

N-Glycosylation and *N*-Glycan Moieties of CTB Expressed in Rice Seeds

Hiroyuki Kajiura · Masashi Wasai ·
Saori Kasahara · Fumio Takaiwa · Kazuhito Fujiyama

Published online: 5 December 2012
© Springer Science+Business Media New York 2012

Abstract Cholera toxin B subunit (CTB) is widely used as a carrier molecule and mucosal adjuvant and for the expression of fusion proteins of interest. CTB-fusion proteins are also expressed in plants, but the *N*-glycan structures of CTB have not been clarified. To gain insights into the *N*-glycosylation and *N*-glycans of CTB expressed in plants, we expressed CTB in rice seeds with an N-terminal glutelin signal and a C-terminal KDEL sequence and analyzed its *N*-glycosylation and *N*-glycan structures. CTB was successfully expressed in rice seeds in two forms: a form with *N*-glycosylation at Asn32 that included both plant-specific *N*-glycans and small oligomannosidic *N*-glycans and a non-*N*-glycosylated form. *N*-Glycan analysis of CTB showed that approximately 50 % of the *N*-glycans had plant-specific M3FX structures and that almost none of the *N*-glycans was of high-mannose-type *N*-glycan even though the CTB expressed in rice seeds contains a C-terminal KDEL sequence. These results suggest that the CTB expressed in rice was *N*-glycosylated through the endoplasmic reticulum (ER) and Golgi *N*-glycosylation machinery without the ER retrieval.

Keywords CTB · *N*-Glycosylation · *N*-Glycan · Rice

H. Kajiura · K. Fujiyama (✉)
International Center for Biotechnology, Osaka University,
2-1 Yamada-oka, Suita-shi, Osaka 565-0871, Japan
e-mail: fujiyama@icb.osaka-u.ac.jp

M. Wasai · S. Kasahara
Pulp and Paper Research Laboratory, Nippon Paper Industries
Co., Ltd., 5-21-1, Oji, Kita-ku, Tokyo 114-0002, Japan

F. Takaiwa
Transgenic Crop Research and Development Center, National
Institute of Agrobiological Sciences, Kannondai 2-1-2,
Tsukuba, Ibaraki 305-8602, Japan

Introduction

Cholera toxin (CT) from *Vibrio cholerae* exhibits the multi-functional attributes of a single toxin A and five identical B subunits (CTB). The toxin A subunit is composed of two distinct domains, A1 and A2, which are responsible for the enzymatic activity and anchoring of A1 to pentameric CTB, respectively, whereas CTB has no functional enzymatic activity but contributes to the binding of CT to the apical membrane [1–4]. CT uptake, delivery, and activation are initiated via G_{M1} ganglioside–CTB binding [1, 4]. G_{M1} ganglioside is one of the glycosphingolipids and is composed of five sugar residues (Galβ1,3-GalNAcβ1,4-(NeuAcα2,6-)Galβ1,4-Glcβ1,1-ceramide) and distributed to the lipid raft fraction of cell surface. CTB has an affinity against this G_{M1} ganglioside, CTB binding mediates the transportation of holotoxin to the host cells. CTB itself lacks toxicity but has the affinity for the G_{M1} ganglioside and is highly immunogenic by an oral route. Therefore, CTB has been widely used as a carrier protein and mucosal adjuvant to increase the uptake of CTB fusion protein and/or antigen to stimulate the mucosal immunity. In addition to this function, CTB has potential as an immunogen for use in oral vaccines.

