A Thesis

For The Degree of Master of Science

Expression and Purification of Soybean Proglycinins Modified by Protein Engineering

Department of Agricultural Chemistry

Graduate School, Cheju National University

Se-Kwang Ko

2003. 2

A Thesis

For The Degree of Master of Science

Expression and Purification of Soybean Proglycinins Modified by Protein Engineering

Se-Kwang KO

(Supervised by Professor Chan-Shick Kim)

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Agricultural Chemistry.

2002. 12

This thesis has been examined and approved.

Thesis director, Jeong-Sam Koh, Prof. of Agricultural Chemistry

---
Hae-Nam Hyun, Prof. of Agricultural Chemistry

Chan-Shick Kim, Prof. of Agricultural Chemistry

2003. 2

Department of Agricultural Chemistry

GRADUATE SCHOOL

CHEJU NATIONAL UNIVERSITY

Expression and Purification of Soybean Proglycinins Modified by Protein Engineering

Se-Kwang KO

(Supervised by Professor Chan-Shick Kim)

Department of Agricultural Chemistry

Graduate School, Cheju National University

Abstract

To improve the functional properties of soybean proglycinins, cDNA encoding soybean proglycinin was manipulated using the synthetic oligonucleotides. Modified expression plasmid $pKGA_{1a}B_{1b}$-3 was used as a template for the insertional mutation. The synthetic oligonucleotides encoding amino acid sequences of Gln-Leu-Ile-Phe-Lys-Leu and Gln-Leu-Phe-Asp-Gln-Thr-Pro-Arg-Val-Phe, respectively, were selectively inserted between Pro^{467} and Gln^{468} in the variable domain V of the clones. The insertion of new oligonucleotide sequence in recombinant plasmid pKG5S and pKG5P was confirmed by the presence of a *Pvu*II restriction site additionally created. Recombinant plasmids carrying modified cDNA were overexpressed in *Escherichia coli* JM105. Modified soybean

proglycinins are soluble proteins; thus, they were purified by consecutive treatments of salt precipitation, ion-exchange chromatography and cryoprecipitation. Purified proglycinins were confirmed via western blotting analysis.

Key words: soybean proglycinin, protein engineering, functional properties, gene expression

Contents

I. Introduction

Soybean proteins are composed of two major components, glycinin (11S globulin) and β-conglycinin (7S globulin) (2). Glycinin is one of the predominant storage proteins with various functional properties. It is composed of six subunits, each of which consists of an acidic and a basic polypeptide that are linked by a disulfide bridge (1). Modification for improving functional properties of food protein has been carried out by chemical, enzymatic and genetic methods (7,12). In particular, genetic modification via protein engineering is now a powerful approach to attain this goal (2,6,9). Glycinin is a suitable target for the improvement of functional properties due to the plethera of information reported about this soybean proglycinin (3,4,10). A reasonable strategy for protein engineering of glycinin protein has been established based on the information available on the primary amino acid and tertiary structure of the protein as well as cloning of cDNA encoding the target protein. An expression of cDNA encoding the proglycinin $A_{1a}B_{1b}$ subunit precursor has been reported (19). Furthermore, the stepwise deletion of the cDNA region encoding the signal sequence and N-terminal regions of the proglycinin $A_{1a}B_{1b}$ resulted in the overexpression and the stability of protein products in *E. coli* (5). From comparing the homology of amino acid sequence among various 11S globulins from various legumes, a series of alternating conserved and variable domains were found. Previously, we succeeded in creating novel proglycinins designed to improve their nutritional and functional properties

