to 0–5 mM calcium, kept at room temperature for 20 min and centrifuged at 5000 rpm for 5 min. The protein of the supernatant after the centrifugation was measured, and the precipitation percent of proteins calculated at various calcium concentrations.

**Results**

**High-level expression of A1aB1b-3 in *E. coli***

We have reported the construction of expression plasmids based on the stepwise deletion of the cDNA region encoding the signal sequence and N-terminal regions of the preproglycinin A1aB1b.20 Among the various modified expression plasmids, the most efficient expression in *E. coli* was observed with pKGA1aB1b-3, and the amount of A1aB1b-3 ultimately produced in *E. coli* was approximately 1.5% of the total bacterial proteins.20 However, the expression level was not enough for use in protein engineering of glycinin. To increase the level of A1aB1b-3 expression, we tried to increase the copy number of plasmids using a runaway vector28 or the replication origin of pUC9.29 and to change the distance between the Shine-Dalgarno (SD) sequence and the initiation codon (ATG).30,31 However, we could not increase the level of A1aB1b-3 expression. Then we changed the culture conditions for *E. coli* JM105 harboring pKGA1aB1b-3. The culture temperature and shaking speed were critical factors for high-level expression. In our previous study,20 we cultured the cells at 30°C with a shaking speed of 175 strokes/min and observed the expression level of around 1.5% of the total bacterial proteins. However, the optimum temperature and shaking speed for high-level expression were 37°C and 85–90 strokes/min. We obtained an expression level corresponding to 20% of the total bacterial proteins as measured by densitometric scanning of lane 3 in Fig. 1. The concentration of proglycinin in extracts of induced bacteria was 39 μg per ml of culture. No expressed protein from the control *E. coli* cells harboring the expression vector pKK233-2 alone and only a little from non-induced cells harboring pKGA1aB1b-3 was detected (Fig. 1, lanes 1 and 2).

**Purification of A1aB1b-3 from *E. coli***

Table I summarizes the purification procedures that we found to be most useful for isolating A1aB1b-3 expressed in *E. coli*. A relatively simple and rapid method for purifying A1aB1b-3 was developed using cryoprecipitation, which is an intrinsic property of the mature glycinin.32 The most important aspect of the purification is that A1aB1b-3 was found almost exclusively in the soluble portion of the cell extract and, therefore, the use of strong denaturants was not required for the recovery at any step of purification of A1aB1b-3.

The details of the purification of A1aB1b-3 are given under Materials and Methods, but briefly, proteins in the extracts of JM105 cells harboring pKGA1aB1b-3 were first fractionated by precipitation with ammonium sulfate. A1aB1b-3 precipitated in the 40–65% ammonium sulfate fraction with a yield of 79% (Table I). The partially purified A1aB1b-3 was

---

**Note:** The text provided is a transcription from the original document, focusing on the expression and purification of proglycinin from *E. coli* as described in the paper. The content includes details on the methodology used for purification, along with observations on the expression levels and the impact of various culture conditions on the protein yield.
then dialyzed extensively against buffer B, and afterwards put on a Q-Sepharose column. Figure 2 shows the elution profile of *E. coli* proteins with a gradient of NaCl in buffer B, and also shows the location of \( A_{13}B_{13} \)-3 found by Coomassie Blue Staining SDS-PAGE analysis. The pooled fractions were cryoprecipitated. Figure 3 shows the SDS-PAGE analysis of the \( A_{13}B_{13} \)-3-containing fractions during the purification. The most purified form of \( A_{13}B_{13} \)-3 was found in the cryoprecipitated fraction (Fig. 3, lane 4). Judging from the gel scan (data not shown) where 5 \( \mu \)g of the purified \( A_{13}B_{13} \)-3 was put onto the gel, the purity was more than 97%. The purified product migrated at the expected molecular mass of 56 kDa on SDS-polyacrylamide gel (Fig. 3, lane 4). The overall yield of the final purified \( A_{13}B_{13} \)-3 was 36%.

