Doctoral Dissertation

Study on the precursor of Gly m Bd 28K, a soybean allergen

2013

Hitomi YUMIOKA

Graduate School of Health & Welfare Science

Okayama Prefectural University

Contents

Abbreviations

BSA	bovine serum albumin
BSA-PBS	PBS/T supplemented with 0.8% (w/v) BSA, 0.1% (w/v) gelatin, and 2 mM NaN_3
CCD	cross-reactive carbohydrate determinants
Con A	concanavalin A
ctVSDs	C-terminal VSDs
DIC	Differential Interference Contrast
2D-PAGE	two-dimensionalpolyacrylamide gel electrophoresis
EGFP	enhanced green fluorescent protein
ERAD	ER-associated degradation
FM4-64	*N*-(3-triethylammoniumpropyl)-4-(6-(4-(diethylamino) phenyl)hexatrienyl)-pyridiniumdibromide
Gm23K	the C-terminal peptide peptide (the 23-kDa peptide) sequence of the Gly m Bd 28K precursor
Gm28K	Gly m Bd 28K
HRP	horseradish peroxidase
mAb	monoclonal antibody
MHLW	Ministry of Health, Labor, and Welfare
mLS medium	modified Linsmaier and Skoog medium
NtADH 5'-UTR	*Nicotianatabacum* alcohol dehydrogenase 5'-untranslated region
ORF	open reading flame
PBS	phosphate-buffered saline, pH 7.4
PBS-T	PBS containing 0.05% Tween 20
PSV	protein storage vacuoles
PVDF	polyvinylidene fluoride membrane
5' RLM-RACE	5' RNA-ligand mediated Rapid amplification of cDNA ends
SDS-PAGE	sodium dodecyl sulfate-poly acrylamide gel electrophoresis
SP	N-terminal 24 amino acid residues of the Gly m Bd 28K precursor
ssVSDs	sequence-specific VSDs
TBS/T	10 mM Tris–HCl buffer containing 0.05% Tween 20, pH 7.4
VSDs	vacuolar-sorting determinants

Chapter 1

Introduction

Food allergy is a serious social problem in both developed and developing countries. Health hazard is frequently caused by foods containing allergens. In order to prevent the allergy induced by the allergic foods, it would be difficult to remove the allergic foods from diets and it would be better to label containing the allergic foods in food products. The General Assembly Joint FAO/WHO Food Standards Commission (Codex Alimentarius Commission) on Food Labeling has listed in 1999 the foods and ingredients that cause the most severe reactions and most cases of food hypersensitivity. The general standards for the labeling of prepackaged foods states: "The following foods and ingredients are known to cause hypersensitivity and shall always be declared: 1) Cereals containing gluten; i.e., wheat, rye, barley, oats, spelt or their hybridized strains and products of these; 2) crustacea and products of these; 3) eggs and egg products; 4) fish and fish products; 5) peanuts, soybeans and products of these; 6) milk and milk products (lactose included); 7) tree nuts and nut products; and 8) sulphite in concentrations of 10 mg/kg or more."

The Japanese government applied the above-mentioned general standards for protection of allergy induced by ingestion of allergic foods. The food labeling was divided into two stages, mandatory and recommended, based on the number of cases of actual illnesses and the degree of seriousness (Table 1). The cause of food allergy may be dependent on changes in the dietary life and environment of social life. Eggs, milk products and wheat are three representative allergic foods among ones inducing immediate-type food allergy and they were shown to account for 70% of the total allergic population induced by ingestion with allergic foods. By the monitoring survey of 2001, 2002, 2005, 2008 and 2011 in Japan (*1*), soybean was shown to rank within the top ten of representative allergic foods in all of the surveys.

**Table 1. Allergenic ingredients designated by the MHLW of Japan*

Specific allergenic ingredients
Mandatory by ministerial ordinance (seven ingredients)
Egg, milk, wheat, buckwheat, peanut, shrimp/prawn, and crab
Subspecific allergenic ingredients
Recommended by ministerial notification (20 ingredients)
Abalone, squid, salmon roe, orange, kiwifruit, beef, walnut, salmon, mackerel, soybean, chicken, banana, pork, matsutake, mushroom, peach, yam, apple, gelatin, cashew nut and sesame

*Based on the notification of March 15, 2001 and newest notification of June 3, 2008 and next newest notification of September 20, 2013 from the Department of Food Safety, Ministry of Health, Labor, and Welfare (MHLW), Consumer Affairs Agency of Japan.

Soybean is known as one of the representative allergic foods but it is one of the major sources of protein and lipid, and is utilized as a supplement for the preparation of many food products and it is also widely used as a food material. In the course of the investigation of the allergens in soybean by an immunoblotting technique with the sera of soybean-sensitive patients, Ogawa *et al*. (*2*) found out that soybean has 16 IgE-binding components. Of the allergens, Gly m Bd 68K, Gly m Bd 30K and Gly m Bd 28K (Gm28K) were shown to be the major allergens in soybean (*2*). Gly m Bd 68K was identified as the α-subunit of β-conglycinin (*3*). Furthermore, the N-terminal amino acid residues of Gly m Bd 30K were found to be identical to those of the 34-kDa soybean seed storage vacuole protein (*4*), which is a distantly related member of the papain superfamily that binds an elicitor secreted from *Pseudomonas* (5).

On the other hand, Gm28K was purified from defatted soybean flakes, characterized and proposed the involvement of its N-linked glycan moiety in the binding to IgE antibody in the sera of soybean-sensitive patients (*6-8*). The finding is compatible with the reports that the

allergenicity of some allergens with N-linked glycan moiety would be connected with their N-linked glycan (*9-13*) . Further Tsuji *et al.* cloned a cDNA encoding Gm28K. The amino acid sequence deduced from the nucleotide sequence of the cDNA cloned was suggested to encompass 473 amino acid residues including 220 amino acid residues of Gm28K (*8*). However, the cloned cDNA did not contain a start codon, which means that the full-length amino acid sequence of the predicted protein has never been elucidated. The protein deduced from the cDNA was shown to have high homology MP27/MP32 protein in the protein storage vacuoles (PSV) in pumpkin seeds and a globulin-like protein of carrot (50.4% identity between Gm28K and the pumpkin protein; 45.9% identity between Gm28K and the carrot protein) (*8*). The MP27/MP32 is biosynthesized as a preproprotein which is composed of a signal peptide, MP27 and MP32 (*14*). After removal of the signal peptide, the proprotein is processed to MP27 and MP32 during the development of cotyledons. Gm28K was estimated to be composed of 220 amino acid residues (*6*), the N-terminal amino acid residue of which is Phe22 on the deduced amino acid sequence (*8*). These findings suggest that Gm28K may be biosynthesized as a preproprotein, and that the allergen may be derived from the preproprotein in the same manner as the MP27/MP32 is. Thus, the precursor of Gm28K would be composed of a hydrophobic signal peptide, Gm28K and the C-terminal peptide (the 23-kDa peptide) (Gm23K) with a molecular mass of approximate 23-kDa. However, the 23-kDa peptide has never been characterized.

In the present study, I elucidated for the first time the full-length cDNA encoding the Gm28K precursor. Moreover, I demonstrated that the N-terminal amino acid sequence of Gm28K precursor functions as a signal peptide and that Gm28K and Gm23K are transported into the vacuoles of suspension-cultured cells of tobacco BY2 (*Nicotianatabacum* L. cv. Bright Yellow 2). In addition, I revealed the presence of Gm23K in soybean and we have also examined the chemical characteristics of the protein, including allergenicity.

Chapter 2

Cloning of cDNA encoding the precursor of Gly m Bd 28K, a soybean allergen and expression of its in tobacco BY2 suspension-cultured cells

As described in Chapter 1, soybean contains 16 allergens (2). Among them, Gly m Bd 68K, Gly m Bd 30K, and Gm28K are the major allergens (2). The former two allergens have been well characterized (3-5). On the other hand, Gm28K was isolated from soybean and was shown to be an unknown protein (6). Furthermore, a cDNA encoding Gm28K has been isolated and it is a 1567-bp nucleotide containing the 3'-untranslated region of 148-bp. The open reading frame (ORF) with a nucleotide sequence of 1419-bp has been revealed. The cDNA was shown to encode a polypeptide with 473 amino acid residues. The ORF of Gm28K encodes the protein with a molecular weight of approximately 50-kDa. However, the molecular mass of Gm28K was 28-kDa (6). Therefore, the amino acid encoding the Gm28K is assumed to be composed by about 250 amino acid residues. Thus, the cDNA cloned encoding the Gm28K has been seems to be a precursor sequence of Gm28K.

Interestingly the deduced amino acid sequence is highly homologous to MP27/MP32 proteins that is a membrane protein of the PSV in pumpkin seeds and a globulin-like protein of carrot (*8*). The MP27/MP32 is biosynthesized as a preproprotein which is composed of a signal peptide, MP27 and MP32 (*14*). After removal of the signal peptide, the proprotein is processed to MP27 and MP32 during the development of cotyledons. By comparison of the amino acid sequence deduced from the nucleotide sequence of the cDNA cloned for Gm28K with that of MP27/MP32, Gm28K is suggested to be biosynthesized as a preproprotein, and that the allergen may be derived from the preproprotein in same manner as the MP27/MP32 is. Thus, the precursor of Gm28K would be composed of a hydrophobic N-terminal signal peptide, Gm28K and Gm23K. However, among the three segments, the presence of N-terminal signal peptide and Gm23K in soybean has never been characterized.

Tobacco BY2 cells is a cell line of plant cells and the cells have been used widely in material production. As shown in Chapter 1, the sugar moiety of Gm28K would be involved in its allergenicity. It is very important to prepare a large amount of Gm28K for further

investigation of its allergenicity. Tobacco BY2 cells is an adequate host organism, because the cells can produce the proteins with sugar chain similar to that of Gm28K. We tried to express Gm28K in tobacco BY2 cells, but failed to express it. This is due to the fact that the cDNA cloned previously did not encode the full-length cDNA. In order to obtain a large amount of Gm28K and to elucidate the relationship between the sugar moiety of Gm28K and its allergenicity, we tried to develop an expression system in tobacco BY2 cells.

In the present chapter, first, I have elucidated the full-length cDNA encoding the Gm28K precursor. And I examined the function of the N-terminal peptide. I investigated the transport of the Gm28K precursor expressed as a fusion protein with enhanced green fluorescent protein (EGFP) in the tobacco BY2 cells.

2.1 MATERIALS AND METHODS

2.1.1 Preparation of total RNA and cDNA

Total RNA was extracted from developing soybean seeds (*Glycine max*, L. cv. Wasesuzunari) using a RNeasy Plant Mini Kit (QIAGEN, North Rhine-Westphalia, Germany). 5' RNA-ligand-mediated rapid amplification of cDNA ends (5' RLM-RACE) was carried out using a First Choice RLM-RACE kit (Ambion, Austin, TX) according to the manufacturer's instructions. Two RNA 5'-adaptor primers were supplied with the RLM-RACE kit and the following three gene-specific reverse primers were used to obtain the Gm28K precursor gene (*8*): 5'-TCCTACCACCATGGCTTTTCA-3' (outer primer), 5'-GTCTTGAAAACCCTCGTGGA-3' (first inner primer) and 5'-GCCACTCCATGGCAAAGAACAA-3' (second inner primer). These PCR products were

purified and subsequently cloned into the pGEM®-TEasyvector (Promega, Madison, WI), and their sequences were elucidated using an ABI PRISM™ 3130 Genetic Analyzer (Applied Biosystems, Foster City, CA).

2.1.2 Homology search

Homology search of amino acid sequence of the Gm28K precursor done using global alignment tools that create an end-to-end alignment of the sequences to be aligned (http://www.ebi.ac.uk/Tools/psa/). The program is stretcher and the matrix is EBLOSUM62.

2.1.3 Plasmid construction

First, the gene encoding the *Nicotianatabacum* alcohol dehydrogenase 5'-untranslated region (*NtADH* 5'-UTR) was amplified using a plasmid containing the *NtADH* 5'-UTR region (*15*) as the template and primers 1 and 2 (Table 2). The N-terminal peptide of the Gm28K precursor (SP) was amplified using cDNA prepared by the 5'RLM-RACE method and primers 3 and 4. The Gm28K- and Gm23K-coding regions were amplified using cDNA encoding the Gm28K precursor and primers 5 and 6. These amplified fragments were then used to construct *Xba*I-*NtADH5*'-UTR-SP- Gm28-Gm23K-*Sac*I with primers 1 and 6.

Each construct shown in Fig. 1 was prepared using PCR. SP-Gm28-Gm23K-EGFP and SP-EGFP-Gm28-Gm23K were constructed using the coding regions of the *Xba*I-*NtADH* 5'-UTR-signal peptide- Gm28K-Gm23K-*Sac*I, EGFP, and the following primer sets: primers 1, 9, 18, and 19 for SP-Gm28-Gm23K-EGFP; and primers 1, 6, 7, 8, 10, and 11 for SP-EGFP-Gm28-Gm23K. EGFP-Gm28-Gm23K was constructed using the coding region of SP-EGFP-Gm28-Gm23K, and primers 1, 6, 16, and 17. EGFP-Gm28K and EGFP were prepared using the coding region of EGFP-Gm28-Gm23K and the following primer sets: primers 1 and 12 for EGFP-Gm28K; and primers 1 and 9 for EGFP. SP-EGFP-Gm28K, SP-EGFP-Gm23K and SP-EGFP were constructed using the coding region of

SP-EGFP-Gm28-Gm23K and the following primer sets: primers 1 and 12 for SP-EGFP-28K; primers 1 and 6 for SP-EGFP-Gm23K; and primers 1 and 9 for SP-EGFP. This resulted in, respectively pBI121-SP-EGFP-Gm28-Gm23K, pBI121-EGFP-Gm28-Gm23K, pBI121-SP-Gm28-Gm23K-EGFP, pBI121-SP-EGFP-Gm28K, pBI121-EGFP-Gm28K, pBI121-SP-EGFP-Gm23K and pBI121-SP-EGFP.

These PCR products were digested with *Xba*I and *Sac*I, and were ligated to plant expression vector pBI121. The vectors were then introduced into the *Agrobacterium tumefaciens* strain LBA4404 by electroporation using a Bio-Rad Gene Pulser (Bio-Rad, Hercules, CA).

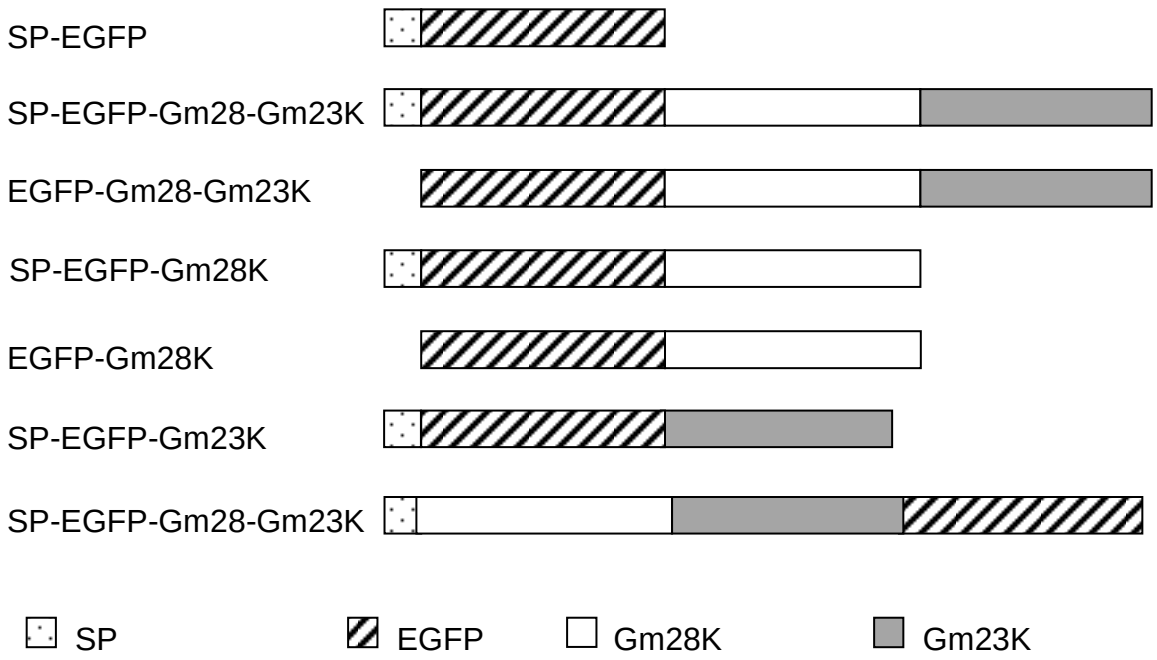

Fig. 1. Schematic models of proteins expressed in tobacco BY2 cells: EGFP, Gm28K and/or 23K with or without signal peptide. Each construct was controlled under the cauliflower mosaic virus 35S promoter, NtADH 5'UTR, and the nopaline synthase terminator in plant expression vectors.

TABLE 2. Sequences of oligonucleotides used as PCR primers used in the present chapter

Primer	Sequence*
1	5’-TTATCTAGAGTCTATTTAACTCAGTATTCA-3’
2	5’-AGTTTTGTTTCCCATTTATTTTTCTTGATT-3’
3	5’-AATCAAGAAAAATAAATGGGAAACAAAACT-3’
4	5’-ATGGAAGGCCATTGTTGTTGTGGCCACTCC-3’
5	5’-ACAACAACAATGGCCTTCCATGATGATGAG-3’
6	5’-TAAGAGCTCAAAAAACATCCATAACCACAT-3’
7	5’-GCCCTTGCTCACCATGGCCATTGTTGTTGT-3’
8	5’-ACAACAACAATGGCCATGGTGAGCAAGGGC-3’
9	5’-TAAGAGCTCTTACTTGTACAGCTCGTCCAT-3’
10	5’-CTCATCATCATGGAACTTGTACAGCTCGTC-3’
11	5’-GACGAGCTGTACAAGTTCCATGATGATGAG-3’
12	5’-TAAGAGCTCAGTTCTCTATCTTCTCATT-3’
13	5’-ACCAGCAGTGTCTTTCTTGTACAGCTCGTC-3’
14	5’-GACGAGCTGTACAAGAAAGACACTGCTGGT-3’
15	5’-TAAGAGCTCTTAAAAAACATCCATAACCAC-3’
16	5’-GCCCTTGCTCACCATTTATTTTCTTGATT-3’
17	5’-AATCAAGAAAAATAAATGGTGAGCAAGGGC-3’
18	5’-GTTATGGATGTTTTTATGGTGAGCAAGGGC-3’
19	5’-GCCCTTGCTCACCATAAAAACATCCATAAC-3’
20	5’-TAATCTAGAATGATTCACACCAACCTGAAG-3’
21	5’-CTCGGTGTTGTCCATCCACACACAGATGAC-3’
22	5’-GTCATCTGTGTGTGGATGGACAACACCGAG-3’
23	5’-TAAGAGCTCCTACTGGGAGCCGGAGTGGCG-3’

*The oliconucleotides used as PCR primers are listed in 5’-3’ orientation.

