DOI:10.1021/jf8034258

JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY
ARTICLE

Production of a Bioactive Peptide (IIAEK) in *Escherichia coli* Using Soybean Proglycinin A1aB1b as a Carrier

KRISNA PRAK AND SHIGERU UTSUMI*

Laboratory of Food Quality Design and Development, Graduate School of Agriculture, Kyoto University, Uji, Kyoto 611-0011, Japan

To produce large amounts of a peptide of fewer than 10 amino acid residues, construction of a gene encoding multimers of the small peptide is necessary. For this study a method was developed to facilitate the gene construction of high multimers of a small peptide with one step of cloning. A hypocholesterolemic peptide, IIAEK, from cow's milk β-lactoglobulin was used as a model peptide for the construction of a gene encoding multimers of IIAEK and for the production of the peptide. Two systems for direct expression of 28-mers of IIAEK sequences (28IIAEK) and expression of 34 IIAEK sequences (4 IIAEK sequences in each of the disordered regions I, II, and III and 14 and 8 IIAEK sequences in disordered regions IV and V, respectively) in a mutant of soybean proglycinin A1aB1b lacking 31 residues in disordered region IV [A1aB1b(Δ31)-34IIAEK] were used. The protein produced from both systems formed inclusion bodies. The expression level of A1aB1b(Δ31)-34IIAEK was 29.9% of the total cell proteins and that of the 28IIAEK was 2.0%. The insoluble A1aB1b(Δ31)-34IIAEK was digested by trypsin without any help from urea or chemicals, and the produced IIAEK was purified using an octadecyl silica column. The yield of IIAEK was 58.6%. The results showed that A1aB1b as a carrier of multiple peptides and use of an *Escherichia coli* expression system are suitable for production of bioactive peptide.

KEYWORDS: IIAEK; peptide; multimers; proglycinin; soybean

INTRODUCTION

For the production of recombinant proteins, various expression systems using prokaryotic cells, yeast cells, or mammalian cells have been established. To achieve a high-level expression of a target protein, *Escherichia coli* is most commonly employed as host cell for an expression system because of its distinct and well-studied genetic background, abundant available plasmids, and ease of gene manipulation. Various studies in large-scale production using the *E. coli* expression system of several small peptides that are composed 21−50 amino acid residues such as buforin II (*1*), dermcidin (*2*), human β-defensin-2 (*3*, *4*), an analogue of human β-defensin-2 (*5*), human β-defensin-3 (*6*), and human β-defensin-4 (*7*) have been reported. Moreover, production of the peptides using the *E. coli* expression system is more environmentally friendly and less expensive compared to chemical synthesis.

However, there has been no report on the large-scale production of a small peptide of fewer than 10 amino acid residues. This may be due to difficulties in the detection or low expression when a fusion protein was not used as an expression partner (*1*) or the tedious and laborious methods in the construction of a gene of high tandem multimers for large-scale production of the peptide (*4*).

*Author to whom correspondence should be addressed (telephone 81-774-38-3760; fax 81-774-38-3761; e-mail sutsumi@kais.kyoto-u.ac.jp).

We hereby report a method of gene construction of high multimers of a small peptide using short single-stranded DNAs encoding for four multimers of IIAEK sequences, requiring only one step of cloning. IIAEK is a hypocholesterolemic peptide, derived from cow's milk β-lactoglobulin, which has greater hypocholesterolemic activity than that of a medicine, β-sitosterol (*8*). Previously, we introduced 8 and 20 of the peptides into soybean proglycinin A1aB1b (*9*): A1aB1b-containing 8 IIAEK sequences was expressed as a soluble protein, but A1aB1b containing 20 IIAEK sequences formed inclusion bodies in *E. coli*. It is expected that the peptides can be released from the inclusion body because the inclusion body formed by A1aB1b containing a peptide sequence (LPLPR) could be digested by trypsin and chymotrypsin (*10*). If we can get IIAEK in a good yield from mutant A1aB1b containing a large number of IIAEK sequences produced in *E. coli*, this system will be suitable for IIAEK production.

In this study, we (a) introduced an additional 14 IIAEKs to the A1aB1b containing 20 IIAEK sequences and (b) constructed a direct expression plasmid for 28 tandem multimers of IIAEKs. This paper reports the comparison of the two systems in terms of production yields, digestion of the production from the former system, and purification of IIAEK.

MATERIALS AND METHODS

Bacterial Strain, Medium, and Plasmid. *E. coli* strains HMS174(DE3), AD494(DE3) and Origami(DE3) (Novagen)

 Published on Web 3/19/2009

were used as hosts for expression, and the culture medium was Luria−Bertani (LB) containing 0.17 or 0.5 M NaCl, pH 7.5. The expression plasmid for a mutant A1aB1b containing 34 IIAEK sequences [pEA1aB1b(Δ31)-34IIAEK] was constructed from the previously constructed plasmids pEA1aB1b-IV-4IIAEK [A1aB1b containing 4 IIAEK sequences in the disordered region IV (IV-L-4), pEA1aB1b-V-8IIAEK (A1aB1b containing 8 IIAEK sequences in the disordered region V] and pEA1aB1b-20IIAEK (A1aB1b containing 20 IIAEK sequences, where 4 IIAEK sequences are present in each of the five disordered regions I, II, III, IV, and V) (*9*). All of the DNAs encoding the mutated A1aB1b were placed between the *Nco*I and *Bam*HI sites of the expression vector pET-21d. For the construction of the direct expression plasmid, pBluescript II KS (−) (Stratagene), pET-21d (Novagen), and pEA1aB1bIV-14 IIAEK were used.

Experimental Procedures. *Preparation of Double-Stranded DNA Encoding Multiple IIAEKs*. Double-stranded DNA (dsDNA) containing large numbers of multimers of IIAEK was produced by annealing a single-stranded DNA1 (ssDNA1) and a single-stranded DNA2 (ssDNA2). SsDNA1 is 5′-(ATC ATC GCC GAG AAA)$_4$-3′, and ssDNA2 is ssDNA1 but encoding for 3′−5′ (**Figure 1**). Before annealing, the ssDNA1 was phosphorylated using T4 polynucleotide kinase (Biolabs). The annealing was done at 67 °C for 15 min after heating at 95 °C for 4 min, followed by slow cooling to room temperature. Then the ligation of the annealed products was done using Takara ligation kit ver. 2.1 (Takara). Because only one side of the dsDNA was phosphorylated at the 5′-end, the ligation was expected to be made orderly for the dsDNAs and give a new dsDNAs as long as possible. The ligation products were separated and purified by agarose gel according to their expected sizes of 4, 8, 12, and 16 multimers of IIAEK genes (**Figure 2A**).