The purified \( A_{13}B_{13} \)-3 predominantly occurred at trimers, judging from the sucrose density gradient centrifugation (Fig. 4), which

---

**Table 1. Purification of \( A_{13}B_{13} \)-3-Expressed in *E. coli***

<table>
<thead>
<tr>
<th>Step</th>
<th>Total protein (mg)</th>
<th>Total ( A_{13}B_{13} )-3 (mg)</th>
<th>Purity (%)</th>
<th>Yield (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Crude extract</td>
<td>2243</td>
<td>471</td>
<td>21</td>
<td>100</td>
</tr>
<tr>
<td>2. Ammonium sulfate fractionation</td>
<td>1438</td>
<td>373</td>
<td>26</td>
<td>79</td>
</tr>
<tr>
<td>3. Q-Sepharose pool</td>
<td>343</td>
<td>240</td>
<td>70</td>
<td>51</td>
</tr>
<tr>
<td>4. Cryoprecipitation</td>
<td>174</td>
<td>169</td>
<td>97</td>
<td>36</td>
</tr>
</tbody>
</table>

* The amount of \( A_{13}B_{13} \)-3 in each sample was measured by densitometric analysis following Coomassie Brilliant Blue staining of the SDS-PAGE.

---

**Fig. 1.** High-level Expression of \( A_{13}B_{13} \)-3. A 11% SDS-polyacrylamide gel is shown displaying the total proteins of *E. coli* strain JM105 growing at 37°C: Lane M, molecular weight markers; Lane 1, JM105 [pKK233-2]; Lane 2, non-induced JM105 [pKG\( A_{13}B_{13} \)-3]; Lane 3, induced JM105 [pKG\( A_{13}B_{13} \)-3]. The arrow indicates the position of \( A_{13}B_{13} \)-3. The numbers on the left denote molecular weights of the markers.

**Fig. 2.** Purification of \( A_{13}B_{13} \)-3 by a Q-Sepharose Column Chromatography. The protein content of each fraction was measured by the method of Bradford, and samples of selected fractions were analyzed by SDS-PAGE on 11% gels. The bar indicates the column fractions that were pooled.
Expression and Functionality of Proglycinin from *E. coli*

Fig. 3. SDS-PAGE Analysis of A₁₂B₁b-3 Samples at Different Stages of Purification.
Lane M, molecular weight markers; Lane 1, crude extract of induced JM105 [pKG₁₂B₁b-3] cells; Lane 2, fraction with 40–65% saturation of ammonium sulfate; Lane 3, Q-Sepharose pool; Lane 4, cryoprecipitate. The arrow indicates the position of A₁₂B₁b-3. The numbers on the left denote molecular weights of markers.

was observed in the endoplasmic reticulum of soybean. A small peak of around 11S was also observed. However, we do not know whether this was derived from hexamers or random aggregates.

**CD spectroscopy of A₁₂B₁b-3 expressed in E. coli**

To examine the potential conformational differences between A₁₂B₁b-3 expressed in *E. coli* and the native glycinin half-molecule (7S), we compared the far-UV circular dichroism spectra of the two proteins (Fig. 5). Analysis of the data for both A₁₂B₁b-3 and the native glycinin half-molecule (7S) using the Hennessey and Johnson algorithm indicated that the two proteins were similar in their secondary structure. The amount of secondary structure derived from the data in Fig. 5 was 17% α-helix, 26% β-sheet, 22% β-turns, and 33% random coil for A₁₂B₁b-3 and 20% α-helix, 24% β-sheet, 22% β-turns and 32% random coil for the native glycinin half-molecule (7S).

**Calcium precipitation properties of A₁₂B₁b-3 expressed in E. coli**

Calcium salt has frequently been used in the fractionation of native conglycinin and glycinin proteins based on their differential solubilities in calcium solutions. The precipitation behavior of A₁₂B₁b-3 expressed in *E. coli*...
Calcium Precipitation Properties of $A_{1a}B_{1b}$-3, the Native Glycinin, and the Glycinin Half-molecule ($\overline{7S}$).

Protein samples were in buffer E. $\bullet$, $A_{1a}B_{1b}$-3; $\triangle$, native glycinin; $\blacksquare$, glycinin half-molecule.