*Underlines mean the restriction sites of *Xba*I and *Sac*I.

2.1.4 Culture of tobacco BY2 cells

Tobacco BY2 cells were subcultured in modified Linsmaier and Skoog (mLS) medium containing vitamins (0.1 mg of thiamine-HCl and 100 mg of myo-inositol per liter), 3% sucrose, and 0.2 mg/1 2,4-dichlorophenoxyacetic acid (2,4-D), pH5.8, once a week at 25^{o}C in the dark on a reciprocal shaker (*16*).

2.1.5 Transformation of tobacco BY2 cells

Agrobacterium carrying a plant expression vector were inoculated onto 2×YT medium containing antibiotics (50 mg/l kanamycin, 200 mg/l streptomycin, and 50 mg/l rifampicin) to transform tobacco BY2 cells by the Agrobacterium-mediated transformation method (*17*). The transformants were selected on mLS medium (*18*, *19*) containing antibiotics (250 mg/l carbenicillin sodium salt and 150 mg/l kanamycin).

2.1.6 Fluorescent microscopy and laser-scanning confocal microscopy

The transformed BY2 cells were examined on an inverted Leica TCS-SPE confocal laser scanning microscope on a DMI 4000 fluorescent microscope version 2.3 (Leica Microsystems, Mannheim, Germany). The laser was focused on cells through a 40×NA 0.75 or 63×NA 1.4 oil objective. The 488-nm excitation wavelength and a 500-531-nm emission filter (Green) were used to detect EGFP.

To visualize the tonoplast in BY2 cells, 50 μM *N*-(3-triethylammoniumpropyl)-4-(6-(4-(diethylamino)phenyl)exatrienyl)-pyridiniumdibromid e (FM4-64; Molecular Probes Inc., Eugene, OR) was added to the cultured-cell medium. The cells were incubated at room temperature for 1 h and washed three times with fresh mLS medium. A 555-700-nm emission filter (red) was used to detect tonoplast stained with FM4-64 at an excitation wavelength of 532 nm (*20*).

To visualize the Golgi apparatus in BY2 cells, the cells were stained by incubation with

medium containing 5 µM BODIPY TR ceramide (Molecular Probes Inc.) for 30 min on ice and 30 min at 37°C, then washed three times with fresh mLS medium. A 530-592-nm emission filter (red) was used to detect the Golgi apparatus stained with BODIPY TR ceramide at an excitation wavelength of 532 nm (*21*). To visualize the ER in BY2 cells, the cells were also stained for 5 min by incubation in a medium containing 1 µM ER-tracker blue-white DPX (Molecular Probes Inc.) transferred to fresh medium without the marker dyes (*22*). Differential Interference Contrast (DIC) and fluorescence images were taken with an AxioCamMRc digital camera controlled by Axio Vision software release 4.8 (Carl Zeiss, Oberkochen, Germany) on an upright microscope Axioskop 2 (Carl Zeiss) fitted for a 40×NA 0.75 or 100×NA 1.30 oil objective. Fluorescence measurements were performed with Image J software (http://rsb.info.nih.gov/ij/index.html).

2.1.7 Protein extraction

Cells were suspended in 10 mM Tris-HCl, pH 7.5 containing a Complete protease inhibitor cocktail (Roche Diagnostics GmbH, Mannheim, Germany) and homogenized on ice using a Handy Sonic model UR-20P (Tomy Seiko Co., Ltd., Tokyo, Japan). The cell lysate was centrifuged at 12,000 ×*g* for 20 min at 4°C and its supernatant was used as the cell extract.

The medium was filtered through a filter paper to remove the cells, and then the filtrate was centrifuged at 5,000 ×*g* for 15 min at 4°C. The obtained supernatant was treated with four volumes of acetone for 1 h, followed by centrifugation. The precipitate was suspended in 5 mM Tris-HCl, pH 7.0 and then centrifuged at 12,000 ×*g* for 20 min to obtain a solution containing the secreted proteins. The protein concentration in the supernatant was determined using a Bio-RadBradford assay (Bio-RaD).

2.1.8 Immunoprecipitation of expressed proteins

The cell extracts were centrifuged at 12,000 ×*g* for 20 min at 4°C. The proteins in the supernatants were immunoprecipitated with a rabbit anti-GFP polyclonal antibody (Medical & Biological Laboratories Co., Nagoya, Japan), and protein G Sepharose™ 4 Fast Flow (GE Healthcare Bio-Science AB, Uppsala, Sweden). The immunoprecipitates linked with protein G Sepharose were washed 5 times with cold 20 mM sodium phosphate buffer, pH 7.0, resuspended in Laemmli's sample buffer (*23*), and subjected to sodium dodecyl sulfate-poly acrylamide gel electrophoresis (SDS-PAGE) and immunoblotting.

2.1.9 SDS-PAGE and Immunoblotting

Total protein extracts from suspention-cultured tobacco BY2 cells and medium were subjected to SDS-PAGE according to the method of Laemmli (*23*) and transferred to an Immobilon-P polyvinylidene fluoride (PVDF) membrane Immobilon-P (Milliopre Co., Bedford, MA) at 1 mA/cm^2 for 1 h in a semi-dry transfer system (transfer buffer: 25 mM Tris, 190 mM glycine, 20% methanol). The membrane was blocked with 3% skim milk in phosphate-buffer saline, pH7.4 (PBS) containing 0.05% Tween 20 (PBS-T) at room temperature for 1 h. After washing with PBS-T, the membranes were incubated with the primary antibodies against either EGFP (anti-GFP polyclonal antibody) (Medical & Biological Laboratories Co.) or Gm23K (monoclonal antibody (mAb) 1G4) for 1 h at room temperature. After washing with PBS-T, the membranes were incubated with the secondary antibody horseradish peroxidase (HRP)-conjugated anti-rabbit IgG (GE Healthcare UK, Ltd., Buckinghamshire, UK) or HRP-conjugated anti-mouse IgG (GE Healthcare UK) for 1 h at room temperature followed by visualization using Luminata Forte Western HRP substrate (Millipore Co.) according to the manufacturer's instruction. For detection of the proteins immunoprecipitated with protein G Sepharose, SDS-PAGE and Immunoblotting were performed in the same manner as described for the experiments with the total proteins, except

that mAb anti-GFP (Medical & Biological Laboratories Co.) was used as a primary antibody and HRP-conjugated anti-mouse IgG was used as a secondary antibody.

2.2 RESULTS

2.2.1 Identification of the 5'-nucleotide sequence of cDNA encoding the Gly m Bd 28K precursor.

In the previous study, it was reported that the sequence of cDNA encoding the Gm28K (*8*). However, the complete N-terminal sequence of the Gm28K precursor remains to be characterized. To determine the complete sequence of the Gm28K precursor, 5'RLM-RACE was performed using the total RNA obtained from developing soybean cotyledons. Consequently, the new nucleotide sequence identified in the present study was revealed to consist of a nucleotide sequence corresponding to 24 amino acids and a 5' untranslated nucleotide sequence with 17 nucleotides, as shown in Fig. 2. The Gm28K would be produced as a precursor with a signal peptide of 24 amino acid residues including MGN. Furthermore, a stop codon, TAA, was located 45 bases upstream of the methionine residue, and there was no methionine within this 45 bases in the genomic sequence from soybean. The full-length sequence of the Gly m Bd 28K precursor is shown in Fig. 2.

We tried homology search to the amino acid sequence of the full-length Gm28K precursor (Table 3). The protein showing high identity of 65.8% with the Gm28K precursor amino acid sequence were revealed to occur in the *Medicago truncatula* is a model for legume plant. In addition, the precursor was also highly homologous with 2-2-like vicilin-like antimicrobial peptides in various plants, for example, chick peas, soybean, barrel medic, European grapes, and cucumbers.

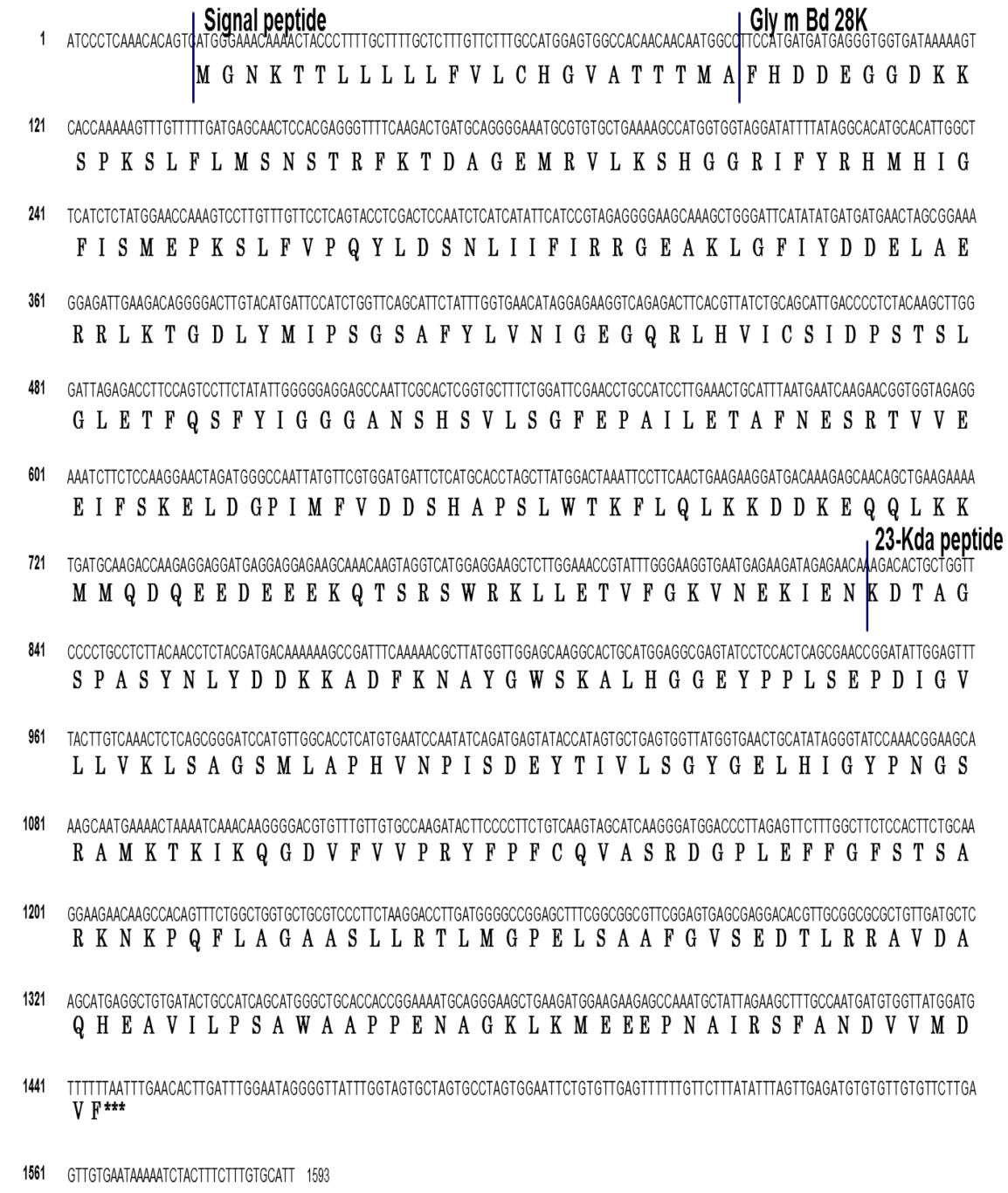

Fig. 2. Nucleotide sequence encoding the Gm28K precursor and its deduced amino acid sequence.

Table 3 Proteins hits in pairwise sequence alignment with the amino acid sequence encoding the Gm28K precursor.

	Protein names	organism	Accession no.	Identity (%)	Similarity (%)	Gaps (%)	Score
1	Allergen Gly m Bd	*Medicago truncatula*	XP_003606476.1	65.8	79.1	1.9	1585
2	PREDICTED: vicilin-like antimicrobial peptides 2-2-like	*Cicer arietinum* (chickpea)	XP_004505960.1	63.7	76.1	2.9	1490
3	PREDICTED: vicilin-like antimicrobial peptides 2-2-like	*Glycine max* (soybean)	XP_003542001.1	59.3	74.3	2.7	1432
4	Vicilin-like antimicrobial peptides 2-2	*Medicago truncatula*	XP_003597178.1	56.3	70.3	7.6	1307
5	Vicilin-like antimicrobial peptides 2-3	*Medicago truncatula*	XP_003597162.1	55.9	70.3	8.2	1286
6	PREDICTED: vicilin-like antimicrobial peptides 2-2	*Vitis vinifera* (common grape vine)	XP_002264047.1	53.5	69.8	3.3	1237
7	predicted protein	*Populus trichocarpa*	XP_002313331.1	53.1	66.7	12.7	1164
8	nutrient reservoir, putative	*Ricinus communis*	XP_002526353.1	52.7	67.7	6.8	1197
9	7S globulin precursor	*Ficus pumila* var. *awkeotsang*	ABK80758.1	52.0	67.9	8.5	1235
10	RmlC-like cupins superfamily protein, putative	*Theobroma cacao*	EOY33094.1	51.9	70.0	5.4	1266
11	probable major protein body membrane protein MP27 / major protein body protein MP32 precursor - cucurbit	*Cucurbita* cv. *Kurokawa Amakuri* (pumpkin)	T10443	50.4	65.1	5.8	1173

2.2.2 Functional analysis of the N-terminal region of Gly m Bd 28K precursor in tobacco BY2 suspension-cultured cells.

The transport of soybean Gm28K is assumed to be transported through the ER and Golgi apparatus because Gm28K has N-linked sugar chain with α1,3-fucose and β1,2-xylose residues (*7*). Therefore, we tried to examine the function of the SP of the Gm28K precursor. EGFP was used as a reporter protein and was fused to SP, Gm28K, Gm23K, and Gm28-Gm23K (Fig. 1). Each construct was stably expressed in tobacco BY2 cells and the detailed localization was imaged by a fluorescent microscope.

When the EGFP gene was introduced into tobacco BY2 cells, the EGFP was localized in

the cytoplasm and the nucleus (Fig. 3A). On the other hand, in the case of the 4-day-old SP-EGFP/BY2 cells, EGFP was mainly localized in the ER (Fig. 3B). However, in the 8-day-old SP-EGFP/BY2 cells, the EGFP was partially transported to the vacuoles, although most of the EGFP fluorescence was still found in the ER (Fig. 3B). Figure 3C also shows that EGFP was predominantly localized in both the ER and vacuoles of 8-day-old SP-EGFP/BY2 cells whose tonoplasts were stained with FM4-64. In an immunoblot of total soluble proteins in 4-day-old SP-EGFP/BY2 cells, the anti-EGFP antibody-positive band was detected at the same position (27 kDa) as that obtained from 4-day-old EGFP/BY2 cells (Fig. 4A). The protein with a molecular mass of 27 kDa was EGFP, and thus these findings indicate that the co-translational cleavage of the SP of SP-EGFP results in EGFP. No band was detected in the immunoblot obtained with empty vector-transformed BY2 cells and the 7S globulin fraction from soybean containing Gm28K and Gm23K (Fig. 4A). However, the immunoblot of the culture medium of 10-day-old SP-EGFP/BY2 cells showed that a large amount of EGFP (27 kDa) was secreted into the medium, although the medium from 4-day-old SP-EGFP/BY2 cells did not contain any EGFP (data not shown). These data suggest that the SP region of the Gm28K precursor functions in intracellular trafficking and secretion of EGFP.