Construction of Expression Plasmid pEA1aB1b(Δ31)-14 IIAEK and pEA1aB1b(Δ31)-34 IIAEK. Four types of dsDNAs encoding the expected sizes of 4, 8, 12, and 16 multimers of IIAEK sequences (**Figure 2B**) were selected for insertion into A1aB1b-4IIAEK (**Figure 3**). To create an insertion site in the middle position of the DNA region encoding four IIAEK in pEA1aB1b-IV-4IIAEK (A1aB1b-IV-4IIAEK contains 4IIAEK in the disordered region IV), one pair of primers was used: 5′-**ATT ATT GCG GAA AAA ATT ATT GCG GAA AAA** AGC AGA AGA AAT GGC ATT GAC GAG-3′ and 5′-**TTT CTC GGC GAT GAT TTT CTC GGC GAT GAT** TTT GCT TTG GCT TCC TCG GGG GCG-3′ (the bold letters indicate the sequences encoding two IIAEK sequences). Thirty cycles of PCR were performed using Pyrobest (Takara) at 94 °C for 15 s, 64 °C for 30 s, and 72 °C for 7 min. The resultant linear plasmid was then ligated with each type of dsDNA to get expression plasmids for mutant A1aB1b containing many IIAEK sequences in the disordered region IV. By this method expression plasmid pEA1aB1b-IV-14IIAEK containing DNA encoding 14 IIAEK sequences in the disordered region IV was constructed. To reduce the total size of the new modified version of A1aB1b, 31 amino acids from the start of disordered region IV were deleted by using primers 5′-CGC CCC CGA GGA AGC CAA AGC AAA A-3′ and 5′-TGG TTT TAT CAC GCT CAG ACC TCC-3′. Thirty cycles of PCR were performed using Pyrobest (Takara) at 94 °C for 15 s, 57 °C for 30 s, and 68 °C for 7 min. The PCR product was phosphorylated and ligated to create the expression plasmid pEA1aB1b(Δ31)-IV-14IIAEK (**Figure 4**).

To construct the expression plasmid pEA1aB1b(Δ31)-34IIAEK, the DNA regions encoding the disordered regions IV-4IIAEK and V-4IIAEK of pEA1aB1b-20IIAEK were substituted with the DNA encoding the disordered regions IV (Δ31)-14IIAEK and V-8IIAEK of pEA1aB1b-IV(Δ31)-14IIAEK and pEA1aB1b-V-8IIAEK, respectively. These substitutions were done successfully using restriction enzyme sites *Nsi*I and *Bsu*36I and *Bsu*36I and *Xba*I for the disordered regions IV and V, respectively (**Figure 4**).

Construction of Direct Expression Plasmid pE28IIAEK. The expression plasmid pEA1aB1b-IV-14IIAEK was used for the construction of the expression plasmid pE28IIAEK for direct expression of 28 IIAEK sequences (**Figure 5**). Two pairs of primers were used. The first pair of primers was 5′-ACC ACG

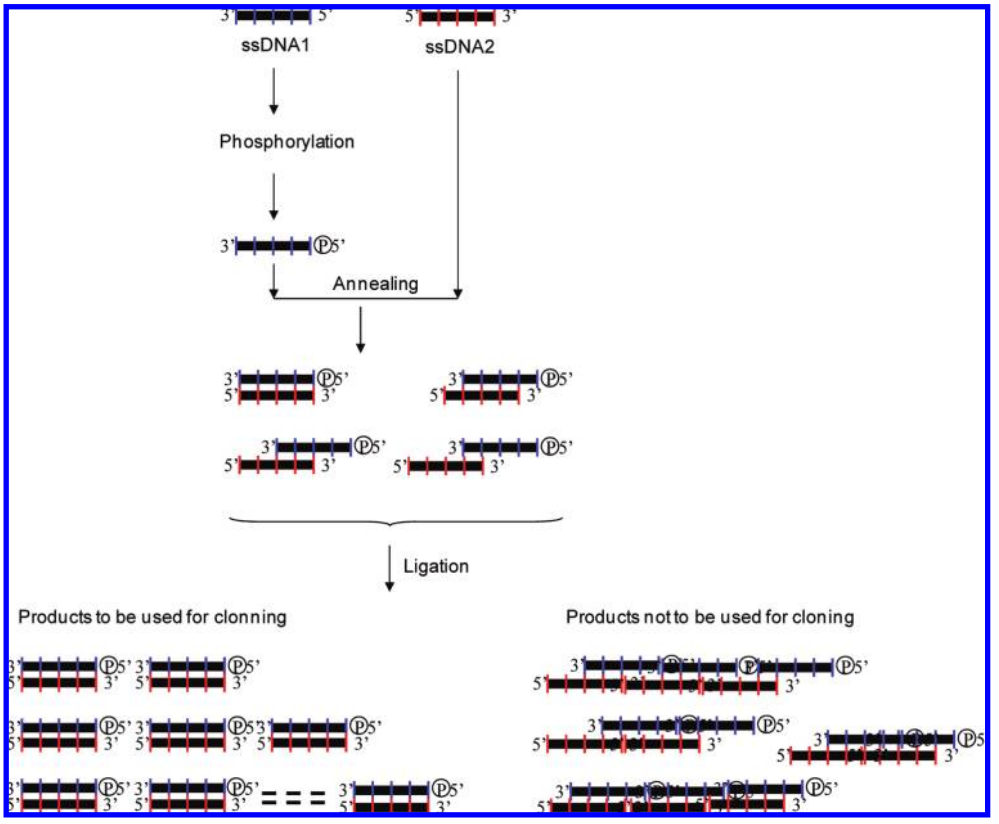

Figure 1. Production of double-stranded DNA containing high multimers of IIAEK sequences. ssDNA1 and ssDNA2 are the single-stranded DNA containing four multimers of IIAEK sequences for 5′−3′ and 3′−5′, respectively.