Discussion

We have enhanced the expression of glycinin $A_{1a}B_{1b}$ cDNA in *E. coli* JM105 by modification of the cultivation conditions, slow shaking at 37°C being necessary for high-level expression. We do not know the exact reason why the high-level expression of glycinin cDNA in *E. coli* is established in slow shaking, but it may be related to the amount of dissolved oxygen in the medium. Although the overall production of $A_{1a}B_{1b}$-3 in the *E. coli* cells appeared to increase, the amount of total *E. coli* proteins was somewhat low due to the slow shaking that decelerates the growth of the *E. coli* cells. However, the total amount of the expressed protein (39 mg/l) is enough to evaluate the functionalities of the expressed proteins, and the high-level expression (approximately 20% of the total *E. coli* proteins) leads to the simple purification procedure (Table I). Therefore, we concluded that the high-level expression system of glycinin cDNA was established.

There are many reports that highly expressed protein formed insoluble inclusion bodies in *E. coli* cells. Formation of unnatural disulfide bonds in expressed protein is partly responsible for the formation of inclusion bodies. Glycinin subunits have several disulfide bonds including an interchain one. Therefore, $A_{1a}B_{1b}$-3 expressed in *E. coli* may form inclusion bodies. However, $A_{1a}B_{1b}$-3 occurred as a soluble protein in *E. coli* cells and existed as a trimer in the cell extract and also after purification (Fig. 4) as observed in the

Gel properties of $A_{1a}B_{1b}$-3 expressed in *E. coli*

$A_{1a}B_{1b}$-3 expressed in *E. coli* formed a transparent gel at the low ionic strength ($\mu = 0.01$) used here, similar to those of the native glycinin and the glycinin half-molecule, although $A_{1a}B_{1b}$-3 required a higher protein concentration to form a gel than the native glycinin and the glycinin half-molecules ($\overline{7S}$) (Fig. 7).
endoplasmic reticulum of soybean.\textsuperscript{8,9} Moreover, \(A_{1\alpha}B_{1\beta}-3\) had a CD spectrum similar to that of the glycinin half-molecule (Fig. 5). These findings suggest that \(A_{1\alpha}B_{1\beta}-3\) expressed in \textit{E. coli} self-assembled to a trimer with a structure similar to that of the glycinin half-molecule.

\(A_{1\alpha}B_{1\beta}-3\) expressed \textit{E. coli} had properties similar to the intrinsic properties of the native glycinin, \textit{i.e.}, cryoprecipitation (Table I), calcium-induced precipitation (Fig. 6), and heat-induced gelation (Fig. 7), although slight differences in the behavior of calcium-induced precipitation and heat-induced gelation were observed between \(A_{1\alpha}B_{1\beta}-3\) and the native glycinin. This together with the fact that \(A_{1\alpha}B_{1\beta}-3\) assembled to a trimer with a structure similar to that of the glycinin half-molecule indicate that the \textit{E. coli} expression system of glycinin cDNA established here may be used for the evaluation of the self-assembly and the food qualities of protein-engineered soybean proteins.

The native glycinin and the glycinin half-molecule are a hetero hexamer and a hetero trimer, respectively, composed of various subunits, each of which is cleaved to the acidic and basic polypeptides.\textsuperscript{2–6} \(A_{1\alpha}B_{1\beta}-3\) expressed in \textit{E. coli} was not cleaved to the acidic and basic polypeptides and existed as a homo trimer (Fig. 4), not a hexamer, since the processing of the junction between the acidic and basic polypeptides is necessary for the assembly to a hexamer.\textsuperscript{11} Therefore, there are three structural differences between the native glycinin and \(A_{1\alpha}B_{1\beta}-3\): \(\text{1} \) hetero or homo, \(\text{2} \) presence or absence of the processed junction, \(\text{3} \) a hexamer or a trimer. The differences in the behavior of calcium-induced precipitation and heat-induced gelation may be caused by these structural differences. Therefore, the processing of the junction between the acidic and basic polypeptide is necessary for the precise evaluation of food qualities of protein-engineered glycinins. We are now purifying the enzyme processing the junction to investigate these problems.

\textbf{Acknowledgments.} This research was supported by Grants-in-Aid from Asahi Breweries Foundation, The Foundation for Promotion of Food Science and Technology, the Ministry of Education, Science and Culture of Japan (to S.U.).

\textbf{References}