Modification Tolerability of Soybean Proglinycin

Andrew B. Gidamis,1 Patrick Wright,2 Zahur U. Haque,3 Tomoyuki Katsube,4 Makoto Kito, and Shigeru Utsumi*

Research Institute for Food Science, Kyoto University, Uji, Kyoto 611, Japan
Received February 27, 1995

Proglincins containing multiple modifications, that is, with two kinds of single modifications, designed previously to improve food functions, were assessed for ability to form proper conformations. This was done to discover the modification tolerability of proglinycin. Based on our results, three new criteria are proposed for judging formation of proper conformations. Our results suggest that proglinycin molecules tolerate different combinations of single modifications without misfolding.

Glycinin (11S globulin) is one of the dominant storage proteins of soybean (Glycine max L.) seeds. Each constituent subunit of glycinin is synthesized as a preproglinycin.1-3 The signal sequence is removed cotranslationally in the endoplasmic reticulum (ER) and the resultant proglinycin subunits assemble into trimers of about 85,2-4,6 These complexes are targeted from the ER to the vacuoles, where a specific posttranslational cleavage occurs,5 where they form protein bodies. The cleavage results in mature subunits, each of which consists of an acidic and a basic polypeptide, and they assemble into hexamers of about 125.2-4,6 Finally, glycins accumulate in a highly packed state in protein bodies. Molecular assembly, targeting from the ER to the vacuoles and accumulation in protein bodies depend on topogenic information contained in the glycinin molecule. Therefore, it is essential to find what kinds of modification glycinin molecules can tolerate without misfolding when creating novel soybean plants that can produce modified glycinins having improved food functions.6

Prior to this study, we attempted to create modified proglucins designed to improve their nutritional and/or functional properties by protein engineering using an Escherichia coli expression system for models of modified glycins.7,8 The modifications were based on structure-function relationships7,8 and genetically variable regions found by comparison of amino acid sequences of glycins-type globulins from various legumes and nonlegumes9-10 (Fig. A). Three kinds of modifications were introduced: (1) deletion of each variable region9; (2) insertion of a hydrophobic oligopeptide (tetramethionine) into hydrophilic regions7; (3) deletion of disulfide bond(s).8 Among them, modified proglucins dI, dV8, IV + 4Met, V + 4Met, Gly12Ser88, and Gly12Ser88 shown in Fig. B could form proper conformations similar to that of native proglucin.7,8 This conclusion was based on the following three criteria: (1) high-level expression (≥10% of E. coli total proteins), (2) solubilities comparable to that of globulins; and (3) self-assembly into trimers.7,8 This conclusion was further supported by the fact that these modified proglucins could form crystals11,12 as the normal recombinant proglucin did.13

In this study, we have assessed multiply modified proglucins with two kinds of single modifications designed as described above for ability to form trimeric structure and for structural stability to assess the modification tolerability of soybean proglucin.

Expression plasmids for the fifteen multiply modified proglucins shown in Fig. C were constructed using the BamHI site in the vector region (pKK233-2) and unique restriction sites (XmaI, EcoT22I, DraIII) between the modification sites in the A16,11β proglucin cDNA region of individual expression plasmids constructed previously.7,8 Nucleotide sequences in the vicinity of the translation initiation site and promoter of each expression plasmid constructed here were either the same as that of pKGA116,11β or pKGA11β.n-11 The expression level of pKGA116,11β and pKGA11β.n-11 were 20% of the E. coli total proteins.7 In this respect, the efficiency of transcription of individual expression plasmids and that of translation of transcripts are expected to be similar. In other words, if the expression level is low, it may be due to some disturbance in the foldings caused by the introduced modifications. Expression levels of proglucins having a single modification are summarized in Fig. B.