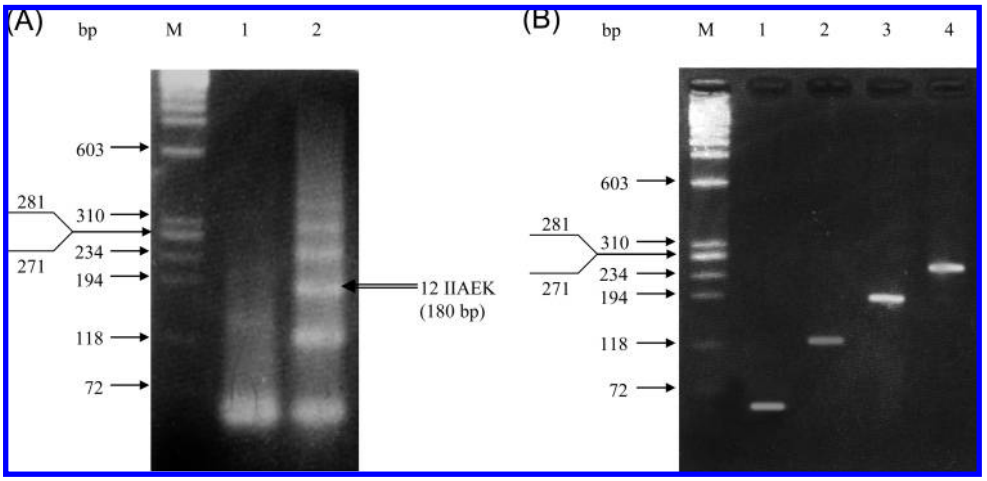

Figure 2. Electrophoretic analysis of the multimers of IIAEK genes. (**A**) Electrophoretic analysis of the multimers of IIAEK genes before purification. Lanes M, 1, and 2 indicate size markers in base pairs (bp) and annealing of single-stranded DNA1 (for 5′−3′) and single-stranded DNA2 (for 3′−5′), each encoding for four mulimers of IIAEK genes and product of lane 1 after ligation, respectively. (**B**) Multimers of IIAEK genes after purification. Lanes 1, 2, 3, and 4 correspond to 4, 8, 12, and 16 multimers of IIAEK genes, respectively.

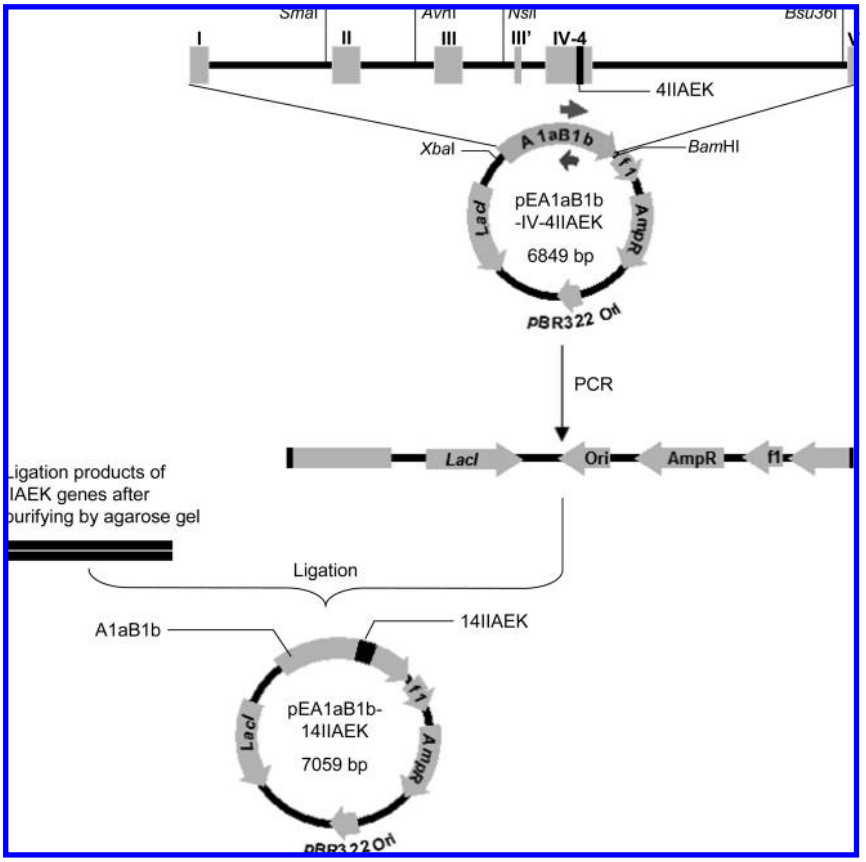

Figure 3. Schematic representation of the construction of expression plasmid pEA1aB1b-IV-14IIAEK. Roman numerals I, II, III, III′, IV-4, and V on the gray boxes indicate the sites of six disordered regions of A1aB1b-IV-4IIAEK. Restriction sites shown are unique sites of pEA1AB1b-IV-4IIAEK. IV-4 here refers to IV-L-4 in previous work (*9*).

CCC CCG AGG AAG **CCA TGG** GCA AAA T-3′ and 5′-CGC TAT CTG CTT GTC CAC GCT GAA TGC ATG TT-3′ (the bold letters indicate an *Nco*I site) to create an *Nco*I site at the start of the first 14 IIAEK sequences (14IIAEK). The second pair of primers was 5′-CAC GCC CCC GAG GAA GCC A**AA GCT T**TA AAA TC-3′ and 5′-CGC TAT CTG CTT GTC CAC GCT GAA TGC ATG TT-3′ (the bold letters indicate a *Hin*dIII site) to create a *Hin*dIII site at the start of the second 14 IIAEK sequences (14IIAEK). Thirty cycles of PCRs were named PCR1.1 and PCR2.1 for the first and second 14IIAEK, respectively, and were performed using Pyrobest (Takara) at 94 °C for 15 s, 62 °C for PCR1.1 and 61 °C for PCR 2.1 for 30 s, and 68 °C for 7 min (**Figure 5**). After amplification by PCR, each PCR product was phosphorylated and ligated, and the resultant new plasmids containing individual new restriction sites were used as templates for the second amplification. The next two pairs of primers were used for the creation of a *Hin*dIII site and a stop codon and a *Sac*I site at the end of the first 14IIAEK fragment and at the end of the second 14IIAEK fragment, respectively. The second amplification was made by PCRs named PCR1.2 and PCR2.2, using Pyrobest (Takara) at 94 °C for 15 s, 57 °C for 30 s, and 68 °C for 7 min, and 94 °C for 15 s, 55 °C for 30 s, and