2.2.3 Effect of the Gm28K precursor on transport in tobacco BY2 suspension-cultured cells

We expressed Gm28K and Gm23K fused to SP-EGFP (SP-EGFP-Gm28-Gm23K) to investigate the behaviors of Gm28K and Gm23K region in tobacco BY2 cells. Fig. 5A shows the localization patterns of EGFP in SP-EGFP-Gm28-Gm23K/BY2 cells during the indicated days after subculture. In the SP-EGFP-Gm28-Gm23K/BY2 cells, the produced protein was mainly found in the ER from the 1st to the 3rd day but was not observed in vacuoles. In contrast, a remarkable accumulation of EGFP-Gm28-Gm23K was observed in vacuoles of 4-day-old

SP-EGFP-Gm28-Gm23K/BY2 cells and the main localization site of EGFP-Gm28-Gm23K shifted from the ER to the ER/vacuoles after 1 week. The vacuolar localization of EGFP in SP-EGFP-Gm28-Gm23K/BY2 cells was confirmed by visualization of the vacuole membrane with FM4-64 (Fig. 5B). The localization pattern of EGFP in EGFP-Gm28-Gm23K/BY2 cells

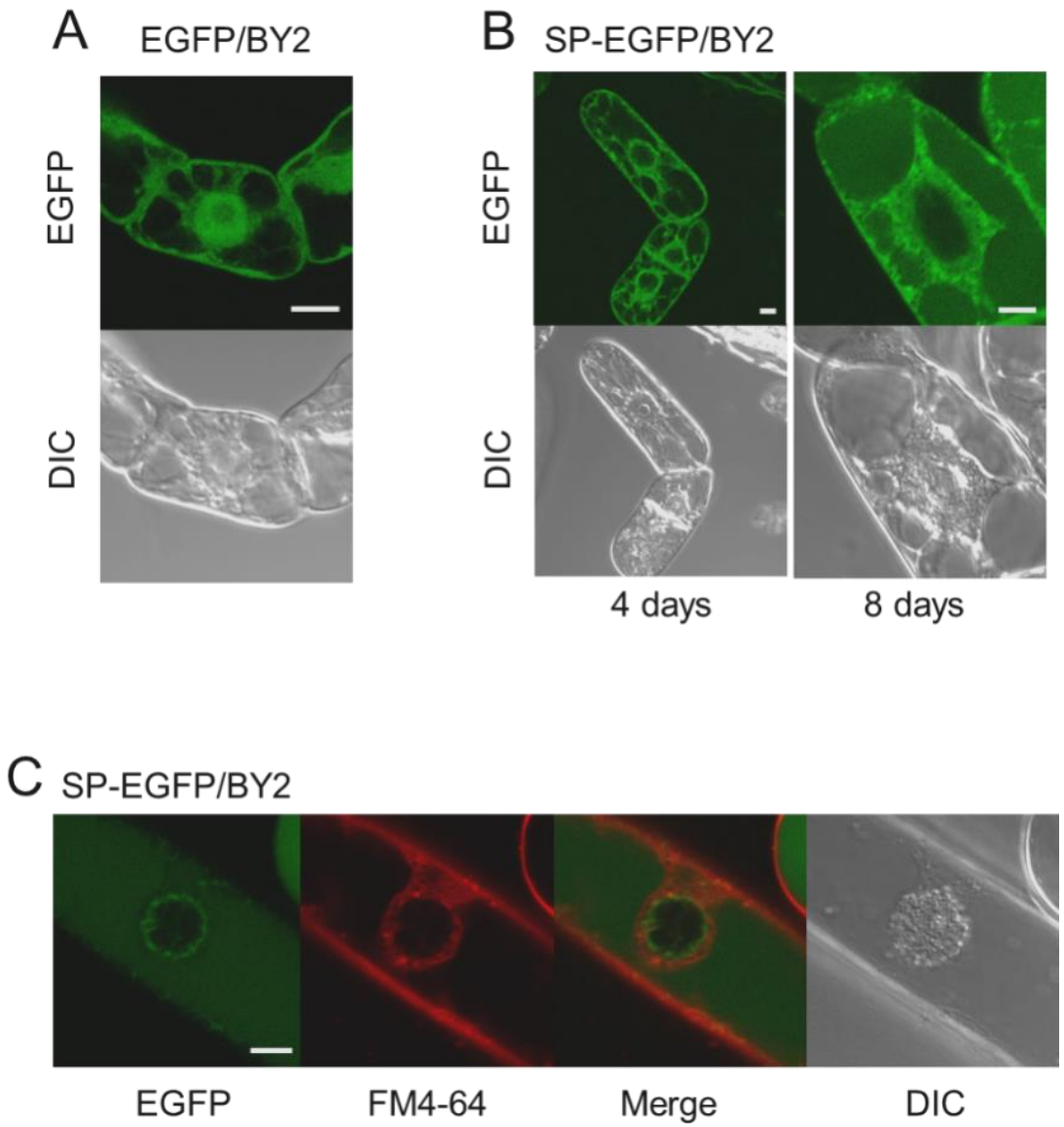

Fig. 3. Localization of EGFP in EGFP/BY2 cells and SP-EGFP/BY2 cells. A: Confocal images of 8-day-old cells stably expressing EGFP in the EGFP/BY2cells; B: Confocal images of 4-day-old and 8-day-old cells stably expressing EGFP in the SP-EGFP/BY2 cells; C: Confocal images of 8-day-old cells stably expressing EGFP in the SP-EGFP/BY2 cells staining with FM4-64. Bars represent a length of 10 μm.

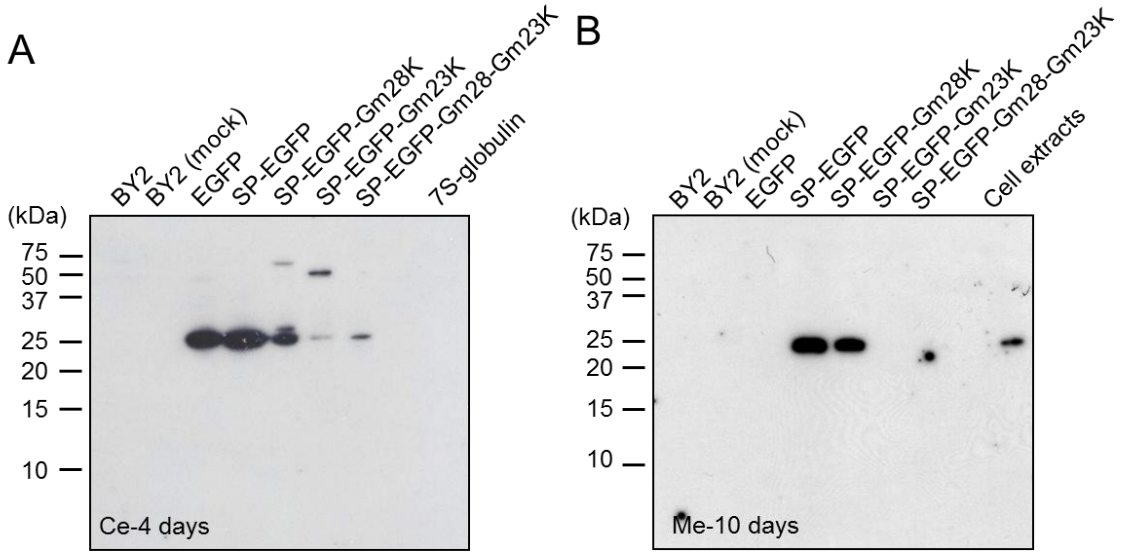

Figure 4. Immunoblotting of recombinant proteins in the cells and the cultured media using anti-GFP antibody. A: Four-day-old cell extracts from the indicated transformants. B: Ten-day-cultured media from the indicated transformants. Mock: BY2 cells transformed with the pBI121-empty vector; 7S-globulin: 7S-globulin fraction from soybean seeds; Cell extracts: Cell extracts of EGFP/BY2 cells in panel A; Ce: Blot of the proteins in the cell extracts; Me: Blot of the proteins in the media.

was similar to that in EGFP/BY2 cells (Fig. 5C). Immunoblotting with an anti-GFP antibody and the total soluble proteins from 4-day-old SP-EGFP-Gm28-Gm23K/BY2 cells showed that a 27-kDa protein was detected but that the EGFP-Gm28-Gm23K with a molecular mass of 78 kDa could not be detected (Fig. 4A). Therefore, the immunoprecipitation with an anti-GFP antibody was applied to concentrate EGFP-Gm28-Gm23K in the cells. A protein band with a molecular mass of 78 kDa corresponding to that of EGFP-Gm28-Gm23K was observed in the immunoblot with the immunoprecipitated sample (Fig. 6A). The putative EGFP-Gm28-Gm23K was also detected with the anti-Gm23K antibody prepared in present study (Fig. 6B). On the other hand, Gm23K was also detected in the cell extracts with anti-Gm23K antibody (Fig. 6C). This finding suggests that the EGFP-Gm28-Gm23K wasproteolytically cleaved between Gm 28K and Gm23K in SP-EGFP-Gm28-Gm23K/BY2 cells. No EGFP-Gm28-Gm23K was found in the culture medium (Fig. 4B).

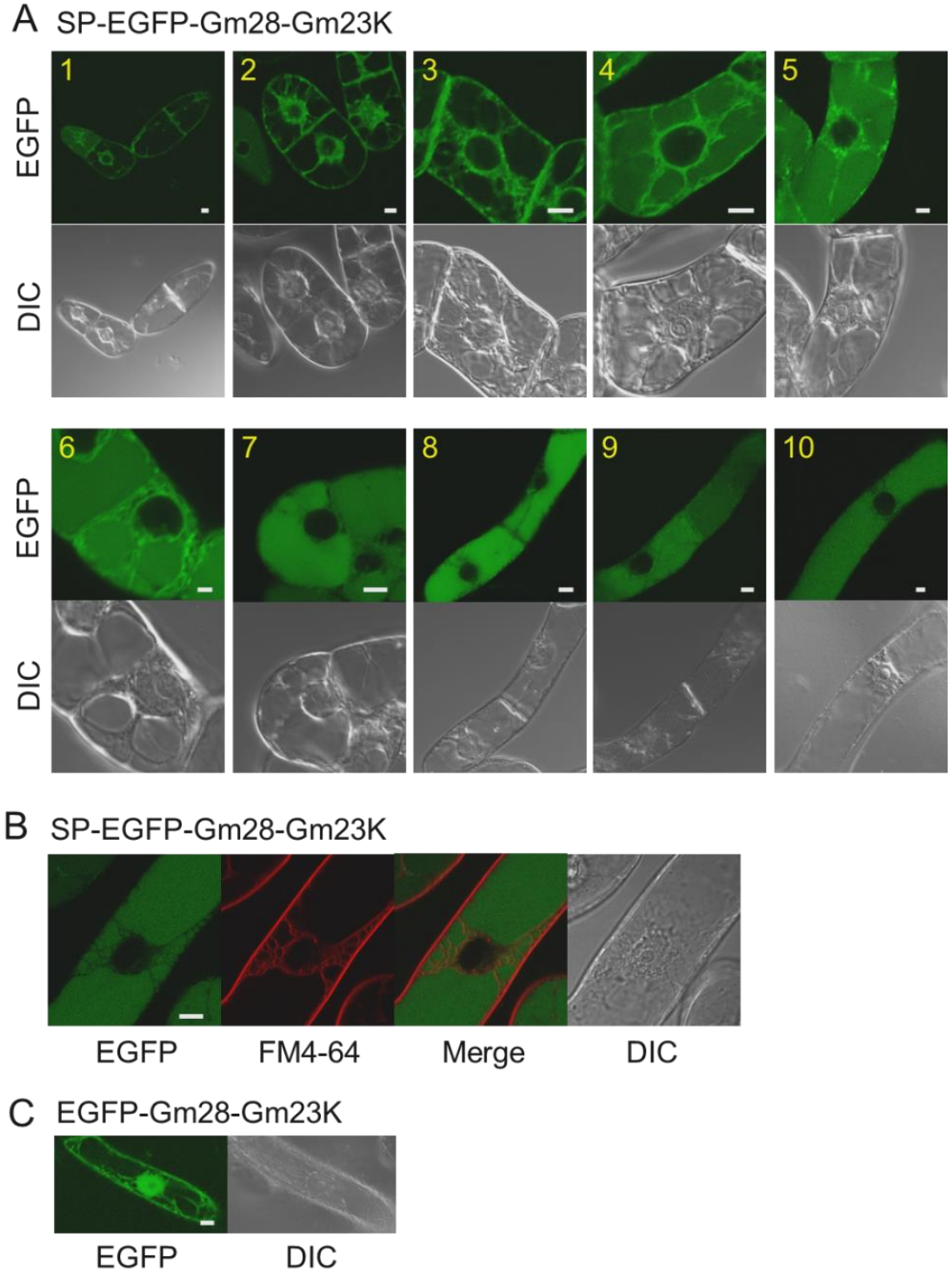

Figure 5. Localization of EGFP-Gm28-Gm23K with or without SP in BY2 cells. A: The time-dependent accumulation of EGFP-Gm28-Gm23K in SP-EGFP-Gm28-Gm23K/BY2 cells cultured for indicated days; B: Colocalization image of SP-EGFP-Gm28-Gm23K/BY2 cells with FM4-64 in 8-day-old cells; C: EGFP-Gm28-Gm23K/BY2 Cells. Bars represent a length of 10 μm.

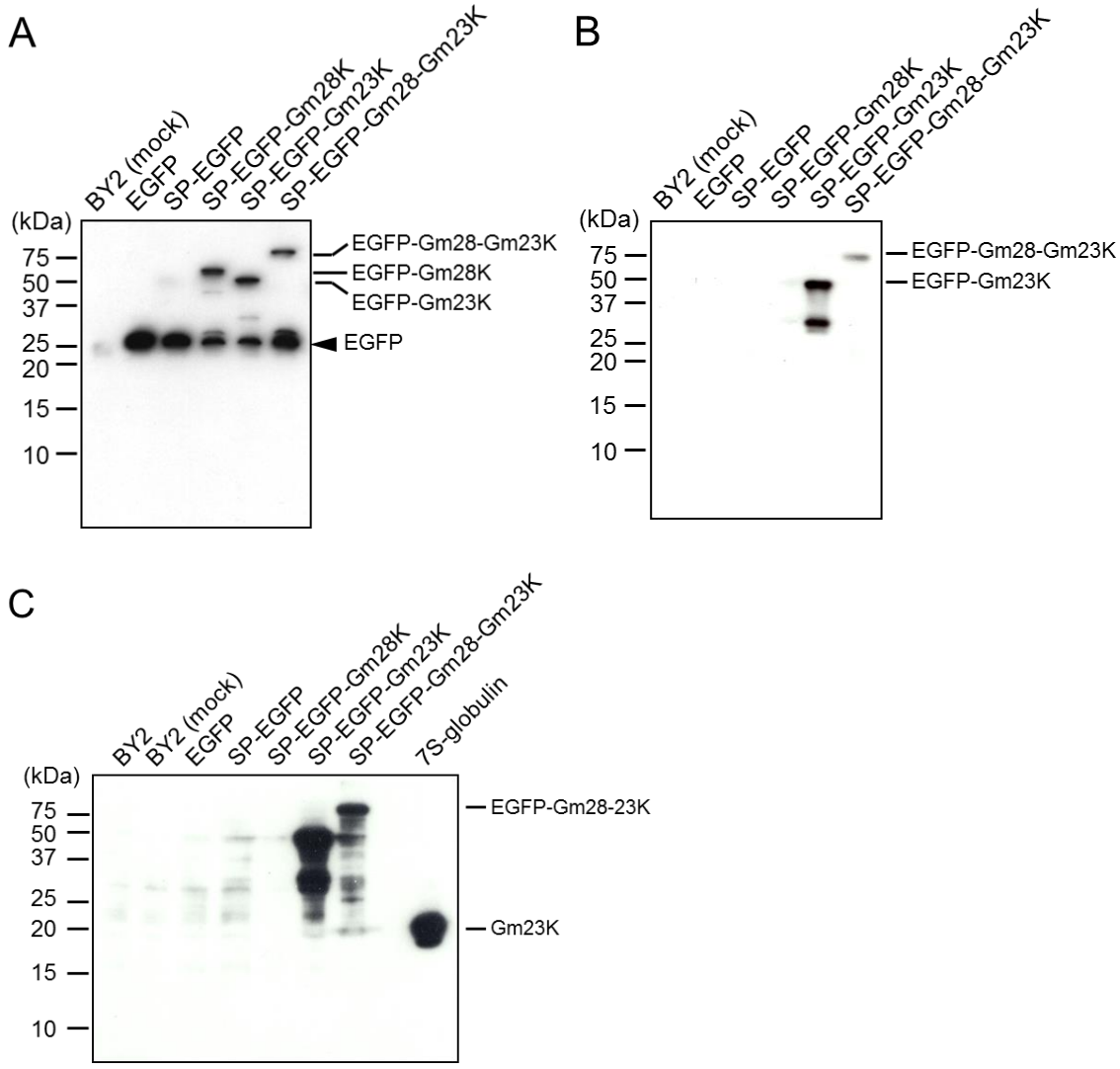

Figure 6. Immunoblotting of recombinant proteins and their derivatives in the indicated transformants. A: Immunoprecipitates using anti-GFP antibody; B: Immunoprecipitates using anti-Gm23K antibody; C: Cell extracts from the indicated cell lines stained with anti-Gm23K antibody.

Further, in order to investigate the influence of EGFP bound to the N-terminus of Gm28K on the transport of the precursor, we tried to express SP-Gm28-Gm23K-EGFP, in which EGFP was fused to the C-terminal of Gm23K. The fluorescence of EGFP was accumulated in the vacuoles of SP-Gm28-Gm23K-EGFP/BY2 cells, which was similar to the localization observed with 10-day-old SP-EGFP-Gm28-Gm23K/BY2 cells (data not shown). In

SP-Gm28-Gm23K-EGFP/BY2 cells, immunoblotting analysis using an anti-EGFP and an anti-Gm23K antibody showed that EGFP was not secreted into the medium. These results suggest that EGFP is not involved in vacuolar transport of the precursor. The above observations clearly demonstrate that both EGFP-Gm28-Gm23K and Gm28-Gm23K-EGFP are transported to the vacuoles.

Next we attempted to prepare SP-EGFP-Gm28K/BY2 cells and SP-EGFP-Gm23K/BY2 cells to examine the contribution of the Gm28K or Gm23K region to vacuolar sorting of the Gm28K precursor in tobacco BY2 cells. The fluorescence of EGFP was very weak in the vacuoles of the SP-EGFP-Gm28K/BY2 cells as compared to that in the vacuoles of SP-EGFP-Gm28-Gm23K/BY2 cells (Fig. 5A, Fig. 7). An immunoblot of the total soluble proteins from 4-day-old SP-EGFP-Gm28K/BY2 cells revealed the presence of EGFP-Gm28K (55 kDa) (Fig. 4A). However, the SP-EGFP-Gm28K/BY2 cells secreted a large amount of EGFP into the 10-day-old culture medium, which was similar to the result for SP-EGFP/BY2 cells (Fig. 4B).

An NPIRL motif derived from sweet potato sporamin has been shown to function as a sequence-specific VSD, and its detailed mechanisms have been analyzed in tobacco BY2 cells (*24*). The Gm28K precursor has an NPIRL-like motif, NPISD, around residue 330 of Gm23K. Therefore, we examined the function of Gm23K as vacuolar-sorting determinant (VSD). We expressed an SP-EGFP-Gm23K without Gm28K in tobacco BY2 cells. The strong and stable fluorescence of EGFP was localized in the ER lumen of the SP-EGFP-Gm23K/BY2, but not in the vacuoles, Golgi, endosomes or cell wall (Fig. 8). This distribution remained constant over the 10-day culture (data not shown). An immunoblot of total soluble proteins from the 4-day-old SP-EGFP-Gm23K/BY2 cells demonstrates that EGFP-Gm23K (50 kDa) was detected with anti-GFP antibody and anti-Gm23K monoclonal antibody (Fig. 4A). Moreover, in an immunoblot with the anti-Gm23K antibody, a degraded protein derived from the EGFP-Gm23K region with a molecular mass of approximately 30 kDa was observed (Fig. 6B).

The EGFP was not secreted into the medium of SP-EGFP-Gm23K/BY2 cells (Fig. 4B).