Recent studies have shown that CTB fusion proteins are expressed in plants [5]. Plants have received attention as a platform for a heterologous protein expression system [6–8]. Plant molecular farming enables cost-effective and environmentally friendly protein production and is free from human pathogen. Molecular farming in plants is expected to be a challenging task compared with the use of already established protein production systems such as bacteria, yeast, and cultured mammalian cells due to its low expression rate and time requirements [9–11]. For efficient plant molecular farming, several strategies are now under

development: large-scale protein production, a high-expression system, and optimization of post-translational modifications including *N*-glycosylation [9]. These strategies enable the effective protein production in plants. CTB and its fusion proteins were also produced in tobacco plants [12, 13], rice seeds [14, 15], potatoes [16–18], and lettuce plants [19, 20]. These CTB fusion proteins showed *in vitro* and *in vivo* activities as expected. Thus, CTB fusion proteins were considered to be effective and useful molecules for human use. However, plant and plant-made CTB fusion proteins have potential as agents for the induction of immune responses mainly mediated by plant-specific β 1,2-xylose and/or α 1,3-fucose residues on *N*-glycans.

In most cells, a majority of proteins are synthesized as glycoproteins through a post-translational modification, glycosylation. Glycosylation is a highly sophisticated post-translational modification in the endoplasmic reticulum (ER) and Golgi apparatus and is categorized into two types: *N*-glycosylation and *O*-glycosylation. *N*-Glycosylation is now well characterized and the synthetic pathways of *N*-glycans in mammals, insects, yeasts, and plants have been proposed [21–25]. In mammals, *N*-glycans contain β 1,4-galactose and *N*-acetylneuraminic acid (NeuAc) residue(s) at the non-reducing terminus of oligosaccharide and bisecting GlcNAc residue, whereas plant complex *N*-glycans do not have NeuAc residues, but rather carry β 1,2-xylose and α 1,3-fucose residues. In addition to the β 1,2-xylose and/or α 1,3-fucose residues, Lewis a structure that is found on the cell-surface glycolipid antigens of red blood cells in humans has also been observed in plants [26, 27]. These plant-specific sugar residues were transferred to *N*-glycan and the transfer was mediated by β 1,2-xylosyltransferase (XYLT), two α 1,3-fucosyltransferase(s) (FUCT), α 1,4-FUCT, and β 1,3-galactosyltransferase. A recent study provided evidence that these plant-specific β 1,2-xylose and/or α 1,3-fucose residues on plant-derived glycoprotein may induce immune responses and, in some cases, result in induction of allergy because these glycan-epitopes are absent in humans [28, 29]. Furthermore, β 1,2-xylose and α 1,3-fucose residues are considered as a IgE-binding carbohydrate determinants, the so-called cross-reactive carbohydrate determinants of plant allergens [29, 30]. Hence, an improved understanding of the *N*-glycans on proteins expressed in plants is critical for human use to avoid induction of immune response.

N-Glycosylation analysis of CTB fusion proteins expressed in plants has been performed in a few studies [13], revealing that the expression of CTB fusion proteins with a C-terminal ER-retention signal, (SE)KDEL, was an effective system for the suppression of β 1,2-xylose and/or α 1,3-fucose modification on *N*-glycans. In addition, these expression systems led to an accumulation of *N*-glycoproteins with a high-mannose-type structure, which are less immunogenic for human. However, these studies analyzed

the CTB fusion proteins but not CTB itself. CTB has one potential *N*-glycosylation site (Asn-Xxx-Ser/Thr) in its amino acid sequence, which is also considered to be *N*-glycosylated. The *N*-glycosylation of the site and the *N*-glycan structures are still unknown. To gain insights into the correlation between the CTB proteins expressed in plants and the *N*-glycan structures of these proteins for medical use, an understanding of the *N*-glycan modification of functional CTB molecules is indispensable.

In this study, to clarify the *N*-glycosylation of CTB, we expressed CTB in rice seeds with a GluB-1 signal peptide and KDEL sequence at the N- and C-terminus, respectively, and the analyzed *N*-glycosylation and *N*-glycan structures of CTB. CTB expressed in rice seeds was *N*-glycosylated at Asn32 with immunoreactive glycan structures against anti-HRP antibody, but some of the CTB was not glycosylated. In addition, *N*-glycosylated CTB carried PNGase F- and Endo H-insensitive *N*-glycan structures, indicating that the *N*-glycans of CTB were modified with β 1,2-xylose and/or α 1,3-fucose residue(s). The detailed *N*-glycan analysis revealed that the majority of the structures were plant-specific M3FX and that the *N*-glycans contained small amounts of high-mannose-type *N*-glycan. These results suggested that the CTBs expressed in rice seeds were *N*-glycosylated through the ER and Golgi *N*-glycosylation machineries without the ER retrieval even though the CTBs expressed in rice seeds contain the C-terminal KDEL sequence.