(heat-induced gelation and emulsification) by protein engineering based on the 3-dimensional structure of glycinin as well as their structure-function relationships (6). For example, the removal of the variable domains and insertions of continuous plural methionines into the variable domains were created for improving the food quality of soybean glycinin (6). In addition to nutritional and functional properties of soybean glycinin, it can be considered as a physiologically functional food material. The angiotensin I-converting enzyme (ACE) participates in the control of blood pressure. A number of ACE inhibitory peptides have been isolated from various food proteins (11,20). In particular, it was reported that protease hydrolyzate derived from soybean product has a potent ACE inhibitory activity (15). Their peptide sequences were also reported (14). Therefore, it is possible to create a glycinin protein with higher content of bioactive peptide sequences. Unfortunately, it has been reported that cDNA manipulation may cause no expression or no formation of the inclusion body in *E. coli*. To evaluate the functional properties of novel glycinins, the soluble glycinin protein need to be first accumulated in *E. coli* and purified as a native form with conformational stability. A recent report indicates that modified glycinins can be processed to mature form and self-assembled into hexamers in transgenic tobacco seeds (16). It indicates that engineered glycinins could be accumulated effectively in the transgenic plants if modified glycinins have the proper conformation. Because soybean glycinin plays an important role as a functional protein as well as a physiologically functional food, the improvement of functional properties

of soybean glycinin is important the food industry (7). In this paper, we describe the expression and purification of soybean proglycinin proteins modified with additional peptide sequences in *E. coli*.

II. MATERIALS AND METHODS

Bacterial strain and plasmids

E. coli JM105 and *E. coli* DH5α were used as the host cells for gene expression and plasmid manipulation, respectively. Both strains were grown in LB medium at 37 ℃.

Plasmid pKG5S and pKG5P were derived from an expression plasmid $pKGA_{1a}B_{1b}$-3, which contained cDNA encoding truncated proglycinin without three N-terminal amino acid sequence. In an expression plasmid, ATG codon is under the control of *tac* promoter in pKK233 (6). DNA manipulation was performed as described in Sambrook *et al.* (13)

Construction of recombinant plasmids

A clone coding proglycinin contained a unique *Eco*81I restriction site in variable region V. The synthetic oligonucleotides were inserted in this region. Oligonucleotides used for mutation were constructed for the purpose of increasing the hydrophobicity of glycinin as well as fortifying the biologically active oligopeptide sequence.

(S) functional oligopeptide sequence: Gln-Leu-Ile-Phe-Lys-Leu-

```
                    Q   L   I   F   K   L
SAM18A  5'-T CAG CTG ATC TTC AAG CT - 3'
PIN18B        3'- C GAC TAG AAG TTC GAA GT - 5'
```

(P) biologically active oligopeptide sequence:

-Gln-Leu-Phe-Asp-Gln-Thr-Pro-Arg-Val-Phe-

```
                    Q   L   F   D   Q   T   P   R   V   F
PEPA30T  5' - T CAG CTG TTC GAC CAG ACC CCT CGC GTG TT - 3‘
PEPB30B        3' - C GAC AAG CTG GTC TGG GGA GCG CAC AAA GT - 5'
```

Underlined oligonucleotides indicate a new *Pvu*II restriction site. Four oligonucleotides were synthesized from Korea Biotech. Inc. To make the double strand oligonucleotide, a pair of oligonucleotides prepared for the synthesis of each oligopeptide sequence was mixed with annealing buffer (0.1M NaCl, 10 mM Tris-HCl, pH 7.8, 1.0 mM EDTA) at a final concentration of 1ng/㎕ (13). The reaction mixture was heated at 65 ℃ for 5 min and then cooled to room temperature for 30 min. To construct the recombinant plasmids, pKG5S and pKG5P containing new synthetic oligonucleotide, an expression plasmid $pKGA_{1a}B_{1b}$-3 was digested with *Eco*81I and then ligated with each paired oligonucleotide at 16 ℃ for 16 h. A 30 ㎕ of ligation mixture was treated with ice cold ethanol and 3 M sodium acetate at -20℃ for 6 h. The ligated DNA was harvested by centrifugation at 12,000 rpm for 10 min. In order to reduce self-ligation, the resuspended DNA was again digested with *Eco*81I and then transformed into *E. coli* DH5α. Transformants were selected on LB agar containing 50 ㎍/ml of ampicillin. Plasmid was prepared with 2.0 ml of LB culture broth by SDS alkaline method. A new *Pvu*II restriction site was confirmed between an expression plasmid $pKGA_{1a}B_{1b}$-3 and

recombinant plasmids pKG5S and pKG5P.