These findings show that EGFP cannot be transported from the ER to the Golgi under the conditions examined here. Thus, it is suggested that the Gm28K-Gm23K region of the Gm28K precursor plays an important role in vacuolar-sorting in tobacco BY2 cells.

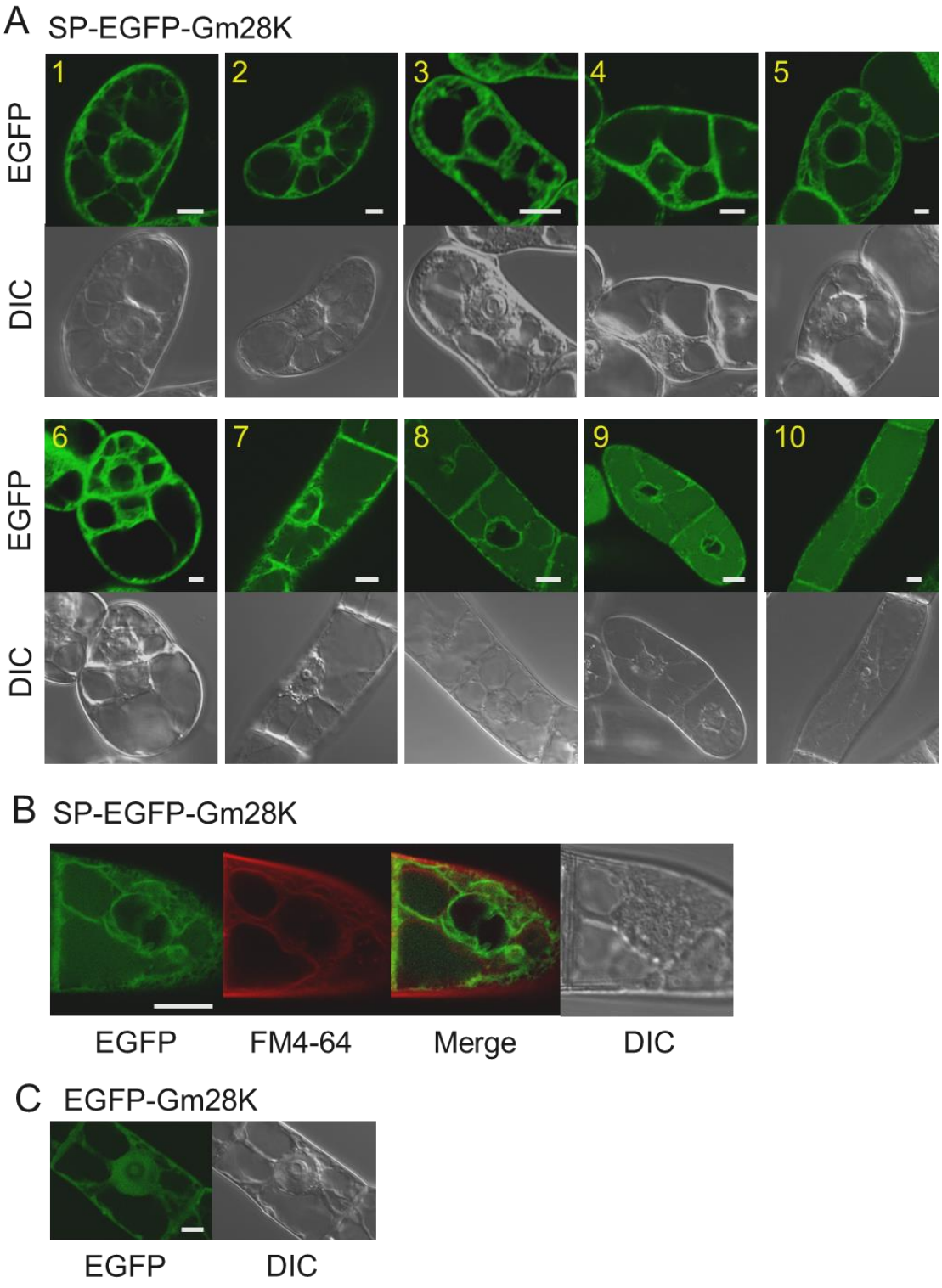

Figure 7. Localization of EGFP-Gm28K with or without SP in BY2 cells. A: The time-dependent accumulation of EGFP-Gm28K in SP-EGFP-Gm28K/BY2 cells cultured for indicated days; B: Colocalization image of SP-EGFP-Gm28K/BY2 cells with FM4-64 in 8-day-old cells; C: EGFP-Gm28K/BY2 Cells. Bars represent a length of 10 μm.

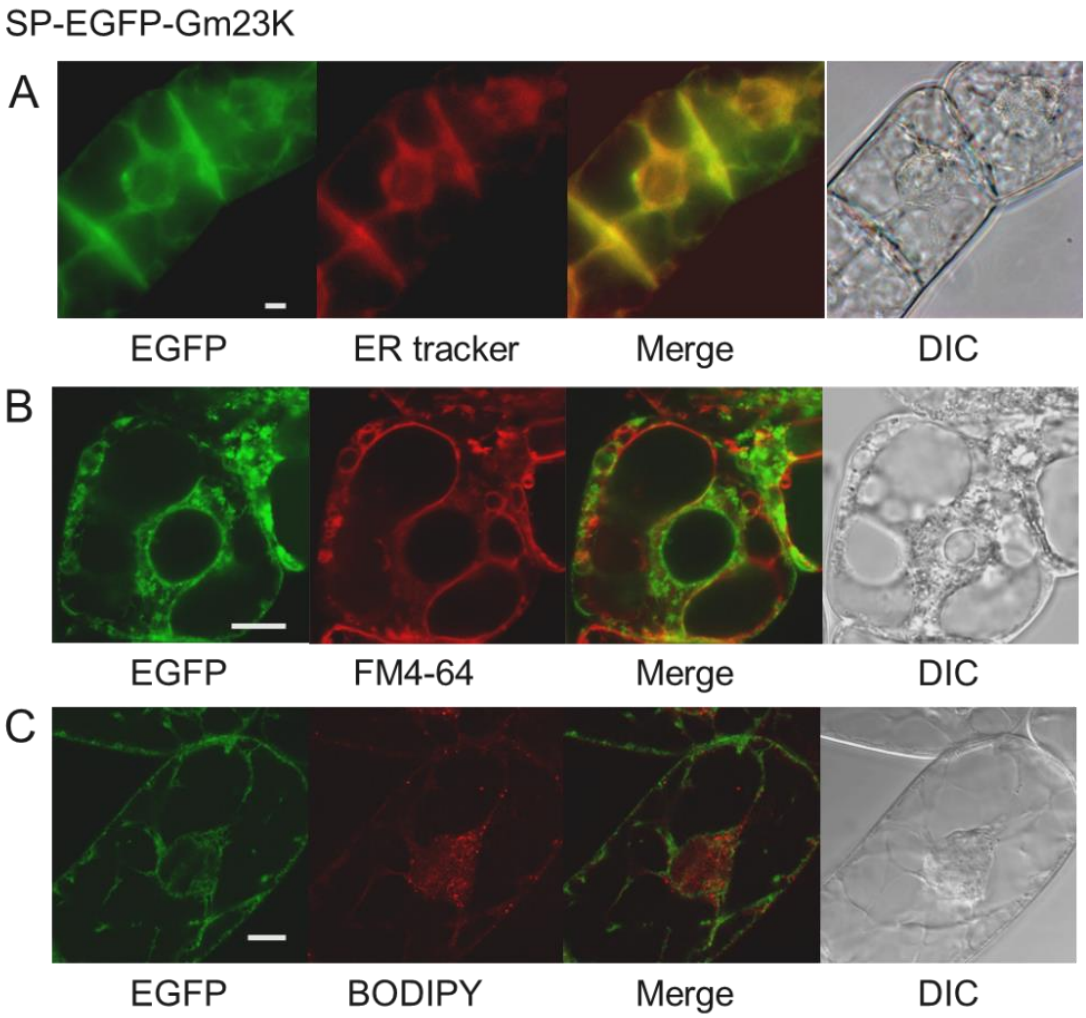

Figure 8. Fluorescence images of SP-EGFP-Gm23K cells. A: The fluorescence images of cells with ER tracker; B: with FM4-64; C: with BODIPY. Eight-old-day cells were analyzed. Bars represent a length of 10 μm.

2.3 DISCUSSION

Previous reports have suggested that Gm28K may be translated as a preproprotein which would be composed of a N-terminal peptide (SP), Gm28K and C-terminal peptide (Gm23K) and then may become each mature protein. However, the N-terminal sequence of its precursor had not been fully characterized. In the present study, we completely identified amino acid sequence of the Gm28K precursor including an SP consisting of 24 amino acids.

A typical signal peptide for transport into the ER lumen has positively charged amino acids followed by 7-15 hydrophobic amino acids and polar amino acid (*25*). In addition, the third amino acid residue from the N-terminal side of the cleavage site is usually alanine or leucine. The SP of the Gm28K precursor also has characteristics similar to the general rule. The N-terminal peptide and Gm28K are cleaved between the Phe25 residue and Ala24 residue of the Gm28K precursor. This suggests that a peptide of 24 amino acids might be required for transport of the Gm28K precursor to the ER lumen and function as a signal peptide in soybean.

What is the actual function of the SP of the Gm28K precursor? To address this question, we attempted to express SP-EGFP-fused proteins in tobacco BY2 cells. The main location of EGFP in SP-EGFP/BY2 cells was the ER and small amounts of EGFP were accumulated in the vacuoles. Moreover, immunoblotting data shows that the SP was cleaved from SP-EGFP (Fig. 4A) and that a large amount of EGFP was secreted into the medium of 10-day-cultured cells (Fig. 4B). These findings suggest that the SP leads proteins to the default secretory pathway. In the other words, the SP acts as a signal peptide for the translocation of the Gm28K precursor into the ER lumen. In SP-EGFP/BY2 cells, EGFP was secreted into the culture medium when SP expressed as the EGFP fusion protein. This result suggests that Gm28K might be also secreted into the culture medium when the fusion protein of the SP and Gm28K is expressed in tobacco BY2 cells.

Therefore, we expressed SP-Gm28-Gm23K-EGFP and SP-EGFP-Gm28K-Gm23K

constructed by fusion of the Gm28K precursor with SP-EGFP, and we analyzed the localization of EGFP using a confocal laser microscope and immunoblotting. The fluorescence of EGFP was observed in the vacuoles in SP-Gm28-Gm23K-EGFP/BY2 cells and SP-EGFP-Gm28K-Gm23K/BY2 cells. Furthermore, the culture media of these transgenic BY2 cells were examined with immunoblotting using anti-GFP antibody. EGFP was not secreted into the culture medium. The above results suggest the possibility that Gm28K-Gm23K sequence is required for the transport to the vacuoles.

Vacuolar proteins are transported to the different types of vacuoles through a vesicle-mediated biosynthetic trafficking pathway that includes the ER, the Golgi apparatus, the TGN and the endosomes/prevacuoles (26). Two types of vacuoles, the PSV and the lytic vacuole, can be found in plant cells (26). Gm28K is a glycoprotein, and it has been reported that the sugar chain is composed of mannose, *N*-acetylglucosamine, xylose, and fucose in a molar ratio of 3:2:1:1 (6). The fact that the allergen has xylose and fucose residues shows that Gm28K was subjected to glycosylation in the Golgi apparatus.

The proteins are transported to the vacuole depend on the VSDs, sequence-specific VSDs (ssVSDs) (*27, 28*), C-terminal VSDs (ctVSDs) (*29-32*) and protein-structure-related VSDs (*33*). The ssVSDs of sporamin, sweet potato tuber protein, has been characterized in detail for BY2 cells (*24, 34*). This motif consists of 5 amino acids, NPIRL, and even if the arginine and leucine residues are replaced by any amino acid residue and any high hydrophobic amino acid residue, respectively, its vacuolar-sorting function was not lost in tobacco BY2 cells (*24*). Ricin and 2S albumin in *Ricinus communis* have an ssVSD similar to sporamin (*35*). The ctVSDs are located in the C-terminal regions of polypeptides. For example phaseolin of *phaseolus vulgaris* (*36*), 2S albumin of *Bertholleticaexcelsa*, α' and α subunits of soybean β-conglycinin have a ctVSD (*38, 39*). A similarity of the ctVSD sequence between common bean phaseolin and soybean β-conglycinin was found (*38*). The 10 C-terminal amino acids of the α' subunit of soybean β-conglycinin contain ssVSD-like sequence, in addition to ctVSD

(*31*). It has been reported that their amino acid sequences are Ser-Ile-Leu and the C-terminal three amino acids (AFV) (*31, 37, 38*). The 11S globulin of field bean (legumin) and soybean have a psVSD (*40, 41*).

In SP-EGFP-Gm28K/BY2 cells, a large amount of EGFP was secreted into medium similar to the case of SP-EGFP/BY2 cells. A C-terminal amino acid sequence often has an important signal for protein trafficking such as the ER retention signal (*42, 43*), membrane association via lipidation (*44*), vacuolar sorting (*29-32*), and so on. However, the Gm28K has neither their specific nor similar vacuolar-sorting motif at its C terminal. The Gm28K precursor contains an NPIRL-like motif, NPISD, on the Gm23K segment. If the NPISD acts as a vacuolar-sorting signal, EGFP-Gm23K should be transported to the vacuoles of the SP-EGFP-Gm23K/BY2 cells. However, confocal images and an immunoblotting analysis resulted in the accumulation of EGFP-Gm23K to the ER and no secretion into the cultured medium. The alteration of SP-EGFP-Gm23K structure by fusion of Gm23K may induce its incorrect folding and accumulation in the ER. As shown in Fig. 6C, the detection of the 30-kDa protein shows that an unpredicted cleavage of SP-EGFP-Gm23K occurred on the Gm23K part in the ER. It was suggested that the ER-accumulation of EGFP-Gm23K inhibited its transport to the Golgi followed by trafficking to the vacuoles or secretion to the extracellular compartment.

The seed storage proteins in the 7S globulin, common bean phaseolin and jack bean canavalin, form their multimers (*45, 46*). For example, common bean phaseolin have a ctVSDs consisted of a four-amino acid hydrophobic propeptide. The GFP fused to the ctVSDs of phaseolin has an ability to carry GFP to the vacuole. However, when phaseolin was expressed as a monomer, it bound to the binding protein (a member of the heat-shock 70 family) and was degraded by ER-associated degradation (ERAD). On the other hand, it was shown that the wild-type formed trimers in the ER to transport them to the vacuole (*47*). Some newly synthesized proteins that cannot fold correctly are degraded by ERAD (*48, 49*). Gm28K

and Gm23K may be required for the multimerization of the Gm28K precursor in the ER and then Gm28K and Gm23K may be also transported to PSV by VSDs through the Golgi apparatus.

Furthermore, as a result of a homology search for the amino acid sequence of the full-length Gm28K precursor, the Gm28K precursor was shown to have high homology with the vicilin-like antimicrobial peptides. The majority of plant food allergens can be classified into just 4 protein families, the prolamin superfamily, the cupin super family, the Bet v 1 superfamily, profilins (*50*). The vicilin belongs the cupin superfamily, and Gm28K also belongs to its family. Cupin motif is found from various proteins of animal, plant, eukaryote, bacteria, and archaea. The protein structure is conserved among these species and various functions have been demonstrated (*51*). The p16 which belongs cupin superfamily, a small protein partitioned between the chromatin and protein bodies of pea (*Pisum sativum*) seeds, is encoded by psp54 (*52*), a large gene whose mRNA would be able to yield a polypeptide of a 54.4 kDa and which has been named p54 (*53*). It has been proposed that one of the roles of the p16 may be to protect seed chromatin from desiccation (*53*). Also the Gm28K precursor might have a function, such as defense against stress proteins in soybean, because it shows high homology with the vicilin-like antimicrobial peptides.

In the present study, we revealed that at least the identified 24 amino acid segment on the N-terminus of the Gm28K precursor functions as a signal peptide to enter the ER lumen on its vacuolar-sorting pathway. After transport through the ER and the Golgi, the Gm28-Gm23K is probably transported to the vacuoles. EGFP is secreted into the medium of SP-EGFP-Gm28K/BY2 cells, suggesting that Gm23K is required for the transport of the Gm28K precursor to the vacuoles. Although, in the SP-EGFP-Gm23K/BY2 cells, the fluorescence was accumulated in the ER but not in the vacuoles, some kinds of VSDs on the Gm23K segment may be recognized by its receptor in the Golgi apparatus after leaving from the ER. These findings suggest that the Gm28K and Gm23K are transported to the soybean

vacuoles from the ER lumen via the Golgi.

Chapter 3

Identification of the 23-kDa peptide derived from the precursor of Gly m Bd 28K and Characterization of the 23-kDa peptide

Gly m Bd 68K, Gly m Bd 30K and Gm28K are the major allergens in soybean. The former two allergens have been well recognized. Gm28K was purified and its properties were well eluchidated.

Gm28K is a glycoprotein with N-linked glycan moiety. It has been demonstrated that its glycan is involved in the binding to IgE antibodies in the sera of the soybean-sensitive patients (*7*). The N-linked glycans of plant glycoproteins are covalently linked to the asparagine residues of a common consensus sequence in polypeptides, Asn-X-Ser/Thr, and its core structure is Man3-GlcNas2-Asn. The plant glycoproteins often contain α1,3-fucose and β1,2-xyose which are rarely found in animal glycoproteins. Many studies reported that rabbits immunized with plant specific glycoproteins produces antibodies recognizing sugar chain (*54-56*). The IgG antibody was shown to have cross-reactivity glycoproteins with the sugar chain of glycoproteins as common structural elements (*55, 57, 58*). On the other hand, the sugar chains of Tri a Bd 27K, a wheat allergen (*59, 60*), and Ole e 1, the major pollen allergen of olive, were reported to be involved in binding to IgE antibodies (*57, 61*). The possibility that the sugar chains of the wheat allergen and olive allergen may be involved in their allergenicities propose the idea that the sugar chain may act as a common antigen, because many plant glycoporotein allergens might have the sugar chains similar to the above-mentioned core structure.

However the relationship between allergenicities and their sugar chains has not been revealed. We hope that the relationship between the chemical structures of the sugar moieties of plant allergens and binding to IgE antibodies in the sera of allergic patients will be well elucidated and that the idea that the sugar chains can act as a common antigen will be verified.