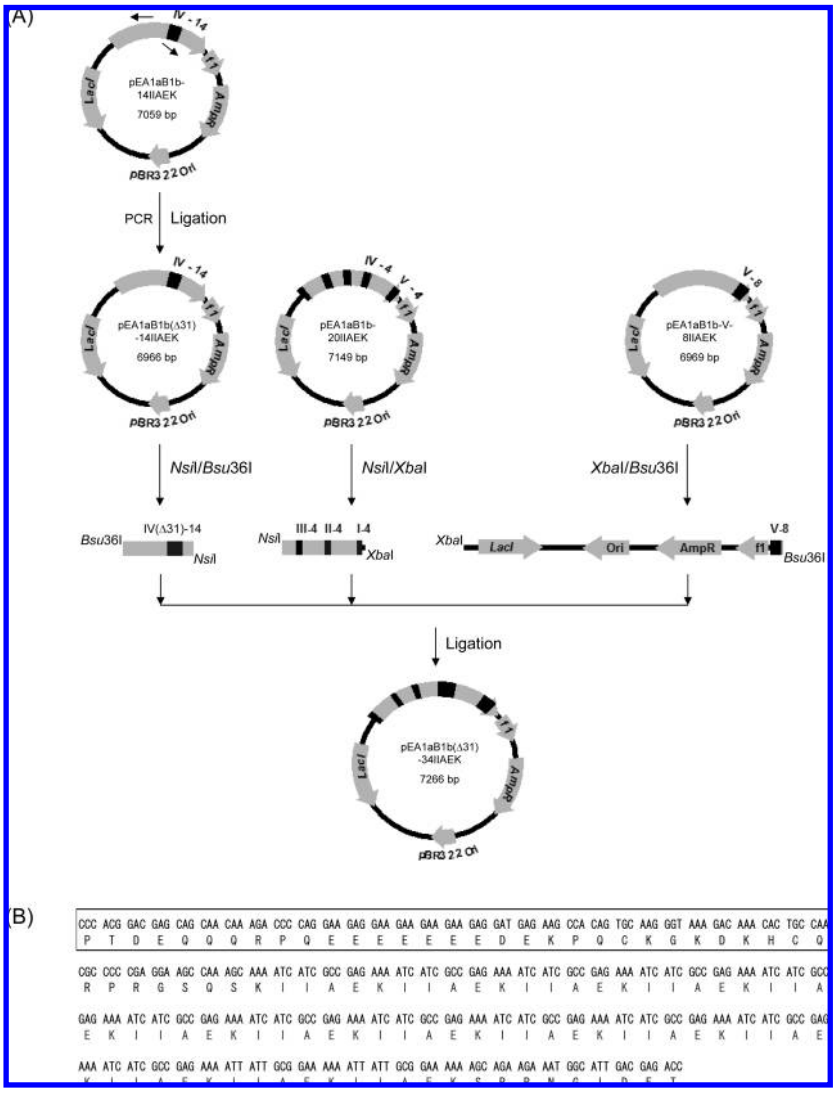

Figure 4. Schematic representation of the gene deletion and construction of expression plasmid pEA1aB1b(Δ31)-34IIAEK. (**A**) Schematic representation of the construction of expression plasmid pEA1aB1b(Δ31)-34IIAEK. (**B**) Deletion of 31 amino acid residues from disordered region IV as shown in the box. Details of the construction are described under Materials and Methods.

68 °C for 2 min, respectively. Two pairs of primers, 5′-AGA CTT CAT CAC CTG ACA TCT ACA ACC CTC A-3′ and 5′-GTC TCG TCA ATG CCA TTT CTT **AAG CTT** TTT TC-3′ (the bold letters indicate a *Hin*dIII site) and 5′-AAC GAG TGC CAG ATC CAA AAA CTC-3′ and 5′-CTC GTC AAT GCC ATT T**GA GCT C***CT A*TT TTT C-3′ (the bold and italic letters indicate a *Sac*I site and a stop codon, respectively) were used for PCR1.2 and PCR2.2, respectively. The product from PCR1.2 was phosphorylated and self-ligated. The resultant plasmid was digested with *Nco*I and *Hin*dIII. The short DNA fragment 1 encodes 14 IIAEK sequences with the two restriction sites *Nco*I and *Hin*dIII at the start and end of the fragment, respectively. The product from PCR2.2 was digested with *Hin*dIII and *Sac*I and was inserted between *Hin*dIII and *Sac*I sites of pBluescript II KS (−) (Stratagene). The plasmid with the insertion was digested with *Hin*dIII and *Sac*I to give fragment 2. The two DNA fragments, fragments 1 and 2, were self-ligated and inserted into *Nco*I and *Sac*I sites of pET-21d (Novagen) to create pE28IIAEK with 28 IIAEK sequences.

Expression and Detection of A1aB1b(Δ31)-14IIAEK, A1aB1b(Δ31)-34IIAEK, and 28IIAEK in E. coli. The expression plasmids for A1aB1b(Δ31)-14IIAEK, A1aB1b(Δ31)-34IIAEK, and 28IIAEK were transformed into *E. coli* expression hosts HMS174(DE3), AD494(DE3), and Origami(DE3). Forty microliters of full-grown cultures was transferred to the fresh 4 mL LB medium containing 0.17 or 0.5 M NaCl and cultured at 37 °C. At A_{600} = 0.3−0.6, isopropyl-1-thio-β-D-galactopyranoside (IPTG) was added to a final concentration of 1 mM to induce expression of the recombinant proteins. Following induction, cultivation was continued at 20 or 37 °C for 40 h. The cells were harvested by centrifugation and stored at −20 °C until used. Frozen cells were resuspended (15.0 mg of cells/mL of buffer) in buffer A [35 mM potassium phosphate, pH 7.6, 0.5 M NaCl, 10 mM 2-mercaptoethanol, 1 mM EDTA, 0.1 mM *p*-amidinophenylmethanesulfonyl fluoride (APMSF), 1 μg/mL leupeptin, 1 μg/mL pepstatine A, and 0.02% NaN_3], and thawed cells were disrupted by sonication on ice. The insoluble materials were separated from the soluble fraction by centrifugation at 17800*g* for 5 min. The total cell extract and soluble and insoluble fractions were analyzed by SDS−PAGE (*11*) using 11% acrylamide (*12*) (**Figure 6A**).

Expression of A1aB1b(Δ31)-34IIAEK and 28IIAEK was confirmed by Western blotting (*13*) using rabbit antiserum against IIAEKIIAEKIIAEK (**Figure 6B**). To decrease nonspecific reaction, anti-IIAEK serum was incubated with the cell extract of *E. coli* AD494(DE3) harboring plasmid pET-21d at 1:1 (v/v) ratio at 4 °C for overnight. The cell extract was obtained by sonication of 15.0 mg of cell in 1 mL of buffer A. After incubation, 1 mL of the mixture containing anti-IIAEK serum and cell extract was centrifuged at 17800*g* for 10 min at 4 °C to separate the supernatant and precipitate. The highly specific anti-IIAEK serum remaining in the supernatant was diluted