Expression of modified proteins in *E. coli*

E. coli JM105 strain harboring an individual expression plasmid was grown for 20 hr at 37 ℃ after inclusion of 2 mM of isopropyl β-D-thiogalactopyranoside (13). Cells harvested by centrifugation were disrupted by sonication in 50 mM of ice cold Tris-HCl buffer (pH 8.0) containing 1 mM EDTA and 1.5 mM phenylmethelsulfonyl fluoride (PMSF). Total cell lysate was analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (8).

Purification of modified proglycinins

Cells were harvested from 4 L of LB broth and disrupted by sonication. The cell debris was removed by centrifugation at 5000 rpm for 10 min. The supernatant was saved. The proglycinin in supernatant was concentrated by ammonium sulfate fractionation in the condition of 40-65% saturation. Q-Sepharose column was equilibrated with buffer A (35 mM potassium phosphate buffer, pH 7.0 containing 0.15 M NaCl, 1.5 mM PMSF, 10 mM 2-ME and 0.02% (w/v) sodium azide). A protein sample loaded was eluted with a 500 ml of salt gradient buffer (0.15 M-0.5M NaCl in buffer A). Finally, the collected protein solution was purified by cryoprecipitation as described previously (5). Amounts of purified proglycinins were measured by densitomertic analysis after staining a gel with coomassie brilliant blue solution.

Immunoblotting

The purified proglycinins were also confirmed by immunoblotting assay with a rabbit antiserum against soybean proglycinin (18). Purified soybean glycinin was injected into a rabbit three times in order to boost antibody production. A rabbit antiserum was collected and purified by a glycinin-Sepharose column. Purified recombinant glycinin protein was analyzed by SDS-PAGE. All proteins were transferred to the nitrocellulose filter and western blotting was performed using glycinin-antiserum as described previously (18).

N-terminal amino acid sequence

To confirm the production of proglycinin in *E.coli*. N-terminal amino acid sequence of purified glycinin was determined. The analyses were done at seoul branch of Korea Basic Science Center. Purified proglycinin was concentrated by SDS-PAGE and then blotted to a polyvinylidene difluoride membrane according to the procedure of Matsudaira (10). N-terminal amino acid sequence was determined by the Edman degradation method with a Milligen 660B protein sequencer.

III. RESULTS AND DISCUSSION

Construction of recombinant plasmids

Previously, pKG$A_{1a}B_{1b}$-3, an expression plasmid was constructed for producing soybean proglycinin in *E. coli* (19). In order to produce various proglycinin variants with enhanced functional properties, annealed synthetic oligonucleotides were inserted into the variable domain V (Fig.1). The recombinant plasmid pKG5S and pKG5P constructed have the functional and biologically active peptide sequence, respectively. The overall scheme for constructing pKG5S and pKG5P is shown in Fig. 2. The insertion of additional nucleotide sequences in cDNA encoding proglycinin was confirmed by the presence of a *Pvu*II restriction site. As shown in Fig. 3, the *Pvu*II digestion of pKG5S and pKG5P generated two DNA fragments (3 kb and 2.5 kb), indicating that a new *Pvu*II restriction site. Also, pKG5S and pKG5P were not digested with *Eco*81I restriction enzyme. These plasmids were used for production of modified soybean proglycinin in *E. coli* JM105.