A cDNA encoding Gm28K has already been cloned but the cDNA was shown to encode the precursor of Gm28K. In the preceding chapter, the full-length cDNA has been elucidate and it has been demonstrated that three segments, SP, Gm28K and Gm23K. It was shown that the SP act a signal peptide and the Gm28K-Gm23K may be transported to the vacuole. Whether Gm23K occurs in soybean seeds has never been examined. If it would occur in

soybean seeds, what is the properties of Gm23K? Furthermore, how is the Gm28K precursor metabolized? In the present chapter, the occurrence of Gm23K in soybean seeds was examined. The properties of Gm23K and the fate of the precursor were elucidated.

3.1 MATERIALS AND METHODS

3.1.1Plant materials

Gm28K was purified from defatted soybean flakes (*Glycine max*, mixture of IOM varieties) in the same manner as described in the previous paper (*6*). The defatted flakes were gifted from Fuji Foundation for Protein Research (Osaka, Japan). Soybean seeds (*Glycine max*, L. cv. Shishioh) were purchased from Takii Co. (Kyoto, Japan). For preparation of RNAs and proteins in the developing cotyledons of soybean, soybean seeds were grown in a garden of Okayama Prefectural University from May to September. Cotyledons of developing seeds after flowering were harvested at several intervals and immediately frozen in liquid nitrogen. The seeds were used for preparation of their proteins and the mRNAs as described below.

3.1.2. Sera of patients sensitive to soybean

The sera of patients sensitized with soybean (radioallergosorbent test (RAST) value 1–4) were obtained from National Sanatorium Minami-Okayama Hospital in Okayama and Shimada Clinic in Tokushima. The informed consents were obtained from the patients involved in the present study. The experiments were done in accordance with the ethical standards as formulated in the Helsinki Declaration.

3.1.3. Analysis of the N-terminal amino acid sequence of the mAb-positive peptide

Proteins obtained by extraction from defatted soybean were separated on a gel by 2D-PAGE and electrophoretically transferred onto a PVDF membrane (Immobilion P, Millipore Co., USA), as described above. The spots corresponding to the peptides recognized by the mAb against the recombinant Gm23K were cut off and N-terminal amino acid sequences of the peptides were analyzed with an Applied Biosystem 473 protein sequencer.

3.1.4. Production of Gm23Kusing Escherichia coli

The construction of an expression system of Gm23K was performed using pET-21d (+) vector (Merck Co., Germany). A DNA fragment encoding Gm23K was amplified by PCR with cDNA encoding the Gm28K precursor, as a template. Primer 24 (5′-GAGAAGATAGAGAACAAAGACAC-3′) and primer 25 (5′-TAATACGACTCACTA-TA-3′) were used as the sense sequence for EKIENKD, the sequence corresponding to 265–271 amino acid sequence and the antisense sequence, which corresponds to T7 promoter in a cloning vector, pBluescript II SK(−) inserted cDNA encoding Gm28K precursor, respectively. The PCR product was digested with *Xho*I. The expression vector pET-21d (+) was digested with *Nco*I, followed by polishing with Klenow treatment using a blunting kit (TaKaRa-Shuzo, Japan), and digested with *Xho*I. The DNA fragment was ligated into the *Nco*I and *Xho*I site of the pET-21d and an expression system was constructed. Gm23K-pET-21d was transformed into *Escherichia coli* BL21 (DE3) and then cultured in Luria–Bertani medium at 37 °C. The recombinant protein was expressed in the presence of 1 mM isopropyl-1-thio-β-D-galactopyranoside and formed as inclusion bodies. The inclusion bodies were sonicated and resuspended in 6 M urea containing 0.2 M 2-mercaptethanol. The recombinant protein was used as an immunogen for the immunization of mice and for the preparation of a mAb against the recombinant protein.

3.1.5. Preparation of a mAb against the recombinant Gm23K

Preparation of a hybridoma producing mAb against the recombinant Gm23K was performed according to the method described previously (*60*). The hybridoma obtained was intraperitoneally injected into female BALB/c mice primed with pristane. After 10–14 days, ascitic fluids formed were collected. The mAb in the ascitic fluids was fractionated by collection of the precipitate formed with ammonium sulfate between 0% and 50% saturation. The mAb was further purified using an Affi-gel Protein A MAPS II kit (Bio-Rad, USA). The obtained mAb is named 1G4 in the present study. The present experiment with the animals has been done according to the National Research Council's guide for the care and use of laboratory animals.

3.1.6. Screening of proteins in defatted soybean flakes recognized by the mAb against the recombinant Gm23K

Proteins in defatted soybean flakes were extracted with 10 mM Tris–HCl buffer including 2 mM 2-mercaptethanol, pH 8.0 by a Polytron homogenizer (PUC-11, Kinematica). The homogenate was squeezed with cheesecloth, and the filtrate was centrifuged at 10,000×*g* for 20 min, 4 °C. The amounts of the proteins in the supernatants were assayed by the method of Lowry *et al.* (*63*) using bovine serum albumin (BSA) as a standard. For two-dimensional (2D) polyacrylamide gel electrophoresis (PAGE), the proteins were separated on short instruction immobiline dry strips, pH 3–10, by isoelectrophoresis focusing with Multiphor II Electrophoresis Unit (Amersham Biosciences Co., USA) according to the manual of Multiphor II 2-D and then the strips were subjected to sodium dodecyl sulfate (SDS)-PAGE according to the method of Laemmli (*23*). The proteins on the gels were electrophoretically transferred onto nitrocellulose membranes, then the proteins on the membranes were visualized with 0.1% Ponceau S in 10 mM acetate, followed by washing with 10 mMNaOH

for destaining of the dye. The proteins on the membranes were immunoblotted with the mAb, 1G4, against Gm23K or C5 against Gm28K, which had been described in a previous paper (*6*).

3.1.7. Detection of the glycan moiety of Gm23K

Gm23Ks separated on a nitrocellulose membrane were investigated with regard to its glycan moiety. The proteins on the membrane were stained with G.P. SENSOR (Honen Co., Japan) according to the manual for the use of G.P. SENSOR. Briefly, the proteins were blocked and treated with 10 mM $NaIO_4$ under darkness. The membrane was washed with PBS, and incubated with biotin-labeled hydrazide diluted with PBS. The membrane was incubated with HRP-labeled streptavidin. After washing, the avidin–biotin complexes on the membrane were visualized with 0.75% 4-chloro-1-naphthol and 0.3% H_2O_2 in 50 mM Tris–HCl buffer (pH 7.3).

For staining with a lectin, concanavalin A (Con A), the proteins on the membrane were blocked with 10 mM Tris–HCl buffer containing 0.05% Tween 20, pH 7.4 (TBS/T) for 10 min at room temperature three times. The membrane was incubated with biotin-labeled Con A (Honen), diluted 1000-fold with TBS/T for 1 h at room temperature. The membrane was washed and then incubated with HRP-labeled streptavidin, diluted 5000-fold with TBS/T, for 1 h at room temperature. Detection of the immunocomplexes on the membrane was performed in the same manner as described above.

3.1.8. Preparation of the probes for Northern blot

For Northern blot, DNA fragments corresponding to Gm28K and Gm23K regions were prepared from the phagemide pBluescript II SK (−), in which the region of the precursor of Gm28K was inserted. The DNA encoding the precursor was prepared according to the plasmid boiling miniprep protocol for a Rapid Excision kit (Stratagene, USA), and the DNA obtained was digested with restriction enzymes, *Nco*I and *Hin*dIII to obtain DNA fragments for Gm28K.

On the other hand, *Bam*HI and *Hin*dIII were used for cutting out the DNA fragment encoding Gm23K. The DNA fragments formed were separated by electrophoresis on a 3% low-melting agarose gel and the gel pieces possessing the fragments were cut off. The pieces were melted at 65 °C, and the DNA fragments in the pieces were extracted with phenol/chloroform (1:1, v/v) and were purified by ethanol precipitation. The DNA fragments were labeled with alkaline phosphatase using an AlkPhos Direct Labeling and Detection kit (Amersham Biosciences) and the labeled DNAs were used for Northern blot analyses in order to detect mRNA(s) for Gm28K and Gm23K as probe 1 and probe 2, respectively.

3.1.9. Northern blot of the Gm28K precursor in developing soybean seeds

Total RNAs were prepared from developing soybean seeds harvested at several stages using RNeasy Plant Mini kit (Qiagen, Japan). Poly $(A)^+$ RNAs were isolated from total RNAs using mRNA Isolation kit (Roche Diagnostics Co., USA), and denatured by formaldehyde and formamide. The denatured RNAs (3 μg/lane) were subjected to electrophoresis on a 1.5% agarose gel and the RNAs on the gel were transferred onto the magnacharge nylon membranes (Micron Separations Inc., USA) in the same manner as described in a previous paper (*62*). The RNAs on the membranes were fixed with an UV linker (Bio-Rad). The mRNA(s) for the two proteins were visualized using an AlkPhos Direct Labeling and Detection System with CDP-Star (Amersham Biosciences). One of the membranes was hybridized with the alkaline phosphatase-labeled probe for Gm28K and the other membrane was hybridized with that for Gm23K in hybridization buffer at 55 °C overnight. After washing the membranes with the washing solution, the specific mRNA for Gm28K and Gm23K were detected using an AlkPhos Detection kit (Amersham Biosciences) onto hyper films ECL (Amersham Biosciences).

3.1.10. Immunoblot of Gm28K and Gm23K in developing soybean seeds

The proteins from developing seeds harvested at several stages were extracted as described above. The proteins were separated on 15% gels by SDS-PAGE. The proteins were transferred onto nitrocellulose membranes, and were immunostained with mAb C5 or 1G4 in the same manner as described above.

3.1.11. Immunoelectron microscopy of Gm28K and Gm23K in developing soybean seeds

Analysis of immunocytochemical localization was carried out essentially as described by Takeuchi and Nishimura (*63*). Developing soybean seeds harvested 35 days after flowering were cut into 1.5–2.0 mm sections and fixed overnight at 4 °C in 3% (v/v) glutaraldehyde solution. Tissues were washed three times (10 min for each wash) with PBS and then dehydrated through a 30%, 50%, 60%, 70%, 80%, 90% (v/v) series of ethanol washes (15 min for each wash). The specimens were then further dehydrated by three washing steps in 95% ethanol (15 min for each wash), followed by three washing steps in 100% ethanol (30 min per wash). The dehydrated tissues were infiltrated for 2 h in a 1:1 (v/v) London white resin/ethanol mixture, in 2:1 (v/v) overnight, followed by 100% London white resin over the next 24 h. Polymerization was done in gelatin capsules for 2–7 days at 52 °C under nitrogen gas.

Ultrathin sections (60–80 nm) were obtained with a glass knife and placed onto formal/carbon-coated grids. The sections were blocked with 5% (w/v) BSA in PBS/T and then incubated for 1 h at room temperature on a drop of the mAb against Gm28K or the recombinant Gm23K, diluted to 0.74 or 0.28 μg/μl in PBS/T supplemented with 0.8% (w/v) BSA, 0.1% (w/v) gelatin, and 2 mM NaN_3 (BSA-PBS). The sections were washed four times for 5 min each on a drop of BSA-PBS and then incubated for 1 h at room temperature on a

drop of goat anti-mouse IgG (H+L) (Auro Probe EM, Amersham Biosciences) diluted 1:25 in PBS/T including 1% BSA. After washing twice with BSA-PBS, the sections were treated with 2.5% (w/v) glutaraldehyde solution for 10 min and washed twice with PBS/T and then once with distilled water (5 min per step). The sections were stained for 30 min with 4% (w/v) uranyl acetate followed by a 1 min incubation in 80 mM lead nitrate, 120 mM trisodium citrate dihydrate, and 0.16 N NaOH for 1 min. The grids were examined and photographed using an electron microscope (model H-700H, Hitachi, Tokyo).

3.1.12. Immunoblot of Gm23K with the sera of soybean-sensitive patients and with anti-HRP antibody

It was very difficult to purify Gm23Ks by conventional methods such as ion chromatography and high-performance liquid chromatography, because they occur in a very low concentration as described in the Discussion section. Therefore, soybean proteins were separated by 2D-PAGE. After electrophoresis and staining with 0.1% Coomassie brilliant blue R250 in methanol/acetic acid/water (3:1:6, by volume), the gel pieces possessing the 23-kDa peptides were cut off and the peptides were electrophoretically eluted from the gel pieces in the same manner as described in a previous paper (*62*). The above-mentioned natural Gm23Ks, Gm28K, or the recombinant Gm23Kwere subjected to SDS-PAGE and transferred onto nitrocellulose membranes. The membranes were blocked with blocking reagents in an ECL Advance Detection kit (Amersham Biosciences) for 1 h, and then, the peptides on the membranes were incubated with the sera of 10 soybean-sensitive patients, or the serum of a patient, who was suffering not from soybean allergy but from mite-allergy, or the serum of a non-allergy person, at 4 °C overnight. The sera used were diluted from 1:20 to 1:30 with the blocking reagents. The membranes were washed with TBS/T, then the membranes were incubated with HRP-labeled anti-human IgE (ICN, USA), diluted to 1:100,000 with the blocking reagents, for 1 h at room temperature. After washing the membranes, the

immunocomplexes were detected using the detection reagents onto hyper films ECL. On the other hand, for examination on the reactivity of each of Gm23Ks, the peptides in 7S globulin fraction were separated by 2D-PAGE and transferred onto nitrocellulose membranes as described above. Thereafter, each of Gm23Ks on the membranes was immunoblotted with the sera of the patients sensitive to soybean.

Immunostaining and inhibition experiments with rabbit polyclonal anti-HRP (Sigma Chemical Co., USA) recognizing N-linked glycan moieties of HRP with the β1,2-xylose and/or α1,3-fucose branches were performed using the proteins in 7S globulin fraction, which were separated by 2D-PAGE, followed by the above-mentioned immunoblotting analyses, as described in a previous paper (*7*).

3.2 RESULTS

3.2.1 Identification of Gm23K derived from the precursor of Gm28K in soybean

It is very interesting to examine whether the postulated Gm23K would be derived from the precursor of Gm28K and whether Gm23K would occur in soybean cotyledons. The mAbs, 1G4, 2E6 and 2D7, against the recombinant Gm23K were prepared as a probe for the detection of Gm23K in growing soybean cotyledons. The class and subclass of the mAb obtained were examined with a mouse monoclonal isotyping kit (Roche Diagnostics). These light chains were shown to be κ, and the heavy chains of 1G4 and 2E6 to be $\gamma_{1,}$ 2D7 to be γ_{2a} (Table 4). The mAb 1G4 recognized several peptides, but these peptides could not be immunoblotted with mAb C5 specific for Gm28K (Fig. 9B and C). These results clearly demonstrate that the peptides recognized belong to a group of Gm23K and that these peptides are different from Gm28K. In the present study, these peptides corresponding to Spots 1–4 in Fig. 9A are named Peptides 1–4. The spot recognized by the mAb on the right side of Spot 3 was observed as a clear spot by immunoblotting, but the peptide corresponding to the spot could not be detected by Ponceau S staining (Fig. 9A). This fact indicates that the concentration of the peptide may be too low to detect it as a peptide. Spot 2 was shown to be the most intense band by immunoblotting with the mAb. Spot 2 consisted of two peptides, which were different in molecular mass from each other. The peptide with the higher molecular mass and the other peptide are named Peptide 2–1 and Peptide 2–2, respectively.

The N-terminal amino acid sequences of the peptides recognized by the mAb against the recombinant Gm23K were determined. The N-terminal amino acid sequences of all the peptides were shown to be KDTAGSPASYNLYDDKKADF. The result clearly shows that the N-terminal amino acid sequences of the mAb-positive peptides correspond to the sequence

Lys270–Phe289 on the deduced amino acid sequence of the precursor of Gm28K, and that Gm23Koriginates from the C-terminal half part of the Gm28K precursor.

Table 4 Class and subclass of three mAbs against the recombinant Gm23K

mAb	light chains	heavy chains
1G4	κ	γ_1
2E6	κ	γ_1
2D7	κ	γ_{2a}

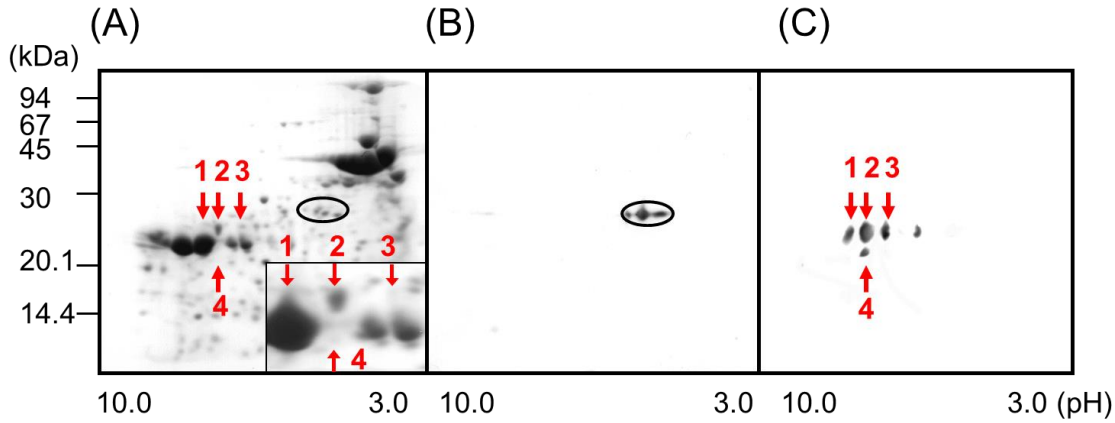

Fig. 9. Immunoblots of the proteins in the 7S globulin fraction with the mAb against the recombinant Gm23K. The 7S globulin fraction was isoelectrically focused on Immobilin Dry Strips and then the proteins on the gels were electrophoresed onto 12% polyacrylamide gels. After 2D-PAGE, the proteins were electrotransferred to nitrocellulose membranes and stained with mAb 1G4 for the recombinant Gm23K or mAb C5 for Gm28K. The pattern of proteins in the 7S globulin fraction stained with Ponceau S was shown in panel A. Immunoblotting patterns with mAb C5 for Gm28K and with mAb 1G4 for the recombinant 23-kDa peptide were shown in panels B and C, respectively. The circles in the panels express Gm28K. The region corresponding toGm23Ks was magnified in the insert in panel A. Arrows 1, 2, 3, and 4 in panel C represent the spots corresponding to Gm23K. Peptides 1–4 marked with arrows in panel A were subjected to N-terminal amino acid sequencing as described in Materials and methods. The standard protein markers are shown on the left side of panel A.