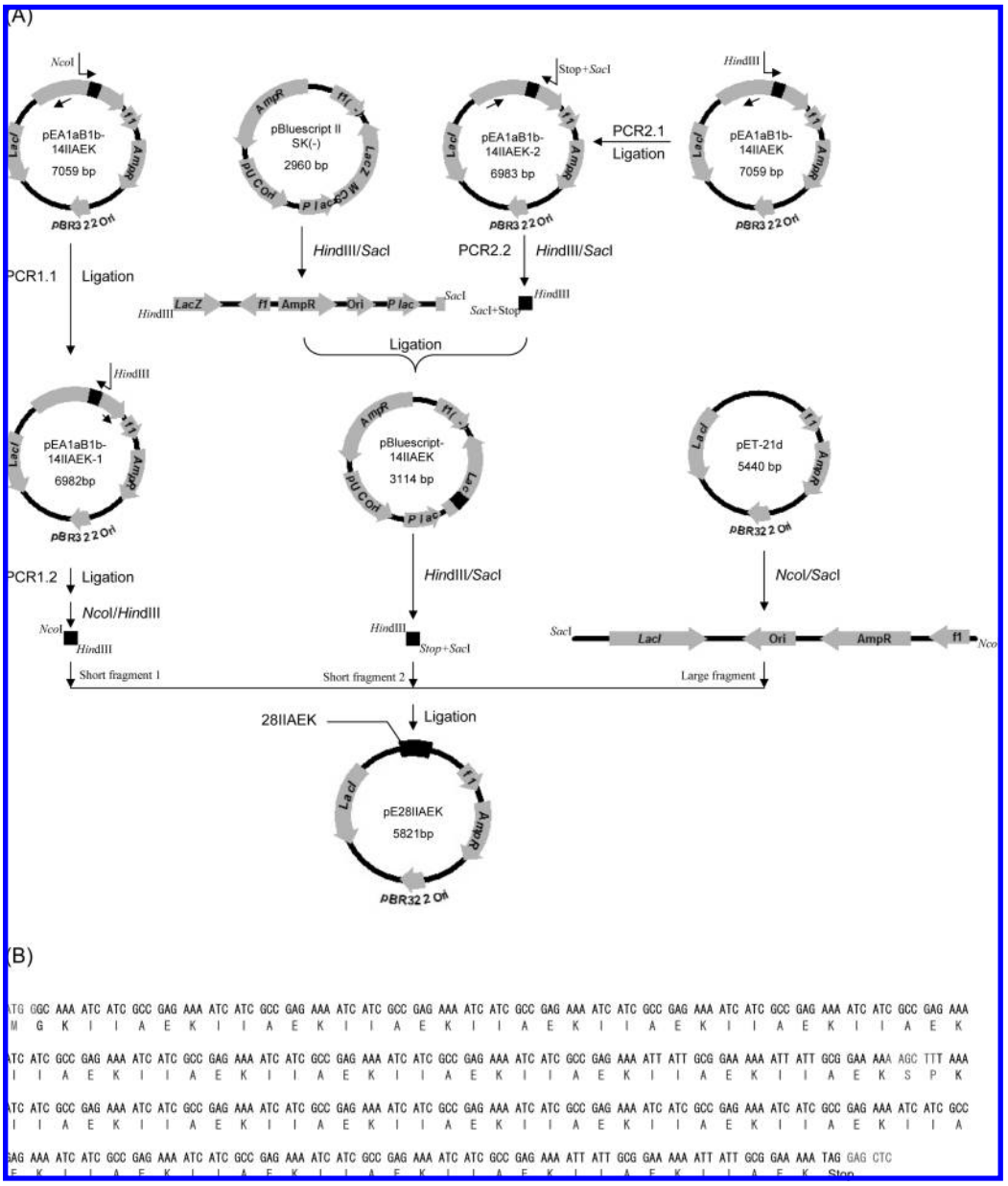

Figure 5. (**A**) Schematic representation of the construction of expression plasmid pE28IIAEK. Details of the construction are described under Materials and Methods. (**B**) Amino acid and nucleotide sequences of DNA encode 28IIAEK.

50 times and was used as antibody followed by goat anti-rabbit IgG−alkaline phosphatase conjugate (Promega) for Western blotting.

Analysis of Protein Expression Level in E. coli and Purity of Protein. The protein expression level and the purity of the protein samples were analyzed by densitometric scan of SDS−PAGE gel followed by estimation using ImageMaster 1D Elite, version 3.0 (Amersham-Pharmachia Biotech, Uppsala, Sweden).

Digestion of A1aB1b(Δ31)-34IIAEK and Purification of IIAEK Peptide. All procedures were carried out at 4 °C, and centrifugations were at 9100*g* for 20 min unless otherwise stated.

The *E. coli* cells (11.0 g) containing recombinant A1aB1b (Δ31)-34IIAEK were suspended in buffer A (40.0 mg of cells/mL of buffer) and disrupted by sonication. Insoluble materials were separated from the soluble fraction by centrifugation. The insoluble fraction containing the expressed protein was washed twice with buffer A and three times with Milli-Q water. The precipitate from 1 mL of the extract was dissolved in 200 μL of SDS buffer [62.5 mM Tris-HCl (pH 6.8), 2% SDS, 10% glycerol] and boiled for 5 min, and then the amount of protein was measured using a protein assay kit (BCA Protein Assay, Pierce, Rockford, IL) with bovine serum albumin as a standard. The pH of the insoluble fraction in Milli-Q water (4 mg/mL) was adjusted to 8.0 with 1 N NaOH, and the proteins in the fraction were digested with trypsin (E/S = 1/100) (w/w) for 12 h at 37 °C. The pH was checked and adjusted every hour in the first 5 h of reaction. The reaction was stopped by adding trifluoroacetic acid (TFA) to a final concentration of 0.1% and boiling for 10 min. The soluble digested products were separated from the undigested samples by centrifugation at 17800*g* and 4 °C for 30 min and stored at −20 °C until used.

IIAEK peptides were purified from the digested products by reverse phase HPLC on an octadecyl silica (ODS) column (Cosmosil 5C18-AR-II, 20 mm × 250 mm, Nakalai Tesque). For the first purification, the digested product was applied onto the column and the column was washed with 1% acetonitrile containing 0.1% TFA for 25 min at the flow rate of 9 mL/min (data not shown). The HPLC was programmed to have the following elution: 1% acetonitrile for 10 min, then a linear gradient of acetonitrile (0.5%/min) starting from 1 to 21% acetonitrile for 40 min and finally 100% acetonitrile for 10 min at a flow rate of 9 mL/min, and the profile was monitored