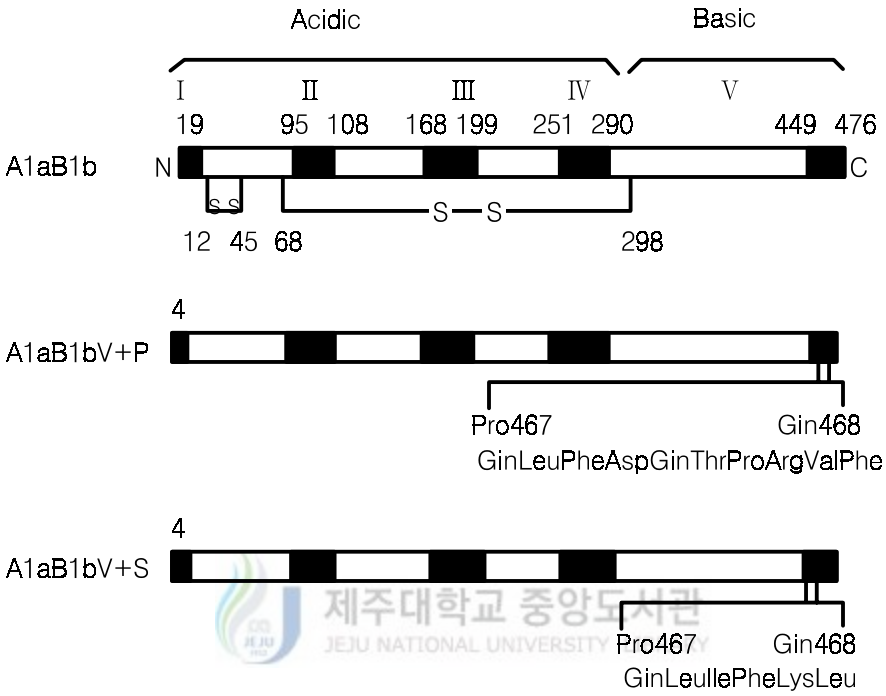

Fig. 1. Schematic representation of the proglycinin $A_{1a}B_{1b}$ subunit and modified proglycinins. The numbers of the residues from the N-terminus are described for the variable domains (I-V) above the alignment. Open bar indicates the conserved region. $A_{1a}B_{1b}$V+P and $A_{1a}B_{1b}$V+S indicate the modified proglycinin, respectively.