3.2.2 The emergence of the Gm28K precursor and localization analysis of Gm28K and Gm23K in soybean

We attempted to elucidate the emergence of Gm28K and Gm23K during the development of soybean cotyledons. Probe 1 for Gm28K and Probe 2 for Gm23K were prepared to detect the mRNA(s) for the two proteins. Probe 1 consisted of 273 bp corresponding to sequence 178–450 and Probe 2 had a nucleotide of 431 bp corresponding to sequence 958–1388. The mRNA fraction prepared from cotyledons harvested at various intervals were subjected to Northern blotting analysis with the alkaline phosphatase-labeled probes. As shown in Fig. 10B and C, the mRNAs for Gm28K and Gm23K were observed to be expressed in the cotyledons harvested at the 21st day after flowering. Only a single band corresponding to mRNA with a length of approximately 1620 bases was observed by Northern blotting using the probes for the two proteins. The amount of the mRNA gradually increased from the 21st day to the 49th day after flowering.

Gm28K and Gm23K in developing soybean cotyledons were investigated by immunoblotting with the mAbs against the two proteins. Both of Gm28K and Gm23K were observed to come out in cotyledons harvested at the 21st days after flowering (Fig.10E and F). The expression profiles of the two proteins were compatible with those of the mRNA for Gm28K and Gm23K. However, the proprotein for Gm28K and Gm23K, which seems to be formed from the preproprotein by removal of the signal peptide, could not be detected.

Electron microscope-immunogold analysis with the specific monoclonal antibodies for Gm28K and Gm23K show that the signals of gold particles were detected on crystalloids in the PSV (Fig. 11B, C).

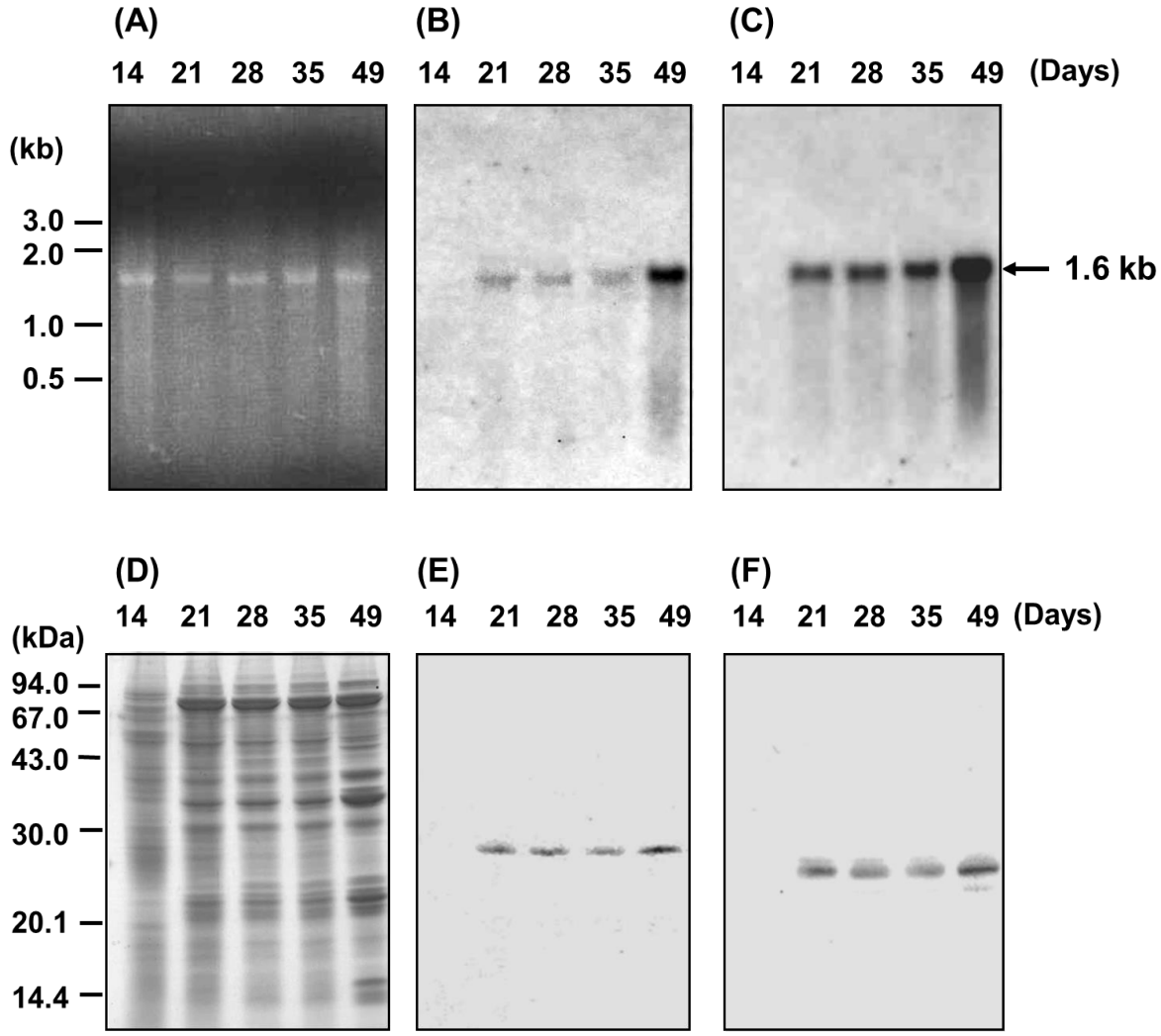

Fig. 10. Northern blot of the mRNA for the Gm28K precursor and immunoblot analysis of Gm28K and Gm23K in growing soybean. Poly $(A)^+$ RNAs from each growing soybean cotyledon harvested from the 14th to the 49th day after flowering were subjected to electrophoresis on 1.4% agarose gels were stained with ethidium bromide (A). The poly $(A)^+$ RNAs on the gels part were transferred onto nylon membranes. Each of the membranes were subjected to Northern blot analysis with alkaline phosphatase-labeled DNA fragment encoding Gm28K (B) or Gm23K (C) as the probes. RNA size markers are indicated on the left of panel A in kb. Proteins were stained with Ponceau S (D) or electrotransferred to nitrocellulose membranes and immunoblotted with mAb C5 for Gm28K (E) and mAb 1G4 for Gm23K (F). Protein size markers are indicated on the left of panel D in kDa.

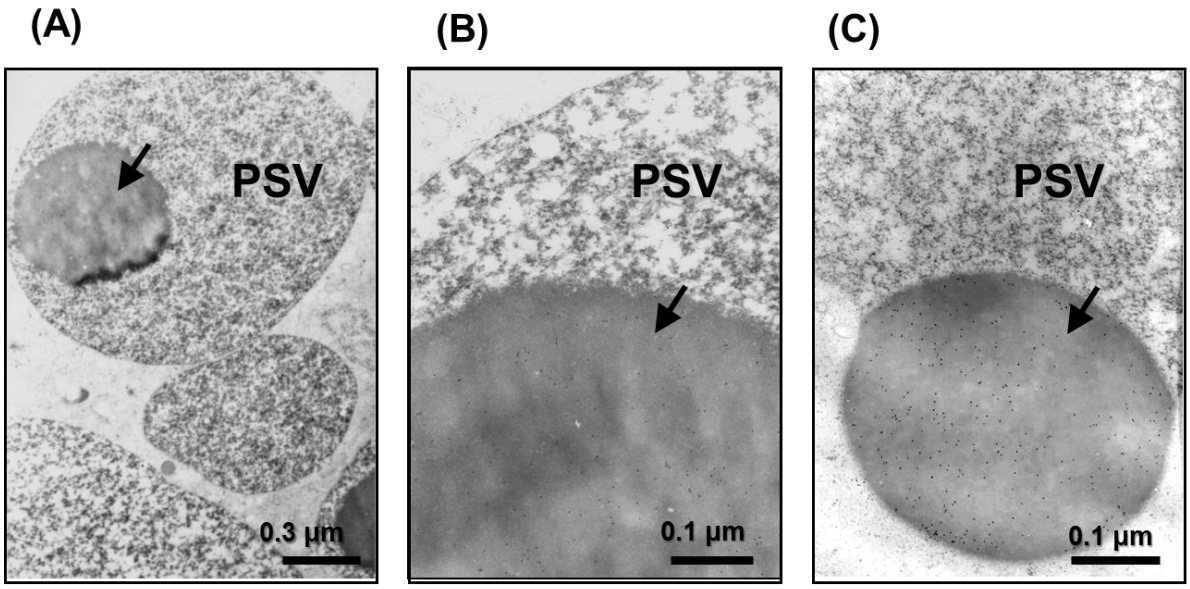

Fig. 11. Immunogold-labeling analysis of Gm28K and Gm23K in a developing soybean cotyledon. The analysis was done in the same manner as described in Materials and methods. The arrows represent the crystalloids of PSV of a developing soybean seed. (A and B) Immunogold-labeling analysis with mAb C5 for Gm28K; (C) immunogold-labeling analysis with mAb 1G4 for Gm23K; panel A, 10,500-fold magnification; panels B and C, 30,000-fold magnification.

3.2.3 Characterization of Gm23K

As described above, Gm23Ks were detected as several spots by immunoblotting with the mAb against the recombinant Gm23K. Interestingly, the peptides were shown to have the same N-terminal amino acid sequence, suggesting that the multiple forms of Gm23K might be formed by the modification of the original Gm23K, such as glycosylation and/or deamidation and/or carbamylation as described in the Discussion.

Gm28K is a glycoprotein with an N-linked glycan moiety. Therefore, we investigated whether Gm23Ks might be glycoproteins. When the Gm23Ks, which were electroblotted on the nitrocellulose membrane after electrophoresis, were analyzed with G.P. SENSOR, Gm23Ks except for Spot 4 were detected with G.P. SENSOR. The results suggest that Gm23Ks for Spots 1–3 may be glycoproteins (Fig. 12B). Staining with Con A specific to a

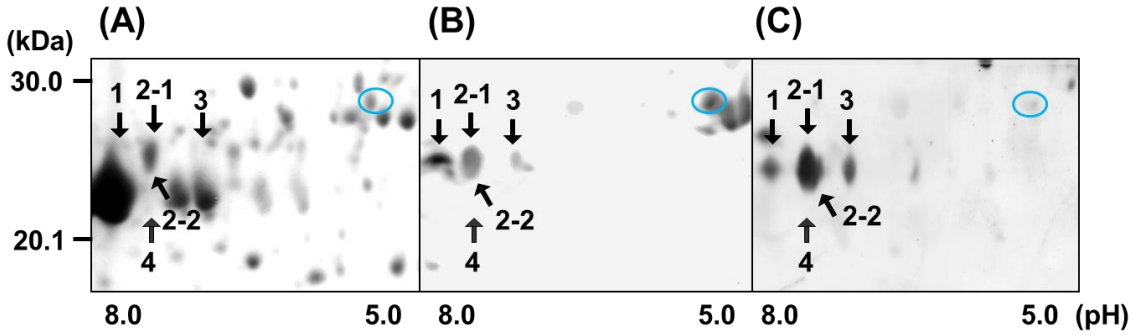

Fig. 12. Detection of the glycan moiety of Gm23K in the 7S globulin fraction. 2D-PAGE and electroblotting to nitrocellulose membranes were conducted in the same manner as described in the legend to Fig. 9. (A) Stained with Ponceau S; (B) stained with G.P. SENSOR; (C) stained with Con A. The arrows in panel A represent the spots corresponding to Gm23Ks. In panels B and C, the closed arrows show the spots positive to G.P. SENSOR and Con A, and the open arrows show the position corresponding to the spots negative to G.P. SENSOR and Con A. The circles in the panels indicate the position for Gm28K. The standard protein markers are shown on the left side of panel A.

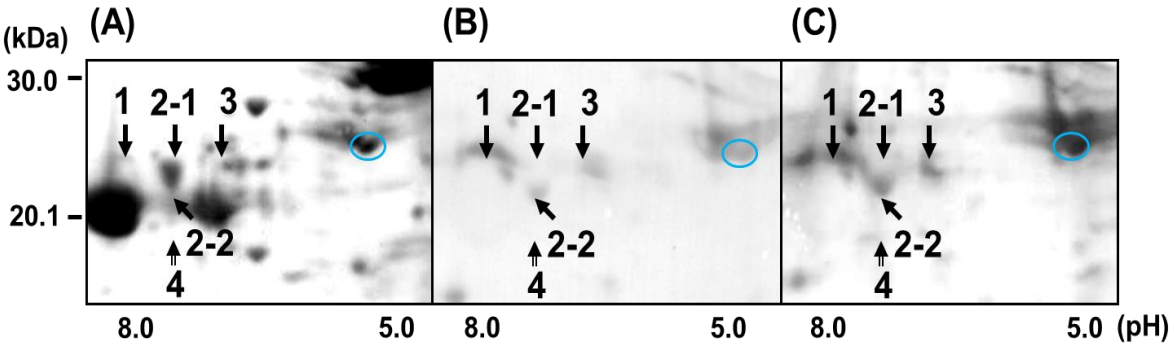

Fig. 13. Immunoblots of Gm23K in the 7S globulin fraction separated by 2D-PAGE. 2D-PAGE and electroblotting to nitrocellulose membranes were conducted in the same manner as described in the legend to Fig. 9. (A) Pattern obtained by staining with Ponceau S; (B) pattern obtained by staining with a serum of a soybean-allergic patient described in Fig. 12; (C) pattern obtained by staining with anti-HRP polyclonal antibody. The arrows represent the spots corresponding to Gm23Ks. In panels B and C, the closed arrows show the spots immunoblotted and the open arrows show the positions corresponding to the peptides recognized with the mAb, but not immunoblotted with the sera of soybean-sensitive patients. The circles in the panels represent the position corresponding to Gm28K. Standard protein markers are shown on the left side of panel A.

mannose residue which inevitably occurs in glycoprotein containing N-linked glycan moiety (*65, 66*) confirmed that these peptides may be N-linked glycoproteins (Fig.12C). Also, some of Gm23Ks were detected by immunoblotting with anti-HRP (Fig. 13C), which recognizes the β1,2-xylose and/or α1,3-fucose branches in N-linked glycan moieties (*57, 58, 67-69*). Although the staining patterns of Gm23Ks with G.P. SENSOR were compatible with those with Con A, both the patterns were different from that obtained by immunoblotting with anti-HRP. Among Gm23Ks immunoblotted with the mAb, Peptide 2-1 and Peptide 4 could not be detected with anti-HRP (Fig. 13C).

3.2.4. Binding of Gm23K to IgE antibodies in the sera of patients sensitive to soybean

The allergenicity of Gm23K was investigated by immunoblotting with the sera of 10 patients sensitive to soybean. All of the patients' sera were observed to recognize Gm28K. The natural Gm23Ks prepared from soybean were recognized with IgE antibodies in most of the sera excepting Patients 4 and 9 (Fig. 14B), although their reactivities were weaker than those of Gm28K. In contrast, the recombinant Gm23K reacted scarcely with the IgE antibodies in the above-mentioned sera (Fig. 14C). In order to examine each reactivity of Gm23Ks components with the above sera, Gm23K components in 7S globulin fraction, which were prepared by Thanh and Shibasaki (*70*), were separated by 2D-PAGE. And then, Gm23Ks were electroblotted onto a nitrocellulose membrane, after which the membrane was subject to immunoblotting with the serum of Patient 1 used in Fig. 14. The immunoblotting pattern is shown in Fig. 13. The peptides corresponding to Spot 2–1, a major spot of Gm23K, and that of Peptide 4 did not react with IgE antibodies in the patient's serum. Immunoblotting with the sera of the other patients positive for the natural Gm23K were observed to exhibit patterns similar to that obtained with the serum of Patient 1 (data not shown). Interestingly, all of Gm23Ks immunoblotted with the sera (Spots 1, 2–2, 3) were recognized with anti-HRP polyclonal antibodies as shown in Fig. 13C. The binding of Gm23K to the IgE antibodies in

the sera of the above-mentioned patients was shown to be inhibited strongly by preincubation with anti-HRP antibody (data not shown).

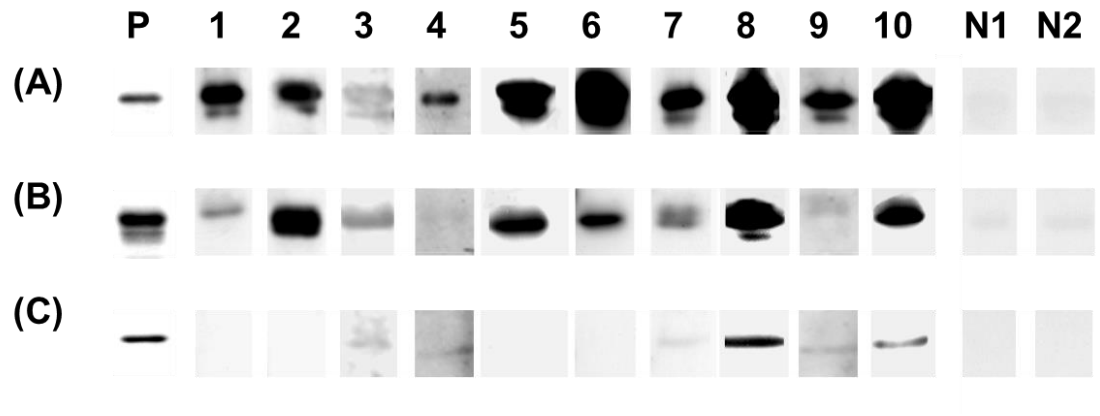

Fig. 14. Immunoblots of Gm23K with the sera of soybean-sensitive patients. Gm28K, the natural Gm23Ks and the recombinant Gm23K were electrophoresed and electroblotted to nitrocellulose membranes. The membranes were cut off into several pieces and immunoblotted with the sera of patients sensitive to soybean. (A) Gm28K; (B) the natural Gm23Ks; (C) the recombinant Gm23K. P shows staining of the proteins on the membranes with Ponceau S. Panels 1–10, the immunoblots obtained with the sera of 10 patients sensitive to soybean; panel 11, that obtained with the serum of a mite-allergy patient who did not have the sensitivity to soybean; panel 12, the immunoblot obtained with the serum of a non-allergic person.