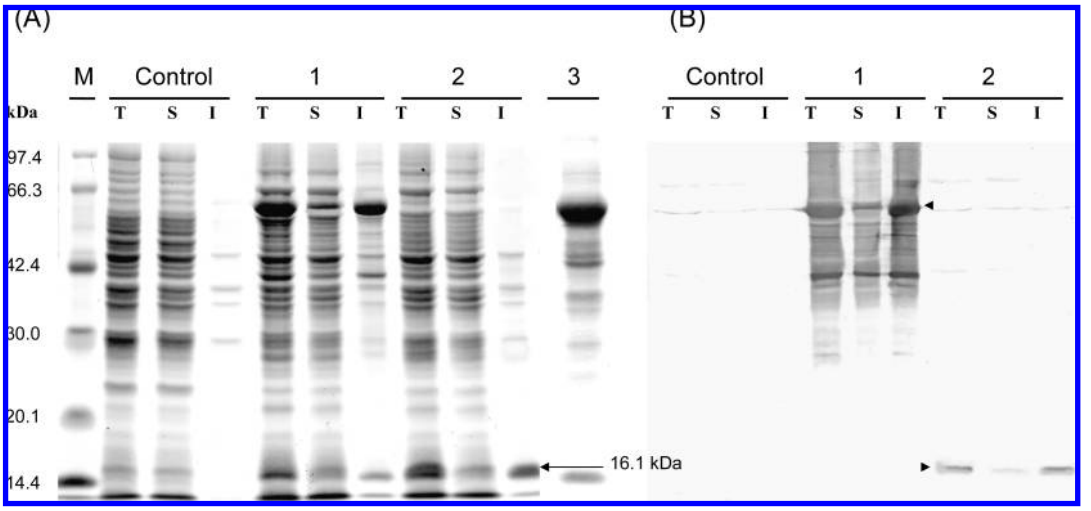

Figure 6. Analysis of the expression of A1aB1b(Δ31)-34IIAEK and 28IIAEK. (**A**) SDS-PAGE (11% acrylamide) profile of *E. coli* expressing protein. (**B**) Western blot analysis of recombinant proteins using rabbit antisynthetic-IIAEK sera. Lanes M, control, 1, 2, and 3 show electrophoretic profiles of standard size markers in kilodaltons (kDa), induced AD494 containing pET-21d, induced AD494 containing pEA1aB1b(Δ31)-34IIAEK, induced AD494 containing pE28IIAEK, and insoluble cell fraction containing A1aB1b(Δ31)-34IIAEK after washing, respectively. T, total; S, soluble; I, insoluble. Arrow indicates where 28IIAEK, a 16.1 kDa protein, should be located.

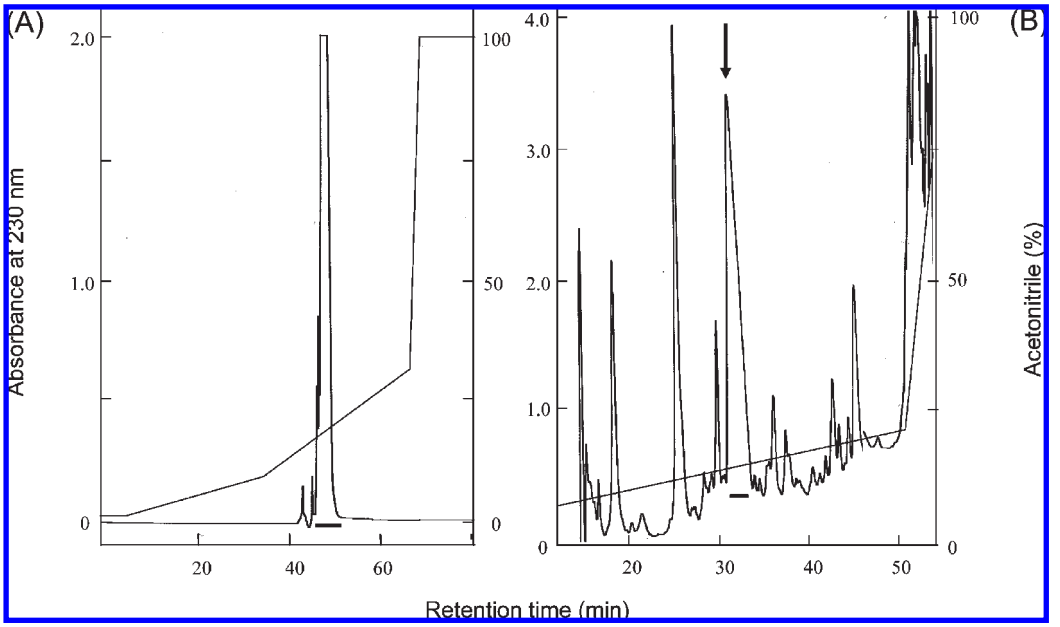

Figure 7. Purification and analysis of IIAEK from tryptic hydrolysate of inclusion bodies of A1aB1b(Δ31)-34IIAEK. (**A**) Final purification of IIAEK by octadecyl silica (ODS) column (20 mm × 250 mm). (**B**) Elution pattern of the peptides from ODS column (4.6 mm × 150 mm) for the analysis of IIAEK production yield after digestion. Arrow and bars indicate peaks of IIAEK.

at 230 nm. The fractions containing IIAEK peptides were collected, freeze-dried, and subjected again to the same column for a higher purity. The second ODS column chromatography was done at 1% acetonitrile for 5 min (**Figure 7A**). The concentration of acetonitrile was increased linearly to 11% at 30 min, then increased linearly to 31% for the next 30 min, and finally increased to 100% in the next minute. Synthetic IIAEK peptide (purity > 99.7%, Sigma Genosys Japan, Hokkaido, Japan) was used as a standard for each HPLC run. To confirm the identity of the peptide, the samples were applied onto a liquid chromatograph−mass spectrometer (Mariner, PerSeptive Biosystems) and a protein sequencer (492 Applied Biosystems).

Yield of IIAEK Peptide from Digestion and Purification. The amount of purified peptide obtained after digestion was estimated on the basis of the comparison of the peak area of absorbance at 230 nm with that of the synthetic peptide as a standard on HPLC using a small ODS column (Cosmosil 5C18-AR-II, 4.6 mm × 150 mm, Nakalai Tesque) (**Figure 7B**) according to the procedure described previously (*9*). The yield (%) of IIAEK after digestion was calculated on the basis of the purity of the protein of interest (65.0%) in the inclusion bodies as follows:

$$\text{yield of IIAEK (\%)} = [\text{amount of purified peptide (mg)}/\text{amount of theoretical peptide released after digestion (mg)}] \times 100$$

$$\text{amount of theoretical peptide released after digestion} = [\text{amount of protein of interest} \times \text{percentage of peptide mass in a mole of recombinant protein}]/100$$

Table 1. Purification of IIAEK from A1aB1b(Δ31)-34IIAEK Expressed in *E. coli*[a]

	total protein[b] (mg)	protein of interest[c] (mg)/[yield (%)]	peptide of interest[d] (mg)/[yield (%)][e]
crude extract[f]	1039 ± 25.9	311 ± 14.5/[100]	
inclusion bodies	571 ± 17.2	285 ± 9.1/[91.6]	
inclusion bodies after second wash	371 ± 13.8	241 ± 9.8/[77.5]	
after first HPLC			44 ± 3.2/[64.5]
after final HPLC			40 ± 2.7/[58.6]

[a] Values are means ± standard deviations of at least two independent experiments. [b] Total protein concentration was determined by BCA protein assay (Pierce, Rockford, IL) using bovine serum albumin as a standard. [c] The amount of expressed A1aB1b(Δ31)-34IIAEK. The calculation based on quantifying the percentage of the protein of interest in the gel compared to the total amount of the protein by using electrophoregrams. [d] The amount of peptides in IIAEK fraction. The calculation was based on the elution peak area of the fraction and compared to that of the standard IIAEK. [e] The yields (%) of IIAEK based on the amount of expressed IIAEK [28.3% mass product from 1 mol of A1aB1b(Δ31)-34IIAEK] from 241 mg of A1aB1b(Δ31)-34IIAEK. [f] The amount of crude extract from 11 g of *E. coli* wet weight.