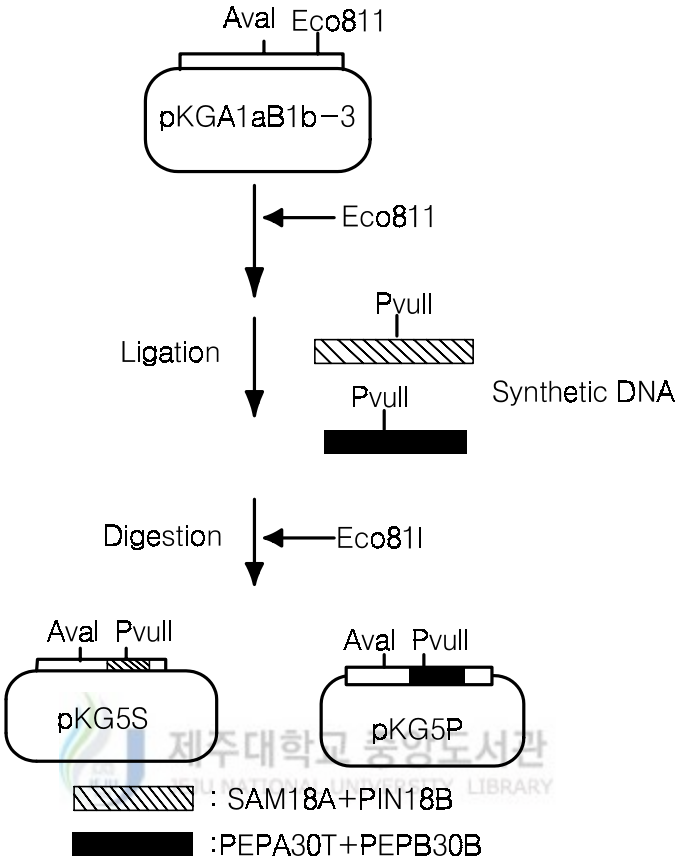

Fig.2. Schematic diagram for construction of recombinant plasmid pKG5S and pKG5P

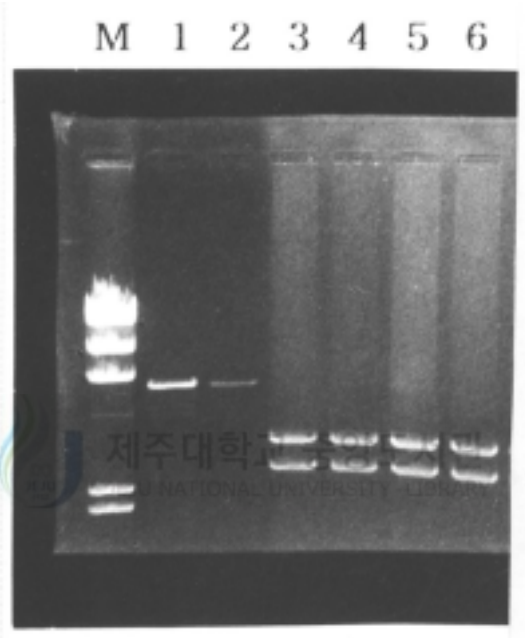

Fig. 3. Restriction patterns of expression plasmid and recombinant plasmids digested with *Pvu*II. λ-DNA digested with *Hin*dIII (lane M), pKGA$_{1a}$B$_{1b}$-3 (lane 1), pKGA$_{1a}$B$_{1b}$V+4Met (lane 2), pKG5S (lane 3,4) and pKG5P (5,6).

Expression and purification of proglycinin protein

The modified proglycinin proteins directed by pKG5S and pKG5P were analyzed with SDS-PAGE. The total cell lysate prepared from *E. coli* containing recombinant plasmid showed different band patterns

compared to those of *E. coli* JM105 harboring pKK233 (Fig. 4, 5). Cells harboring a recombinant plasmid produced a modified protein with a molecular weight of 55 kDa. The total cell lysate obtained from *E. coli* JM105 containing plasmid pKG5P was purified. Fig. 4 shows the protein patterns between the total cell lysate and the modified glycinin protein purified by ammonium sulfate fractionation, ion-exchange chromatography and cryoprecipitation. Cells harboring pKG5P produced a protein with a molecular weight of 55 kDa which is identical with the estimated molecular weight. The expressed protein was significantly concentrated by ion-exchange chromatography and further purified by cryoprecipitation. The yield of purified proglycinin was 21 mg/ L of culture. It also indicates that the modified glycinin is accumulated as a soluble protein in *E. coli.* Which allow us to evaluate the functional properties of modified proglycinin protein.

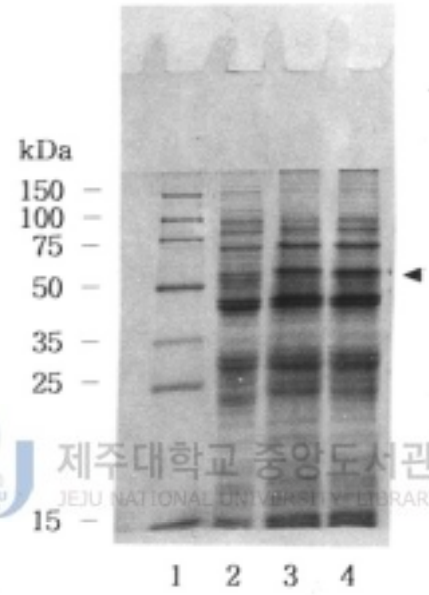

Fig. 4. SDS-PAGE analysis of purified various proglycinin obtained from recombinant plasmids. Lane 1, high molecular weight standard proteins (15, 25, 35, 50, 75, 100 and 150 kDa); Lane 2, total cell lysate of JM105/pKK233(control): Lane 3, JM105/pKG5S; Lane 4, JM105/pKG5P.