3.3 DISCUSSION

In Chapter 2, it is suggested that Gm28K is translated as a preproprotein consisting of a signal peptide, Gm28K, and a putative C-terminal peptide (Gm23K). After the signal peptide is cleaved in the ER, the proprotein is transported to the vacuole through the Golgi apparatus.

In this chapter, in order to investigate whether Gm23K actually exists in soybean seeds, we prepared a mAb against the peptide derived from the recombinant Gm23K expressed in *Escherichia coli*, which would be a useful probe for detection of the peptide. When proteins in soybean were separated by 2D-PAGE, followed by immunoblotting with the mAb, several peptides with different isoelectric points and molecular masses, which specially reacted with the mAb, were observed (Fig. 9). These peptides positive to the mAb were shown to have quite identical N-terminal amino acid sequences corresponding to sequence Lys294–Phe314 on the deduced amino acid sequence of the precursor for Gm28K. These results indicate that Gm23K actually exists in soybean seeds as the C-terminal counterpart of the precursor of Gm28K, suggesting that the precursor would be converted to Gm28K and Gm23K by cleavage at the carboxyl side of the Asn294 residue with an asparaginyl endopeptidase. The processing enzyme involved in the metabolism of the precursor might be the same asparaginyl endopeptidase as that participating in the process of plant proteins, for example, MP27/MP32 (*71*) and 11S globulins (*72*, *73*). The occurrence of an asparaginyl endopeptidase in PSV of soybean seeds has ever been demonstrated (*73*, *74*). The above-mentioned observations propose that the processing enzyme would be connected with the metabolism of the proprotein consisting of Gm28K and Gm23K after removal of the signal peptide.

Gm23K was observed to be detected as several spots positive to the mAb against the recombinant Gm23K. This phenomenon indicates that Gm23K may essentially occur as multiple forms in soybean or that its multiplicity may be due to modification arising during electrophoresis. When 7S globulin fraction containing Gm28K was subjected to 2D-PAGE,

followed by immunoblotting with the mAb specific for the allergen, Gm28K had shown a phenomenon similar to the multiplicity of Gm23K as described in a previous paper (*6*). This phenomenon suggests that the multiple forms of Gm23K with different isoelectric points might be formed due to deamidation of glutamine and/or asparagine residues or carbamylation of lysine and/or arginine residues with cyanate or isocyanate derived from urea during isoelectric focusing (*75, 76*). However, the difference between in the molecular masses of Peptides 1, 2, and 3 and that of Peptide 4 could not be explained only by the above-mentioned deamidation and carbamylation. Glycosylation of the original Gm23Kwould result in the formation of molecular forms with larger molecular masses than that of original peptide. We tried to examine whether Gm23Ks have any glycan moieties using G.P. SENSOR. As shown in Fig. 12, Peptides 1, 2–1, 2–2, and 3 among the positive peptides were stained with the G.P. SENSOR and Con A, suggesting that these peptides may be glycoproteins with N-linked glycan moieties. Since the deduced amino acid sequence of the precursor for Gm28K includes N-X-S/T, a potential binding site of an N-linked glycan moiety, at the Asn349 residue, the residue would be a candidate as an N-linked glycan-binding site of Gm23K. However, among the mAb-positive peptides, Peptide 4 with the lowest molecular mass was shown to have no glycan moiety. The above findings clearly support the idea that the glycosylation may also contribute to the multiplicity of the original Gm23K.

It is interesting to elucidate the expression of the Gm28K precursor during the development of soybean cotyledons to obtain the information as to the synthesis of the two proteins. Therefore, we cultivated soybean seeds in the garden of our university and harvested developing cotyledons at several intervals.

The mRNA(s) for Gm28K and Gm23Kwere examined using by Northern blotting with the DNA fragments corresponding to Gm28K and Gm23K and the two proteins were immunoblotting with the mAbs against the two proteins. The mRNA for Gm28K and Gm23K

was detected as a single band corresponding to the Gm28K precursor with a polynucleotide of approximately 1600 bases, suggesting that only one mRNA comes out at the 21st day after flowering and that the mRNA detected would encode the full-length of the Gm28K precursor (*8*). Concomitant with the expression of the mRNA for Gm28K and Gm23K, these proteins were shown to be expressed in the cotyledons of soybean harvested at the 21st day after flowering. However, the protein corresponding to the Gm28Kprecursor was not detected. Based on the experiments using the tobacco BY2 cells in chapter 2, a proprotein form the Gm28K precursor would be suggested to be transported from the ER to the Golgi apparatus. However the proprotein form exists in the ER, but the proprotein could not be detected, because it may be in a trace amount.

The localization of Gm28K and Gm23K in developing cotyledons of soybean was also examined. Both of the proteins were observed to colocate in the crystalloid parts of the PSV of the cotyledons. As cotyledons grew up, the amount of the two proteins were shown to increase, and, dry seeds of soybean would contain a large amount of the proteins.

The fact that Gm28K and Gm23K are transported to the vacuole in soybean seeds is consistent with the localization of tobacco BY2 cells, in which the Gm28K precursor form were expressed, as described in chapter 2. The Gm28K precursor gene was shown to be transcribed and translated during the long period of time from the early stage to one of late maturing seeds. The proprotein, which is, derived from the preproprotein of Gm28K and contains Gm28K and Gm23K segments, is likely to be glycosylated to form a high-mannose-type of glycoprotein in the ER. The glycosylated proprotein would be transferred to the Golgi apparatus and elaborated under a strict regulation, by removal of glucose and/or mannose and/or addition of xylose and/or fucose. Thereafter, the proprotein would move to vacuoles, in which the proproteins would be converted to Gm28K and Gm23K by cleavage with the above-mentioned asparaginyl endopeptidase.

Gm28K is a glycoprotein with N-linked sugar chain, and it has been demonstrated that its sugar chains are involved in binding to IgE antibodies in the serum of soybean-sensitive patients. Therefore, the allergenicity of Gm23Ks was examined by immunoblotting with the sera of 10 patients sensitive to soybean. Since Gm23K exists in a very low concentration, the peptide could not be purified directly from soybean flake. Therefore, Gm23Ks were prepared by elution from the gel pieces of 2D-PAGE of the 7S globulin fraction. The native Gm23K from soybean was shown to be clearly recognized with IgE antibodies in the sera of most patients (Fig, 14). These results demonstrate that Gm23K have the binding activity to IgE antibodies in the sera of the patients, showing that Gm23K may be a new allergen in soybean.

In a previous paper (*2*), 16 IgE-binding proteins had been reported to exist in soybean. An IgE-binding protein corresponding to the above-mentioned Gm23K had been observed in the extract obtained from soybean flakes, but Gm23K could not have been identified as an IgE-binding protein in the previous paper (*2*). Gm23K is generally contained in the 7S globulin fraction prepared by isoelectric precipitation according to Thanh and Shibasaki (*68*). However, the 7S globulin fraction used in the previous work (*2*) had been obtained by further purification of the original 7S globulin fraction with ultracentrifugation. Therefore, Gm23Kwould be removed during the ultracentrifugation and the 7S globulin fraction would have no Gm23K, explaining why Gm23K as an IgE-binding protein had been missed in the previous paper (*2*).

The involvement of N-linked glycans in plant allergens in the binding to IgE antibodies has not been well characterized. In the present study, Gm23K was examined with regard to the relationship between the glycan moiety and the allergenicity by immunoblotting with the sera of 10 patients sensitive to soybean. As shown in Fig. 14, the sera of 8 of 10 patients recognized clearly the natural Gm23K. However, the sera stained weakly or scarcely the recombinant peptide without any glycan moiety. Similar results had been observed with regard

to the IgE-binding of Gm28K (*8*). It occurred to us that the glycan moiety may be connected with the binding to IgE antibodies in the sera of the patients sensitive to soybean. Further investigation on the glycan moiety of each Gm23K in soybean seeds has been done using 2D-PAGE. Among Gm23Ks positive to mAb, Peptides 1, 2–2 and 3 were shown to be recognized with the sera of the soybean-sensitive patients (Fig. 13B). Interestingly, these three peptides were also recognized with anti-HRP antibodies, which are specific for the glycan moiety of HRP. Furthermore, when Gm23Ks were incubated with anti-HRP and then reacted with IgE antibodies in the patients' sera, the IgE-binding of the peptides, which were positive with the above-mentioned sera, was strongly inhibited (data not shown). These facts suggest that the glycan moiety of Gm23Ks positive to anti-HRP may bind to IgE antibodies in the patients' sera. Gm28K is recognized with anti-HRP antibodies in the same manner as Gm23K is. The above findings mean that the N-linked glycan moiety of Gm23Ks with α1,3-fucose residue and/or β1,2-xylose residue may be involved in its allergenicity. Because, with regard to a number of allergens with N-linked glycan moieties in plants, such as wheat, barley and olive pollen (12, 13, 78), their glycan moieties containing α1,3-fucose residue and/or β1,2-xylose residue have been reported to function as a common epitope. Although there are several studies on the involvement of the above-mentioned glycan moieties in the allergenicity of the allergens, the exact chemical structures of the glycan moieties responsible for binding to IgE antibodies have never been elucidated. The relationship between the IgE-binding activity and the chemical structure of the glycan moieties in allergens with N-linked glycan moieties still remains unknown.

Chapter 4

Conclusion Remarks

The major allergens in soybean are Gly m Bd 68K, Gly m Bd 30K and Gm28K. The former two allergens have been well studied. Since Gm28K was an unknown protein, it was purified and its properties have been well elucidated. Furthermore, the cloning of a cDNA encoding Gm28K has been done and the analysis of the cDNA suggested that the precursor for Gm28K might be synthesized to release Gm28K during its metabolism. However, the cloned cDNA did not encode the complete amino acid sequence of the precursor, although the precursor was suggested to consist of SP, Gm28K and the 23-kDa peptide (Gm23K). In the present study, I focused on the elucidation of the precursor of Gm28K.

In Chapter 2, in order to reveal the nucleotide sequence of the full-length cDNA encoding the Gm28K precursor, the cloning of the cDNA was done using PCR. Next, the amino acid sequence of the Gm28K precursor with regard to its transport in the cells has been examined using tobacco BY2 cells. We succeeded in cloning of the full-length cDNA encoding the precursor and elucidated the complete amino acid sequence of the SP part with 24 amino acid residues including newly determined three additional amino acid residues. On the basis of the newly cloned cDNA, various expression vectors were constructed and transformed into tobacco BY2 cells. For the detection of the proteins expressed in tobacco BY2 cells, the gene for EGFP as an indicator protein, which exhibits fluorescence, was incorporated. Localization of the proteins expressed in tobacco BY2 cells were examined with a fluorescence microscopy and a laser-scanned confocal microscopy.

When the expression vector SP-EGFP was expressed in the tobacco BY2 cells, EGFP located in the ER of the cells and was secreted in their medium. These findings mean that the SP part acts as a signal peptide, which leads its precursor to the ER. When the cells with expression vectors for SP-EGFP-Gm28K-Gm23K and SP-Gm28K-Gm23K-EGFP were cultured, EGFP was observed in the vacuoles of SP-EGFP-Gm28K-Gm23K/BY2cells and SP-Gm28K-Gm23K-EGFP/BY2 cells. On the other hand, in the SP-EGFP-Gm28K/BY2 cells, EGFP was secreted in the medium similar to SP-EGFP/BY2 cells. Furthermore, in the

SP-EGFP-Gm23K/BY2 cells, EGFP was observed in the ER but not in the medium. The above observations demonstrate that the Gm28K-Gm23K part is essential for the transport of the Gm28K precursor to vacuoles in the tobacco BY2 cells. In conclusion of the three segments of the precursor, SP, Gm28K and Gm23K, SP may act as a signal peptide for the transport of Gm28K and Gm23k across the ER membrane, and the Gm28K and Gm23K segments can transport the proprotein to vacuoles in the cells co-operatively. Finally, the Gm28K and Gm23K segments are cleaved between Gm28K and Gm23K in the vacuoles.

As describe above, the Gm28K precursor might consist of the three segments, SP, Gm28K and Gm23K. However, the occurrence of the latter segment, Gm23k, in soybean has never been examined. In Chapter 3, the occurrence of Gm23K in soybean, and its properties including allergenicity have been investigated. Furthermore, the fate of the Gm28K precursor during the development of soybean seeds has been examined.

We produced a mAb against the recombinant Gm23K in *E. coli* to detect Gm23K in soybean seeds specifically. By immunoblotting with the above mAb, the occurrence of Gm23K in soybean seeds was demonstrated. Interestingly, several spots positive to the mAb were observed. The multiplicity of Gm23K would depend on the modification of Gm23K undertaken during 2D-PAGE such as deamination and carbamylation, or glycosylation.

During development of soybean seeds, the metabolism of the Gm28K precursor was investigated by Northern blotting and immunoblotting with the mAbs against native Gm28K and a recombinant Gm23K. As expected, only one band corresponding to the mRNA for the precursor with a nucleotide sequence of approximate 1,600 bases was observed from 14 days after flowering, when the mRNA samples obtained during development of soybean seeds were subjected to Northern blotting. Furthermore, when the proteins in the developing soybean seeds were subjected to SDS-PAGE and then immunoblotting with the mAbs against Gm28K and Gm23K, only the bands corresponding to Gm28K and Gm23K were detected, but the preproprotein (the Gm28K precursor) could not be detected. These findings indicate that

Gm28K and Gm23K may be synthesized as a preproprotein and that Gm28K and Gm23K are produced from the preproprotein by hydrolysis. Also Gm28K and Gm23K were shown to locate in the crystalloids of the vacuoles in soybean seeds. This fact is compatible with the findings obtained from the experiments in which the metabolism of the precursor in tobacco BY2 cells was examined as described in Chapter 2, and demonstrated that the proprotein (the Gm28K-Gm23K segment) is hydrolyzed to Gm28K and Gm23K by an asparaginyl endopeptidase.

As described above, Gm23K was observed to occur in the multiple forms by 2D-PAGE. Some of multiple formed-Gm23K with N-glycans also exhibited the binding to IgE antibodies in the sera of patients sensitive to soybean, which means that Gm23K with sugar chains has allergenicity. It is very interesting to confirm the allergenicity of Gm23K. Therefore, in order to purify Gm23K from soybean and clearly demonstrate its allergenicity, soybean proteins were subjected to SDS-PAGE, Gm23K in a gel stained with Coomassie Blue R250 was cut out and was eluted by electrophoresis. The purified Gm23K were shown to be detected by immunoblotting with IgE antibodies in most of the sera of soybean-sensitive patients, but the recombinant Gm23K expressed in *E. coli* was not immunoblotted or weakly immunoblotted with most of the sera. The above observations demonstrate that the sugar moiety of Gm23K contributes to the binding to IgE antibodies in the sera of patients sensitized with soybean. The similar phenomena has been observed in Ole e 1, Gm28K and wheat allergen, Tri a Bd 27K.

Gm23K was shown to be a glycoprotein with N-linked sugar chain by staining with G.P. SENSOR and lectin blotting. The binding of the sugar chain of the allergen and IgE antibodies in the patients' sera may induce immune response and cause allergy in humans (*12*, *78*). Furthermore, β1,2-xylose and α1,3-fucose residues are considered as a IgE binding carbohydrate determinants, so-called cross-reactive carbohydrate determinants (CCD) of plant allergens (*78*). IgE antibody against β1,2-xylose and α1,3-fucose residues is induced by exposure to grass and tree pollens and results in extensive cross-reactivity toward plant foods

(*79*). In recent studies, in the binding activity of CCD to IgE antibodies, α1,3-fucose residues have been demonstrated to react strongly than xylose residues (*79*). In addition, 3D structure of plant specific N-glycan, displaying core β1,2-xylose and α1,3-fucose and locating at opposite side (*80*), proposing that β1,2-xylose and α1,3-fucose function as independent epitopes (*78*). Thus, the specific plant glycan is likely to be an epitope against for IgE antibodies in the patients' sera as well as Gm28K. In Japan, RAST is generally a method used for examination of food allergy, and its results reported that a concentration of soybean-specific IgE antibody is high.

Acknowledgement

The completion of this doctoral dissertation is owing to the support of the following persons. The author would like to express the sincere gratitude to Dr. Masumi Kimoto, Professor of Okayama Prefectural University, for her kind instruction and encouragement during the course of this study. I would like to gratefully and sincerely thank my supervisor, Dr. Hideaki Tsuji for his support, encouragement and providing me with an excellent situation for doing research during my graduate studies. I would like to express acknowledgement to Professor Kazuhito Fujiyama at international Center for Biotechnology, Osaka University, a debt of gratitude to a lot of his great help during this research. I am also truly indebted and thankful to Associate Professor Ryo Misaki for his great helps to do the experiments at Osaka University. I would also like to express my gratitude to the two members of our graduate committee, Professor Koichiro Yamamoto and Hiromi Yamashita. I am heartily thankful to Dr. Makiko Suzuki, Dr. Miki Hiemori, Dr. Takao Ohashi, Dr. Hiroyuki Kajiura for their valuable support and advice.

I would like to acknowledge all of teachers in the Department of Nutritional Science, Faculty of Health and Welfare Science, Okayama Prefectural University and Iternational Center for Biotechnology, Osaka University for their positive helps and supports.

I would like to show my gratitude to all students who are/were present in Dr. Tsuji's laboratory and Professor Fujiyama's laboratory for their supports.

Finally, I would like to express my hearty thanks to my friends and family. My friends were always supporting me and encouraging me with their best wishes. And, importantly, I am deeply appreciative of my family's personal support and understanding at all times.

References

1) Urisu A, Ito H, Shiomi K, Ebisawa M, Adati R, Horiguti I. Study on QOL improvement and safety management of food allergy patients based on scientific knowledge. *Research report 2010 of the MHLW science research grants*. 201033018A (2011).

2) Ogawa T, Bando N, Tsuji H, Okajima H, Nishikawa K, Sasaoka K. Investigation of the IgE-binding proteins in soybeans by immunoblotting with the sera of the soybean-sensitive patients with atopic dermatitis. *J Nutr Sci Vitaminol.* 37 (1991) 555-565.

3) Ogawa T, Bando N, Tsuji H, Nishikawa K, Kitamura K. α-subunit of β-conglycinin, an allergenic protein recognized by IgE antibodies of soybean-sensitive patients with atopic dermatitis. *Biosci Biotech Biochem.* 59 (1995) 831-833.