Percentage of 34 IIAEK peptide mass in a mole of recombinant A1aB1b(Δ31)-34IIAEK was 28.3%.

The yield (%) and degree of purification (fold) were calculated on the basis of the amount of peptide of interest.

RESULTS

Gene Construction. To produce a DNA fragment coding highly repeated IIAEKs, two single-stranded DNAs coding four multimers of IIAEKs, one for 5′−3′ and the other for 3′−5′, were used. IIAEK is a short peptide composed of five amino acids, two isoleucines, an alanine, a glutamic acid, and a lysine, and ATC, GCC, GAG, and AAA for Ile, Ala, Glu, and Lys, respectively, were selected to code $(IIAEK)_4$ for high expression of IIAEK in *E. coli*. The final nucleotide sequence was $(ATCATCGCCGAGAAA)_4$. Due to the same nucleotide sequences for the individual units of the four IIAEK sequences, the annealing of two single strands for four-repeated IIAEK resulted in a double-stranded DNA fragment encoding one, two, three, or four IIAEK sequences, leaving the overhang of three, two, one, and zero of single-stranded DNA, respectively. However, the mobilities of these DNAs in an agarose gel were not clearly different from each other (**Figure 2A**, lane 1). We found that after insertion of the DNA fragment with an expected size of 12 IIAEK double-stranded DNA sequences (**Figure 2B**, lane 3) into the coding region for the variable region IV of A1aB1b containing 4IIAEK sequences, only 14 and not 16 IIAEK sequences were obtained. Furthermore, we could not obtain any clone for a complete double-stranded DNA using the longer DNA fragment containing 16 IIAEK double-stranded DNA sequence (**Figure 2B**, lane 4). This might be due to the very low efficiency of getting a complete double-stranded DNA encoding longer than 10 IIAEK.

Expression and Solubility of A1aB1b(Δ31)-14 IIAEK, A1aB1b(Δ31)-34 IIAEK, and 28IIAEK. On the basis of the results of the expression, AD494(DE3) was selected as the expression host for the two constructs. In high-salt (0.5 M NaCl) medium, the expression levels of the proteins were higher at 20 °C than at 37 °C (data not shown). Even at the low culture temperature, the expressed proteins were mainly insoluble (**Figure 6A**).

A1aB1b(Δ31)-14IIAEK, in which 31 amino acids were deleted from the disordered region IV, was expressed as an insoluble protein at a high level (29.9% of the total cell proteins). The expression level is similar to that of A1aB1b-14IIAEK (data not shown), indicating that reducing the size of mutant A1aB1b containing many IIAEKs had no effect on the expression level of the mutant.

The expression of 28IIAEK was not clearly observed because of its low expression level. The size of 28IIAEK is 16.1 kDa. However, its mobility on SDS-PAGE cannot be predicted because it is composed of highly repetitive sequences (*14*). To confirm the expression of the recombinant proteins, Western blotting was done using anti-IIAEK serum. Results of the Western blot (**Figure 6B**) and estimation by analyzing the gel image with ImageMaster 1D Elite, version 3.0 (Amersham-Pharmachia Biotech, Uppsala, Sweden), showed that the mobility of 28IIAEK on SDS-PAGE was reasonable according to its molecular size, although its expression was low (2.0% of the total proteins). **Figure 6A** shows that A1aB1b(Δ31)-34IIAEK was expressed at a much higher level (29.9% of the total proteins) than 28IIAEK. On the basis of the expression level, A1aB1b(Δ31)-34IIAEK was selected for release and purification of IIAEK peptide.

Purification of IIAEK Peptide from Digested A1aB1b (Δ31)-34IIAEK. Densitometric scanning of the SDS-PAGE profile (**Figure 6A**) showed that 93.8% of A1aB1b(Δ31)-34IIAEK which accumulated in *E. coli* was insoluble. The purity of the protein in the insoluble fraction was 65.0% (**Figure 6**A, lane 3). After overnight digestion with trypsin, no insoluble materials remained.

Chromatography of the digested product in the supernatant on the ODS column resolved the IIAEK peptide estimated to account for 28.3% of total mass product (**Figure 7B**). On the basis of the peak area of the standard IIAEK and the amount of expressed IIAEK (28.3% mass product from 1 mol of A1aB1b-34IIAEK), the yield of IIAEK after digestion was calculated to be 64.5% after the first purification (**Table 1**). The IIAEK fraction from the first purification was collected and subjected to the same column for the final purification (**Figure 7A**), which provided 40 mg of IIAEK with a purity of 97.8% in 58.6% overall yield (**Table 1**).

DISCUSSION

The need for large quantities of bioactive peptides has prompted investigations into developing efficient biological production methods. We have worked to produce a large quantity of the IIAEK peptide for medical purposes by employing two methods using either direct expression system of a multimer of IIAEKs or construction of A1aB1b containing multimers of IIAEKs and expressed them in *E. coli*. We have earlier reported that a mutant of A1aB1b containing multimers of IIAEK repeats $[(IIAEKIIAEKIIAEKIIAEK)_n]$ can be easily digested by trypsin, and a high yield of IIAEK was obtained after purification (*9*). Therefore, in constructing large multimers of IIAEK, we wanted to have only the IIAEK sequence without the amino acid residues obtained from the DNA encoding restriction sites. To attain this, we developed a new method to get a DNA encoding highly repeated multimers of IIAEK in only one step of cloning (see Materials and Methods). However, it was difficult to get a construct containing DNA encoding more than 14 multimers of IIAEK

sequences due to improper annealing caused by codon repetition (**Figures 1** and **2A**). To produce DNA encoding a highly repeated multimer of IIAEK sequences, a connection for tandem multimers of IIAEK sequences is necessary. This connection could be easily inserted by introducing restriction sites at the end of the DNA fragment and later digestion by the restriction enzymes at the same sticky end (*4*, *15*), and the resulting fragments could be purified and ligated. However, it is desirable to use unique restriction sites that are not present in pEA1aB1b-14IIAEK for ease of DNA identification after digestion and purification from agarose gel. Therefore, we introduced restriction sites that could be digested by *Nco*I, *Hin*dIII, and *Sac*I. The DNA region encoding 14 IIAEK sequences in pEA1aB1b-14IIAEK was important as a template for the construction of direct expression plasmid containing 2 14 IIAEK (28IIAEK) or 3, 4, and *n* 14 IIAEK sequences, and a large DNA containing a large number of IIAEK sequences could be introduced easily to pEA1aB1b by changing only the restriction sites. Therefore, the present method for the first step of construction of tandem multimers of a peptide should be powerful for construction of other short peptide production systems.