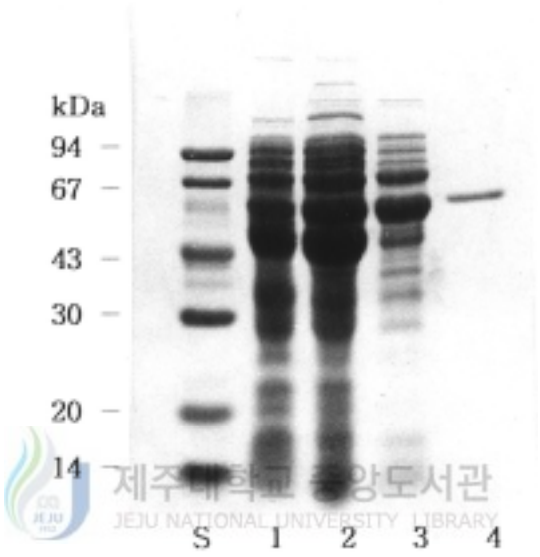

Fig. 5. SDS-PAGE analysis of the total cell lysate . Lane S, low molecular weight standard proteins; Lane 1, total cell lysate of JM105/pKG5P; Lane 2, ammonium sulfate fractionation (40-65%); Lane 3, Q-Sepharose pool; Lane 4, cryoprecipitation. The arrow indicates the modified proglycinin. The numbers denote molecular weights of marker proteins.

Immunoblotting

Immunoblotting analysis was carried out using a polyclonal antibody prep.. Modified proglycinin proteins were obtained from *E.coli* cells harboring plasmid pKG5P or pKG5S. Purified proglycinins showed the positive signal indicating that the soybean proglycinin protein was expressed in *E. coli* (Fig. 6).

Therefore, the proglycinins modified with additional oligopeptide sequences were successfully expressed and purified via simple purification steps. The soluble form of modified proglycinins will facilitate the purification as well as the evaluation of the functionality of proglycinin in *E. coli.* The level of production of proglycinins modified with additional oligopeptide sequence was lower than that of the proglycinin from an expression plasmid pKG$A_{1a}B_{1b}$-3. The modified proglycinin was subjected to analysis of its N-terminal amino acid sequence analysis. The modified proglycinin obtained from pKG5P showed the 8 *mer* amino acid sequence which is identical with that predicted from the nucleotide sequences of the deleted $A_{1a}B_{1b}$-3 cDNA. Design and production of novel proglycinins which are expected to exhibit better food and/or physiological functions than the normal soybean glycinin is based on the relationships between the structure and the functional properties. It is believed that modified proglycinins are valuable for the study of structure-function relationships, in soybean proglycinins.

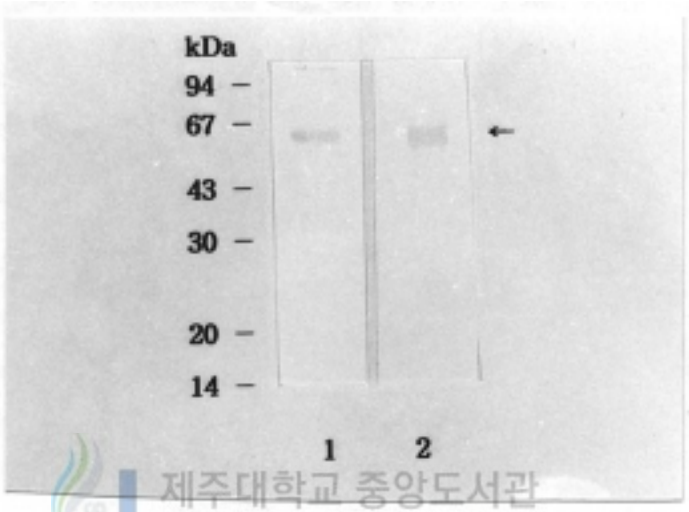

Fig. 6 . Western blot analysis of recombinant proglycinins.

Lane 1, proglycinin purified from *E. coli*/pKG5S; Lane 2, proglycinin purified from *E. coli*/pKG5P. The numbers indicate the molecular weight of marker proteins low molecular weight marker.