E. coli cells (strain JM105) harboring individual expression plasmids were cultured at 37°C as described.12 Cells harvested by centrifugation were disrupted in extraction buffer (35 mM potassium phosphate, pH 7.6, 0.4 M NaCl, 1 mM EDTA, 1.5 mM PMSF) by sonication, and cell debris and supernatant were fractionated by centrifugation.12 Aliquots of total cells, cell debris, and supernatant were dissolved in SDS-sample buffer (62.5 mM Tris-HCl, pH 6.8, 2% SDS, 30% glycerol, 0.2% 2-mercaptoethanol) and analyzed by SDS-PAGE13 to measure the expression level and solubility of the expressed proteins. The amount of the expressed proteins was measured by densitometric analysis following Coomassie brilliant blue staining. Self-assembly into trimers of the expressed proteins was analyzed by sucrose density gradient centrifugation using 2S, 7S, and 11S fractions purified from soybean seeds as size markers as described previously.14

Expression level, solubility, and ability for self-assembly of the multiply modified proglucins are summarized in Table. Most (>90%) of all the multiply modified proglucins were in the soluble fractions. They mostly sedimented in fractions corresponding to the size of trimers. However, the expression level was variable among them. The modified proglucins with the combinations between dI, dV8, IV + 4Met, Gly12Ser88, and Gly12Ser88 were accumulated in the cells at the level of ≥15% of the total bacterial proteins. This suggests that the modifications introduced into dI, dV8, IV + 4Met, Gly12Ser88, and Gly12Ser88 do not disturb proper folding. On the other hand, any modified proglucins with the combination of V + 4Met with dI, IV + 4Met,
Fig. Schematic Representation of Proglicynin $A_{15}B_{15}$, Subunit, Normal Recombinant Proglicynin $A_{15}B_{15}$-3, and Modified Proglicynins.

(A) The variable and conserved regions of proglucin A$_{15}B_{15}$ subunit as aligned by Wright. The numbers of the residues from N-terminus are described for the variable regions above the alignment. The five variable regions were labeled I-V. Numbers below the figure indicate positions of the cysteine residues that form disulfide bonds. Shaded and open areas are variable and conserved regions, respectively. (B) Schematic representation of normal recombinant proglucin A$_{15}B_{15}$-3 and modified proglucynins with a single modification. A$_{15}B_{15}$-3 lacks N-terminal three amino acids. $\Delta I$, N-terminal 11 residues deleted; $\Delta V8$, C-terminal 8 residues deleted; IV$_V^+$+4Met, oligopeptide Arg Met Met Met Met Met Gly inserted between Pro281 and Arg282 in the variable region IV; V$_V^+$+4Met, oligopeptide Glu Met Met Met Met Met His inserted between Pro467 and Gly468 in the variable region V; Gly12, Cys12 replaced with glycine; Ser88, Cys88 replaced with serine; Gly12Ser88, cysteines 12 and 88 replaced with glycine and serine, respectively. Expression plasmids for these modified proglucynins except for $\Delta I$ were constructed using pKGA$_{15}B_{15}$-3 which expresses A$_{15}B_{15}$-3. For details, see references 7, 8. N-Terminal methionine was retained in all the proglucynins except for $\Delta I$. $\Delta V8$ had two extra amino acids Leu-Leu at its C-terminal. The numbers on the right of individual modified proglucynins with a single modification are the expression levels in E. coli. (C) Schematic representation of modified proglucynins with the combination of the modifications designed in (B). In case of disulfide bonds, only the modified one is shown. G, S, and C indicate glycine, serine, and cysteine residues, respectively.

Table Expression Level of Multiply Modified Proglucynins in Escherichia coli and Properties of Expressed Proteins