4) Ogawa T, Tsuji H, Bando N, Kitamura K, Zhu YL, Hirano H, Nishikawa K. Identification of the soybean allergenic protein, Gly m Bd 30K, with the soybean seed 34-kDa oil-body-associated protein. *Biosci Biotechnol Biochem.* 57 (1993) 1030-1033.

5) Ji C, Boyd C, Slaymaker D, Okinaka Y, Takeuchi Y, Midland SL, Sims JJ, Herman E, Keen N. Characterization of a 34-kDa soybean binding protein for the syringolide elicitors. *Proc Nat Acad Sci USA* 95 (1998) 3306–3311.

6) Tsuji H, Bando N, Hiemori M, Yamanishi R, Kimoto M, Nishikawa K, Ogawa T. Purification of characterization of soybean allergen Gly m Bd 28K. *Biosci Biotechnol Biochem.* 61 (1997) 942-947.

7) Hiemori M, Bando N, Ogawa T, Shimada H, Tsuji H, Yamanishi R, Terao J. Occurrence of IgE antibody-recognizing *N*-linked glycan moiety of a soybean allergen, Gly m Bd 28K. *Int Arch Allergy Immuno.l* 122 (2000) 238-245.

8) Tsuji H, Hiemori M, Kimoto M, Yamashita H, Kobatake R, Adachi M, Fukuda T, Bando N, Okita M, Utsumi S. Cloning of cDNA encoding a soybean allergen, Gly m Bd 28K. *Biochim Biophys Acta.* 1518 (2001) 178-182.

9) Sanchez-Monge R, Gomez L, Barber D, Lopez-Otin C, Armentia A, Salcedo G.. Wheat and barley allergens associated with baker's asthma. Glycosylated subunits of the alpha-amylase-inhibitor family have enhanced IgE-binding capacity. *Biochem.J*. 281 (1992) 401–405.

10) Batanero E, Villalba M, Rodríguez R. Glycosylation site of the major allergen from olive tree pollen. Allergenic implications of the carbohydrate moiety. Mol. Immunol. 31 (1994) 31–37.

11) Garcia-Casado G, Sanchez-Monge R, Chrispeels MJ, Armentia A, Salcedo G, Gomez L. Role of complex asparagine-linked glycans in the allergenicity of plant glycoproteins. *Glycobiology* 6 (1996) 471–477.

12) Fötisch K, Altmann F, Haustein D, Vieths S. Involvement of carbohydrate epitopes in the IgE response of celery-allergic patients. *J. Allergy Clin. Immunol.* 120 (1999) 30–42.

13) Batanero E, Crespo JF, Monsalve RI, Martín-Esteban M, Villalba M, Rodríguez R.

IgE-binding and histamine-release capabilities of the main carbohydrate component isolated from tmajor allergen of olive tree pollen Ole e 1. *J. Allergy Clin. Immunol.* 103 (1999) 147–153.

14) Inoue K, Motozaki A, Takeuchi Y, Nishimura M, Hara-Nishimura I. Molecular characterization of proteins in proteinbody membrane that disappear most rapidly during transformation of protein bodies into vacuoles. *Plant J.* 7 (1995) 235–243.

15) Satoh J, Kato K, Shinmyo A. The 5'-untranslated region of the tobacco alcohol dehydrogenase gene functions as an effective translational enhancer in plant. *J Biosci Bioeng.* 98 (2004) 1-8.

16) Murashige, Toshio, Skoog, Folke. A Revised Medium for Rapid Growth and Bio Assays with Tobacco Tissue Cultures. *Physiologia Plantarum. 15* (1962) 473–497.

17) Hans C. Rempel, Louise M. Nelson. Analysis of conditions for *Agrobacterium*-mediated transformation of tobacco cells in suspension. *Transgenic Res.* 4 (1995) 199-207.

18) Stevens LH, Stoopen GM, Elbers IJ, Molthoff JW, Bakker HA, Lommen A, Bosch D, Jordi W. Effect of climate conditions and plant developmental stage on the stability of antibodies expressed in transgenic tobacco. *Plant Physiol.* 124 (2000) 173-182.

19) Misaki R, Sakai Y, Omasa T, Fujiyama K, Seki T. N-terminal vacuolar sorting signal at the mouse antibody alters the *N*-linked glycosylation pattern in suspension-cultured tobacco BY2 cells. *J Biosci Bioeng.* 112 (2011) 476-484.

20) Bolte S, Talbot C, Boutte Y, Catrice O, Read ND, Satiat-Jeunemaitre B. FM-dyes as experimental probes for dissecting vesicle trafficking in living plant cells. *J Microsc.* 214 (2004) 159-173.

21) Xie B, Deng Y, Kanaoka MM, Okada K, Hong Z. Expression of Arabidopsis callose synthase 5 results in callose accumulation and cell wall permeability alteration. *Plant Sci.* 183 (2012) 1-8.

22) Jun L, Saiki R, Tatsumi K, Nakagawa T, Kawamukai M. Identification and subcellular localization of two solanesyl diphosphate synthases from Arabidopsis thaliana. *Plant Cell Physiol.* 45 (2004) 1882-1888.

23) Laemmli UK. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227 (1970) 680-685.

24) Matsuoka K, Nakamura K. Large alkyl side-chains of isoleucine and leucine in the NPIRL region constitute the core of the vacuolar sorting determinant of sporamin precursor. *Plant Mol Biol.* 41 (1999) 825-835.

25) Gunnar von Heijne. The Signal Peptide. *J.Membrane Biol.* 155 (1990) 195-201.

26) Xiang L, Etxeberria E, Van den Ende W. Vacuolar protein sorting mechanisms in plants. *FEBS J.* 280 (2013) 979-93.

27) Brown JC, Jolliffe NA, Frigerio L, Roberts LM. Sequence-specific, Golgi-dependent vacuolar targeting of castor bean 2S albumin. *Plant J.* 36 (2003) 711-719.

28) Frigerio L, Jolliffe NA, Di Cola A, Felipe DH, Paris N, Neuhaus JM, Lord JM, Ceriotti A, Roberts LM. 2001. The internal propeptide of the ricin precursor carries a

sequence-specific determinant for vacuolar sorting. *Plant Physiol.* 126 (2001) 167-175.

29) Nishizawa K, Maruyama N, Satoh R, Higasa T, Utsumi S. A vacuolar sorting determinant of soybean β-conglycinin β subunit resides in C-terminal sequence. *Plant Sci.* 167 (2004) 937-947.

30) Frigerio L, de Virgilio M, Prada A, Faoro F, Vitale A. Sorting of phaseolin to the vacuole is saturable and requires a short C-terminal peptide. *Plant Cell* 10 (1998) 1031-1042.

31) Nishizawa K, Maruyama N, Utsumi S. The C-terminal region of α' subunit of soybean β-conglycinin contains two types of vacuolar sorting determinants. *Plant Mol Biol.* 62 (2006) 111-125.

32) Saint-Jore-Dupas C, Gilbert MA, Ramis C, Paris N, Kiefer-Meyer MC, Neuhaus JM, Faye L, Gomord V. Targeting of proConA to the plant vacuole depends on its nine amino-acid C-terminal propeptide. *Plant Cell Physiol.* 46 (2005) 1603-1612.

33) Saalbach G, Jung R, Kunze G, Saalbach I, Adler K, Muntz K. Different legumin protein domains act as vacuolar targeting signals. *Plant Cell* 3 (1991) 695-708.

34) Matsuoka K, Nakamura K. Propeptide of a precursor to a plant vacuolar protein required for vacuolar targeting. *Proc Natl Acad Sci USA.* 88 (1991) 834-8.

35) Brown JC, Jolliffe NA, Frigerio L, Roberts LM. Sequence-specific, Golgi-dependent vacuolar targeting of castor bean 2S albumin. *Plant J.* 36 (2003) 711-9.

36) Frigerio L, de Virgilio M, Prada A, Faoro F, Vitale A. Sorting of phaseolin to the vacuole is saturable and requires a short C-terminal peptide. *Plant Cell* 10 (1998) 1031-42.

37) Saalbach G, Rosso M, Schumann U. The vacuolar targeting signal of the 2S albumin from Brazil nut resides at the C terminus and involves the C-terminal propeptide as an essential element. *Plant Physiol.* 112 (1996) 975-85.

38) Nishizawa K, Maruyama N, Satoh R, Fuchikami Y, Higasa T, Utsumi. A C-terminal sequence of soybean β-conglycinin α' subunit acts as a vacuolar sorting determinant in seed cell. *Plant J.* 34 (2003) 647-659.

39) Mori T, Maruyama N, Nishizawa K, Higasa T, Yagasaki K, Ishimoto M, Utsumi S. The composition of newly synthesized proteins in the endoplasmic reticulum determines the transport pathways of soybean seed storage proteins. *Plant J.* 40 (2004) 238-49.

40) Saalbach G, Jung R, Kunze G, Saalbach I, Adler K, Müntz K. Different legumin protein domains act as vacuolar targeting signals. *Plant Cell* 3 (1991) 695-708.

41) Maruyama N, Mun LC, Tatsuhara M, Sawada M, Ishimoto M, Utsumi S. Multiple vacuolar sorting determinants exist in soybean 11S globulin. *Plant Cell* 18 (2006) 1253-73.

42) Munro S, Pelham HR. A C-terminal signal prevents secretion of luminal ER proteins. *Cell* 48 (1987) 899-907.

43) Pelham HR. Evidence that luminal ER proteins are sorted from secreted proteins in a post-ER compartment. *EMBO J.* 7 (1988) 913-918.

44) Lane KT, Beese LS. Thematic review series: lipid posttranslational modifications. Structural biology of protein farnesyltransferase and geranylgeranyltransferase type I. *J Lipid Res.* 47 (2006) 681-699.

45) Ceriotti A, Pedrazzini E, Bielli A, Giovinazzo G, Bollini R, Vitale A. Assembly and intracellular transport of phaseolin, the major storage protein of *Phaseolus vulgaris* L. *J Plant Physiol.* 145 (1995) 648-653.

46) Ko TP, Ng JD, McPherson A. The three-dimensional structure of canavalin from jack bean (*Canavalia ensiformis*). *Plant Physiol.* 101 (1993) 729-744.

47) Foresti O, Frigerio L, Holkeri H, de Virgilio M, Vavassori S, Vitale A. A phaseolin domain involved directly in trimer assembly is a determinant for binding by the chaperone BiP. *Plant Cell* 15 (2003) 2464-2475.

48) Pedrazzini E, Giovinazzo G, Bielli A, de Virgilio M, Frigerio L, Pesca M, Faoro F, Bollini R, Ceriotti A, Vitale A. 1997. Protein quality control along the route to the plant vacuole. *Plant Cell* 9 (1997) 1869-1880.

49) Meusser B, Hirsch C, Jarosch E, Sommer T. ERAD: the long road to destruction. *Nat Cell Biol.* 7 (2005) 766-772.

50) Radauer C, Breiteneder H. Evolutionary biology of plant food allergens. *J Allergy Clin Immunol.* 120 (2007) 518-525.

51) Dunwell JM, Khuri S, Gane PJ. Microbial relatives of the seed storage proteins of higher plants: conservation of structure and diversification of function during evolution of the cupin superfamily. *Microbiol Mol Biol Rev.* 64 (2000) 153-79.

52) Castillo J, Rodrigo MI, Márquez JA, Zúñiga A, Franco L, A pea nuclear protein that is induced by dehydration belongs to the vicilin superfamily. *Eur J Biochem.* 267 (2000) 2156-65.

53) Castillo J, Genovés A, Franco L, Rodrigo MI. A multifunctional bicupin serves as precursor for a chromosomal protein of Pisum sativum seeds. *J Exp Bot.* 56 (2005) 3159-69.

54) Kurosaka A, Yano A, Itoh N, Kuroda Y, Nakagawa T, Kawasaki T. The structure of a neural specific carbohydrate epitope of horseradish peroxidase recognized by anti-horseradish peroxidase antiserum. *J Biol Chem.* 266 (1991) 4168-72.

55) Hijikata A, Matsumoto I, Kojima K, Ogawa H. Antigenicity of the oligosaccharide moiety of the Japanese cedar (Cryptomeria japonica) pollen allergen, Cry jI. *Int Arch Allergy Immunol.* 105 (1994) 198-202.

56) Ogawa H, Hijikata A, Amano M, Kojima K, Fukushima H, Ishizuka I, Kurihara Y, Matsumoto I. Structures and contribution to the antigenicity of oligosaccharides of Japanese cedar (Cryptomeria japonica) pollen allergen Cry j I: relationship between the structures and antigenic epitopes of plant N-linked complex-type glycans. *Glycoconj J.* 13 (1996) 555-66.

57) Batanero E, Villalba M, Monsalve RI, Rodríguez R. Crossreactivity between the major allergen from olive pollen and unrelated glycoproteins: evidence of an epitope in the glycan moiety of the allergen. *J. Allergy Clin. Immunol.* 97 (1996) 1264–1271.

58) Wilson IB, Harthill JE, Mullin NP, Ashford DA, Altmann F. Core α1,3-fucose is a key part of the epitope recognized by antibodies reacting against plant N-linked oligosaccharides

and is present in a wide variety of plant extracts. *Glycobiology* 8 (1998) 651–661.

59) Kimoto M, Suzuki M, Komiyama N, Kunimoto A, Yamashita H, Hiemori M, Takahashi K, Tsuji H. Isolation and molecular cloning of a major wheat allergen, Tri a Bd 27K. *Biosci Biotechnol Biochem.*73 (2009) 85-92.

60) Hiemori M, Yosida Y, Kimoto M, Yamashita H, Takahashi K, Takahashi K, Komiyama N, Tsuji H. Investigation of multiple forms of Tri a Bd 27K, a major wheat allergen, by immunoblotting analysis. *Biosci Biotechnol Biochem.* 74 (2010) 199-202.

61) Lowry O.H., J. Rosebrough, A.L. Farr, R.J. Randall. Protein measurement with the folin phenol reagent. *J. Biol. Chem.* 193 (1951) 265–275.

62) Tsuji H, Oka T, Kimoto M, Hong YM, Natori Y, Ogawa T. Cloning and sequencing of cDNA encoding 4-aminobenzozte hydroxylase from Agaricus bisporus. *Biochim. Biophys. Acta*. 1309 (1996) 31–36.

63) Takeuchi Y, Nishimura M. Studies on intracellular localization of proteins by immunoelectric microscopy using colloidal gold. *Plant Cell Technol.* 5 (1993) 115–123.

64) Tsuji H, Kimoto M, Watanabe H, Sasagawa T, Oka T, Yamashita H, Okita M. Epitope mapping of monoclonal antibodies against 4-aminobenzoate hydroxylase from Agaricus bisporus. *Biochim. Biophys. Acta.* 1425 (1998) 628–631.

65) Goldstein IJ, Hollerman CE, Smith EE. Protein-carbohydrate interaction: II. Inhibition studies on the interaction of concanavalin A with polysaccharides. *Biochemistry* 41 (1965) 876–883.

66) Poretz RD, Goldstein IJ. An examination of the topography of the saccharide binding sites of concanavalin A and of the forces involved in complexation. *Biochemistry* 9 (1970) 2890–2896.

67) McManus MT, McKeating J, Secher DS, Osborne DJ, Ashford D, Dwek RA, Rademacher TW. Identification of a monoclonal antibody to abscission tissue that recognizes xylose/fucose containing N-linked oligosaccharides from higher plants. *Planta* 175 (1988) 506–512.

68) Kurosaka A, Yano A, Itoh N, Kuroda Y, Nakagawa T, Kawasaki T. The structure of a neural specific carbohydrate epitope of horseradish peroxidase recognized by anti-horseradish peroxidase antiserum. *J. Biol. Chem.* 266 (1991) 4168–4172.

69) Prenner C, Mach L, Glössl J, März L. The antigenicity of the carbohydrate moiety of an insect glycoprotein, honey-bee (Apismellifera) venom phospholipase A2. The role of alpha 1,3-fucosylation of the asparagine-bound N-acetylglucosamine. *Biochem. J.* 284 (1992) 377–380.

70) Thanh VH, Shibasaki K. Major proteins of soybean seeds. A straightforward fractionation and their characterization. *J. Agric. Food Chem.* 24 (1976) 1117–1121.

71) Hara-Nishimura I, Nishimura M. Proglobulin processing enzyme in vacuoles isolated from developing pumpkin cotyledons. *Plant Physiol.* 85 (1987) 440–445.

72) Hara-Nishimura I, Inoue K, Nishimura M. A unique vacuolar processing enzyme responsible for conversion of several proprotein precursor into the mature forms. *FEBS*

Lett. 294 (1991) 89–93.

73) Scott MP, Jung R, Muntz K, Nielsen NC. A protease responsible for post-translational cleavage of a conserved Asn-Gly linkage in glycinin, the major seed storage protein of soybean. *Proc. Natl. Acad.Sci.* 15 (1992) 658–662.

74) Shimada T, Hiraiwa N, Nishimura M, Hara-Nishimura I. Vacuolar processing enzyme of soybean that converts proproteins to the corresponding mature forms. *Plant Cell Physiol.* 35 (1994) 713–718.

75) Robinson AB. Evolution and the distribution of glutaminyl and asparaginyl residues in proteins. *Proc. Natl. Acad. Sci.* 71 (1974) 885–888.

76) O'Farrell PH. High resolution two-dimensional electrophoresis of proteins. *J. Biol. Chem.* 250 (1975) 4007–4021.

77) van Ree R, Cabanes-Macheteau M, Akkerdaas J, Milazzo JP, Loutelier-Bourhis C, Rayon C, Villalba M, Koppelman S, Aalberse R, Rodriguez R, Faye L, Lerouge P. β (1, 2)-xylose and α (1, 3)-fucose residues have a strong contribution in IgE binding to plant glycoallergens. *J. Biol. Chem.* 275 (2000) 11451–11458

78) Altmann F. The role of protein glycosylation in allergy. *Int Arch Allergy Immunol.* 142 (2007) 99-115.

79) Jin C, Altmann F, Strasser R, Mach L, Schähs M, Kunert R, Rademacher T, Glössl J, Steinkellner H. A plant-derived human monoclonal antibody induces an anti-carbohydrate immune response in rabbits. *Glycobiology* 18 (2008) 235-41.

80) Bouwstra JB, Spoelstra EC, De Waard P, Leeflang BR, Kamerling JP, VliegenthartJF. Conformational studies on the N-linked carbohydrate chain of bromelain. *Eur J Biochem.* 190 (1990) 113-22.