IV. REFERENCES

1. Badley, R.A., D. Atkinson, H. Hauser, D. Oldani, J.P. Green and M.M. Stubbs. 1975. The structure, physical and chemical properties of the soybean protein glycinin. *Biochem. Biophys. Acta* **412**: 214-228

2 Hoffman, L.M., D.D. Donaldson, E.M. Herman. 1988. A modified storage protein is synthesized, processed and degraded in the seeds of transgenic plant. *Plant Mol. Biol.* **11**: 717-729.

3. Gidamis, A.B., B. Mikami, T. Katsube, S. Utsumi and M. Kito. 1994. Crystallization and preliminary X-ray analysis of soybean proglycinins modified by protein engineering. *Biosci. Biotech. Biochem.* **58**: 703-706.

4. Katsube, T., A.B. Gidamis, J. Kanamori, I.J. Kang, S. Utsumi and M. Kito. 1994. Modification tolerability of the hypervariable region of soybean proglycinin. *J. Agric. Food Chem*. **42**: 2639-2645.

5. Kim, C-S., S. Kamiya, J. Kanamori, S. Utsumi and M. Kito. 1990. High-level expression, purification and functional properties of soybean proglycinin from *Escherichia coli*. *Agric. Biol. Chem*. **54**: 1543-1550.

6. Kim, C-S., S. Kamiya, T. Sato, S. Utsumi and M. Kito. 1990. Improvement of nutritional value and functional properties of soybean glycinin by protein engineering. *Protein Engineering* **3:** 725-731.

7. Kinsella, J.E. 1979. Functional properties of soy proteins. *J. Am. Oil Chem. Soc*. **56**: 242-258

8. Laemmli, U.K. 1970. Cleavage of structural proteins during assembly of the head of bacteriophae T4. *Nature* **227**: 680-685.

9. Lee, Sam-Pin. 1993. Enhancing the functional properties of bovine β -lactoglobulin A by genetic modification, Ph.D. Thesis, Cornell Univerity. 1-161.

10. Matsudaira, P. 1987. Sequence from picomole quantities of proteins electroblotted onto polyvinylidene difluoride membranes. *J. Biol. Chem.* **262**: 10035-10038.

11. Matsufuji, H., T. Matsui, E. Seki, K. Osajima, M. Nakashima and Y. Osajima. 1994. Angiotensin I-converting enzyme inhibitory peptides in an alkaline protease hydrolyzate derived from sardine muscle. *Biosci. Biotech. Biochem.* **58**: 2244-2245.

12. Richardson, T. 1985. Chemical modifications and genetic engineering of food proteins. *J. Dairy Sci.* **68:** 2753-2762

13. Sambrook, J., E.F. Fritsch, and T. Maniatis. 1989. Molecular cloning: a laboratory manual. 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harvor, NY.

14. Kawamura, Y. 1997. Food protein and antihypertensive peptides. *Farming*

Japan **31**: 14-19.

15. Shin, Z-I., C-W, H-S. Ahn, H-S. Nam, H-J. Lee, and T-H. Moon. 1995. Fractionation of angiotensin converting enzyme inhibitory peptides from soybean paste. *Korean J. Food Sci. Technol.* **27**: 230-234.

16. Utsumi, S., S. Kitagawa, T. Katsube, T. Higasa, M. Kito, F. Takaiwa and T. Ishige. 1994. Expression and accumulation of normal and modified soybean glycinins in potato tubers. *Plant Science* **102**: 181-188.

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

18. Utsumi, S., C-S. Kim, M. Hohno, and M. Kito. 1987. Polymorphism and expression of cDNAs encoding glycinin subunits. *Agric. Biol. Chem.* **51**: 3267-3273.

19. Utsumi, S., C-S. Kim, T. Sato and M. Kito. 1988. Signal sequence of preproglycinin affects production of the expressed protein in *Escherichia coli*. *Gene* **7:** 349-358

20. Yano, S., K. Suzuki and G. Funatsu. 1996. Isolation from α-zein of thermolysin peptides with angiotensin I-converting enzyme inhibitory activity. *Biosci. Biotech. Biochem.* **60**: 661-663.