<table>
<thead>
<tr>
<th>Name of modified proglucynins</th>
<th>Expression level (%)</th>
<th>Solubility</th>
<th>Assembly</th>
<th>Stability</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta I$ + $\Delta V8$</td>
<td>20</td>
<td>Soluble</td>
<td>Trimer</td>
<td>Stable</td>
</tr>
<tr>
<td>$\Delta I$ + (IV$_V^+$+4Met)</td>
<td>20</td>
<td>Soluble</td>
<td>Trimer</td>
<td>Stable</td>
</tr>
<tr>
<td>$\Delta I$ + (IV$_V^+$+4Met)</td>
<td>3</td>
<td>Soluble</td>
<td>Trimer</td>
<td>Stable</td>
</tr>
<tr>
<td>$\Delta I$ + Ser88</td>
<td>20</td>
<td>Soluble</td>
<td>Trimer</td>
<td>Stable</td>
</tr>
<tr>
<td>(IV$_V^+$+4Met) + $\Delta V8$</td>
<td>15</td>
<td>Soluble</td>
<td>Trimer</td>
<td>Stable</td>
</tr>
<tr>
<td>(IV$_V^+$+4Met) + (V$_V^+$+4Met)</td>
<td>2</td>
<td>Soluble</td>
<td>Trimer</td>
<td>Stable</td>
</tr>
<tr>
<td>Gly12 + $\Delta V8$</td>
<td>20</td>
<td>Soluble</td>
<td>Trimer</td>
<td>Stable</td>
</tr>
<tr>
<td>Gly12 + (IV$_V^+$+4Met)</td>
<td>20</td>
<td>Soluble</td>
<td>Trimer</td>
<td>Stable</td>
</tr>
<tr>
<td>Gly12 + (IV$_V^+$+4Met)</td>
<td>5</td>
<td>Soluble</td>
<td>Trimer</td>
<td>Stable</td>
</tr>
<tr>
<td>Ser88 + $\Delta V8$</td>
<td>20</td>
<td>Soluble</td>
<td>Trimer</td>
<td>Stable</td>
</tr>
<tr>
<td>Ser88 + (IV$_V^+$+4Met)</td>
<td>20</td>
<td>Soluble</td>
<td>Trimer</td>
<td>Stable</td>
</tr>
<tr>
<td>Ser88 + (V$_V^+$+4Met)</td>
<td>5</td>
<td>Soluble</td>
<td>Trimer</td>
<td>Stable</td>
</tr>
<tr>
<td>Gly12Ser88 + $\Delta V8$</td>
<td>20</td>
<td>Soluble</td>
<td>Trimer</td>
<td>Stable</td>
</tr>
<tr>
<td>Gly12Ser88 + (IV$_V^+$+4Met)</td>
<td>20</td>
<td>Soluble</td>
<td>Trimer</td>
<td>Stable</td>
</tr>
<tr>
<td>Gly12Ser88 + (V$_V^+$+4Met)</td>
<td>5</td>
<td>Soluble</td>
<td>Trimer</td>
<td>Stable</td>
</tr>
</tbody>
</table>

* % of total E. coli proteins.

Gly12, Ser88, and Gly12Ser88 were accumulated at low levels (≤5%), although the modified proglucynin V$_V^+$+4Met itself was at the level of 10% (Fig. B). This indicates that the modification which was introduced into the fifth variable region interfered leniently with the folding of V$_V^+$+4Met and fairly in the case of modified proglucynins having the combination of V$_V^+$+4Met with other modifications.

The accumulation level of V$_V^+$+4Met in E. coli is lower than that of unmodified protein (Fig. B). However, V$_V^+$+4Met accumulated at a similar level compared to normal glucynin in the seeds of tobacco and in potato tubers. Moreover, a modified proglucynin DI$_V^+$+4Met, which is degraded completely in E. coli cells, can be accumulated at high levels in yeast cells. The normal and modified proteins expressed in plant and yeast cells move co-translationally into the ER and target from the ER to the vacuoles, since the cDNAs encoding preproformos are introduced. These results show that the modified proteins expressed in plant and yeast cells are not recognized as misfolded proteins by the quality control system in the E.R, and are not proteolyzed by the degradation system in the vacuoles. The modified proglucynins are stable after formation of the proper conformation.

To examine the stability of the multiply modified proglucynins, soluble fractions of E. coli cells harboring individual expression plasmids were dialyzed against extraction buffer without EDTA and PMSF for 42h at room temperature. Each dialyze was analyzed by SDS-PAGE and the proteins were detected by Coomassie brilliant blue staining. The intensities of the bands of all the multiply modified proglucynins were almost equal to those before dialysis. In other words, they were stable during dialysis.
Modification Tolerability of Proglycinin

(1595)

We concluded from the results obtained here that the criteria for judging formation of proper conformation should be as follows: (1) solubility should be comparable with that of globulins, (2) there must be self-assembly into trimers, and (3) conformation should be stable.

All single modifications to improve food functions of soybean glycinin, that followed these three criteria, could be combined without misfolding.

Acknowledgment. This research was supported in part by Grants-in-Aid for Scientific Research from the Ministry of Education, Science, and Culture of Japan (to S.U.).

References