The expression level of 28IIAEK was low and similar to those of the other proteins with similar molecular masses: modified α′ extension region (*16*) and a tandem repeat of a hexa-multimeric buforin II (*1*). However, high expression level of the tandem repeat of buforin II was obtained by using fusion proteins with a carrier protein (*1*). The carrier was suggested to neutralize the positive charge of the antimicrobial peptide, which is lethal to the *E. coli* cell. IIAEK and the modified α′ extension region are not antimicrobial peptides, and they tend to be a neutral or negative peptide, respectively. Therefore, the low expression of the peptides might be due to inefficient intracellular synthesis and/or degradation of such polypeptides. To get a higher level of expression of the protein containing 28 multimers of IIAEK sequences, a fusion protein as a carrier for the expression should be introduced.

Increasing the number of IIAEK from 20 to 34 did not reduce the expression level of mutant A1aB1b in *E. coli*. Hence, a large number of IIAEKs can be introduced to A1aB1b without reducing high expression level. On the basis of the results, A1aB1b seems to play the role of a fusion partner in the expression of the peptide, which was not surprising considering that A1aB1b has a chaperone-like property (*17*).

The useful properties of A1aB1b that make it suitable for the introduction of many bioactive peptides can be summarized as follows: (1) it has a chaperone-like property, (2) it is expressed at a high level in *E. coli*, (3) its inclusion body such as A1aB1b(Δ31)-34IIAEK caused by its mutation is easily digested by proteinases without the help of urea or other chemicals, and (4) the yield of 58.6% of IIAEK after digestion and purification from the inclusion body was lower than the 80% yield for the soluble protein, A1aB1b-containing 8 IIAEK (*9*). In addition to the above properties, another advantage of A1aB1b is that its gene has many unique restriction sites (5), so that it is easy to introduce DNA for a greater number of the tandem multimer. Thus, A1aB1b is a really good model protein for the introduction into and production of bioactive peptides in *E. coli*.

LITERATURE CITED

(1) Lee, J. H.; Minn, I.; Park, C. B.; Kim, S. C. Acidic peptide-mediated expression of the antimicrobial peptide buforin II as tandem repeats in *Escherichia coli*. *Protein Express. Purif.* **1998**, *12*, 53–60.

(2) Čipáková, I.; Hostinová, J. Expression and purification of human antimicrobial peptide, dermcidin, in *Escherichia coli*. *Protein Express. Purif.* **2006**, *45*, 269–274.

(3) Chen, H. Q.; Xu, Z. N.; Xu, N. Z.; Cen, P. L. Efficient production of a soluble fusion containing human β-defensin-2 in *E. coli* cell-free system. *J. Biotechnol.* **2005**, *115*, 307–315.

(4) Rao, X.; Hu, J.; Li, S.; Jin, X.; Zhang, C.; Cong, Y.; Hu, X.; Tan, Y.; Huang, J.; Chen, Z.; Zhu, J.; Hu, F. Design and expression of peptide antibiotic hPAB-β as tandem multimers in *Escherichia coli*. *Peptides* **2005**, *26*, 721–729.

(5) Rao, X.; Li, S.; Hu, J.; Jin, X.; Hu, X.; Huang, J. A novel carrier molecule for high-level expression of peptide antibiotics in *Escherichia coli*. *Protein Express. Purif.* **2004**, *36*, 11–18.

(6) Huang, L.; Wang, J.; Zhong, Z.; Peng, L.; Chen, H.; Xu, Z.; Cen, P. Production of bioactive human β-defensin-3 in *Escherichia coli* by soluble fusion expression. *Biotechnol. Lett.* **2006**, *28*, 627–632.

(7) Xu, Z.; Zhong, Z.; Huang, L.; Peng, L.; Wang, F.; Cen, P. High-level production of bioactive human β-defensin-4 in *Escherichia coli* by soluble fusion expression. *Appl. Microbiol. Biotechnol.* **2006**, *72*, 471–479.

(8) Nagaoka, S.; Futamura, Y.; Miwa, K.; Awano, T.; Yamauchi, K.; Kanamaru, Y.; Tadashi, K.; Kuwata, Y. Identification of novel hypocholesterolemic peptides derived from bovine milk β-lactoglobulin. *Biochem. Biophys. Res. Commun.* **2001**, *281*, 11–17.

(9) Prak, K.; Maruyama, Y.; Maruyama, N.; Utsumi, S. Design of genetically modified soybean proglycinin A1aB1b with multiple copies of bioactive peptide sequences. *Peptides* **2006**, *27*, 1179–1186.

(10) Takenaka, Y.; Utsumi, S.; Yoshikawa, M. Introduction of a low molecular weight agonist peptide for complement C3a receptor into soybean proglycinin A1aB1b subunit by site-directed mutagenesis. *Biosci., Biotechnol., Biochem.* **2001**, *65*, 1202–1205.

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

(12) 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.* **1990**, *54*, 1543–1550.

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

(14) Matoba, N.; Doyama, N.; Yamada, Y.; Maruyama, N.; Utsumi, S.; Yoshikawa, M. Design and production of genetically modified soybean protein with anti-hypertensive activity by incorporating potent analogue of ovokinin(2−7). *FEBS Lett.* **2001**, *497*, 50–54.

(15) Doi, N.; Kakukawa, K.; Oishi, Y.; Yanagawa, H. High solubility of random-sequence proteins consisting of five kinds of primitive amino acids. *Protein Eng. Des. Sel.* **2005**, *18*, 279–284.

(16) Onishi, K.; Matoba, N.; Yamada, Y.; Doyama, N.; Maruyama, N.; Utsumi, S.; Yoshikawa, M. Optimal designing of β-conglycinin to genetically incorporate RPLKPW, a potent anti-hypertensive peptide. *Peptides* **2004**, *25*, 37–43.

(17) Choi, S.-K.; Adachi, M.; Utsumi, S. Improved bile acid-binding ability of soybean glycinin A1a polypeptide by the introduction of a bile acid-binding peptide (VAWWMY). *Biosci., Biotechnol., Biochem.* **2004**, *68*, 1980–1983.

Received for Review November 2, 2008. Accepted February 26, 2009. Revised manuscript received February 12, 2009. This work was supported in part by a grant to S.U. from the Ministry of Agriculture, Forestry and Fishery of Japan.