초 록

단백질 공학에 의한 재조합 대두 Proglycinins의 발현과 정제

고 세 광

(지도교수: 김찬식)

제주대학교 대학원 농화학과

Glycinin은 대두의 주요 저장 단백질중의 하나로써 대두단백질의 기능성에 중요한 역할을 담당하고 있다. 대두 단백질의 기능특성과 영양가등을 향상시키기 위해서 $A_{1a}B_{1b}$ proglycinin subunit를 glycinin의 구조와 기능과의 관계를 기초로하여, 여러종류의 식물종자 globulins단백질의 아미노산서열을 비교함으로써 개량 가능한 영역(variable domains)에 modification을 시도하였다.

본 연구에서는 $A_{1a}B_{1b}$ Proglycinin을 code한 cDNA의 variable domains에서 Pro467과 Gln468사이에 혈압강하작용을 나타내는 Angiotensin Converting Enxyme(ACE)저해 peptide인 Asp-Gln-Thr-Pro-Arg-Val-Phe 및 유화성을 개량하기 위한 Ile-Phe-Lys-Leu의 아미노산 서열을 code한 합성 oligonucleotides을 각각 도입하여 modified expression plasmids를 조제하였다. 각각의 modified expression plasmides를 대장균 JM105균주를 사용하여 발현을 시도하였다. 각각의 개량된 단백질들은 대장균 균체내에 용해성 단백질 형태로 축적되었으며, 분자집합능력을 가지고 있었다. 개량된 단백질들은 유안암모늄 분획, 이온교환크로마토그라피, 그리고 대두 native glycinin의 고유한 성질인 6mM Tris-HCl(pH 6.3) buffer, 4℃에서 침전하는 성질을 이용하여 정제할 수가 있

었다.

본 연구결과를 통해서 glycinin cDNA의 대장균 발현 시스템이 단백질 공학적으로 설계한 대두단백질의 식품품질과 분자집합능력을 평가하는데 사용할 수 있으며, 이론적으로 생리적 기능성을 겸비하고, 우수한 기능특성을 가진 대두 단백질을 창제할 수 있다는 가능성을 제시하고 있다.

Key words: soybean proglycinin, protein engineering, functional properties, gene expression

감사의 글

지도교수님이신 김찬식 교수님과의 인연은 강의시간에서의 인연으로 시작하여 10여년 동안 지도교수님과 학생으로 선배님과 후배로 주임교수님과 조교로 때로는 엄한 부모님의 역할을 해주셨습니다. 제게 누구보다도 채찍과 당근을 함께 보내주신 교수님이 계셨기에 제가 지금의 이 자리에 설 수 있었습니다. 무엇보다도 실험 뿐만 아니라 다방면에 걸쳐 경험을 쌓게 해주신 교수님께 진심으로 감사드립니다.

학부시절부터 아낌없는 충고와 관심을 아끼지 않으신 강순선 교수님, 유장걸 교수님, 고정삼 교수님, 류기중 교수님, 현해남 교수님께 진심으로 감사드립니다. KAIST 식물발생학 연구실의 정원일 교수님과 김인중 박사님께 감사드립니다. 일본식품총합연구소의 Chikafusa Fukazawa 박사님께 감사드립니다. 계명대학교의 이삼빈 교수님께 감사드립니다. 강지용 학장님, 김원길 선생님께 감사드립니다. 항상 저를 보이지 않게 도와주신 고석형 선생님께 감사드립니다. 실험실 생활을 하면서 정이 들었던 고경철, 남태식 후배님께도 감사드립니다. 그리고, 공개발표하는 날 저에게 파이팅을 외쳐 주셨던 김기훈 전무님과 이원후 상무님 그리고 동료직원들에게 감사드립니다.

마지막으로 제가 이 자리에 오기까지 많은 뒷바라지와 희생, 사랑을 베풀어 주신 부모님과 사랑하는 부인, 그리고 우리 귀염둥이 연우에게 이 모든 영광을 돌립니다.

2002년 12월 이른 새벽에