Materials and Methods

Construction of CTB Expression Vector and Rice Transformation

The gene cassette consisted of the endosperm-specific 2.3-kb GluB-1 promoter, a coding sequence for the signal peptide of GluB-1 and the codon-optimized CTB coding gene of *V. cholerae*, and a 0.65-kb GluB1 terminator [31, 32], and it was cloned into the *EcoRI*-*Sse8387I* site of pTL7 [33] to produce pTL-CTB. The *Sse8387I* fragment of pNPI130Hm [33] was cloned into the *Sse8387I* site of pTL-CTB to generate pTL-CTB130Hm. A binary vector, pTL-CTB130Hm, was transformed in rice plants, *Oryza sativa* L.cv. Nipponbare, using an agrobacterium-mediated method. Transformation of rice plants and regeneration of shoots were carried out according to the procedure described previously [34]. The transgenic rice plants were grown under controlled environment conditions and water with a nutrient solution in a closed green house. Plants were grown with a 12-h photoperiod and 50 % relative humidity in 23–28 °C air during the photo-/dark period.

Expression Analysis of CTB and Its Purification

Soluble CTB proteins were extracted as described previously [35]. The crude protein samples were subjected to 15 % SDS-PAGE under reducing conditions, followed by CBB staining and western blotting. In western blotting, the proteins fractionated by SDS-PAGE were electroblotted onto a PVDF membrane. The membrane was incubated in phosphate-buffered saline (PBS) with 5 % skim milk at room temperature and probed with anti-CTB antibody and anti-KDEL antibody (StressMarq, Canada) as primary antibodies and anti-mouse IgG antibody and anti-rabbit IgG conjugated to horseradish peroxidase (GE Healthcare, Buckinghamshire, England) as secondary antibodies, respectively. Specific bands were visualized using the POD Immunostain Set (Wako, Osaka, Japan).

Rice seeds were ground to a fine powder and total proteins were extracted in 50 mM Tris–HCl buffer (pH 7.5), 500 mM NaCl, 5 mM imidazole, and 5 % 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulphonate (CHAPS) to solubilize CTB by vortex and centrifuged at $17,000\times g$ for 15 min at room temperature. The supernatant was subjected to a TALON[®] Resin column (TaKaRa, Shiga, Japan) and purified following the manufacturer's protocols.

N-Glycosylation Analysis by Western Blotting and Endoglycosidase Digestions

The purified CTB was separated by 15 % SDS-PAGE and detected by CBB staining or western blotting using anti-CTB antibody or anti-HRP antibody as described above. For N-glycosylation analysis, anti-HRP antibody (Sigma, St. Louis, MO) and anti-rabbit antibody conjugated to horseradish peroxidase (GE Healthcare) were used as a primary antibody and a secondary antibody, respectively, and visualized as described above.

The purified CTBs were digested with peptide:N-glycosidase F (PNGase F, TaKaRa) or endoglycosidase H (Endo H) (NEB, Tokyo, Japan) following the manufacturer's protocols, followed by separation by 15 % SDS-PAGE and CBB staining.

NanoLC–MS/MS Analysis of CTB for De Novo Sequencing

The purified CTB was separated by 15 % SDS-PAGE and stained by CBB. The CBB-stained bands were excised from the gel, destained with 50 mM NH_4HCO_3 in 50 % acetonitrile and dehydrated using acetonitrile, followed by reduction and alkylation. The CTB was digested using chymotrypsin (Sigma) in ProteaseMAX[™] Surfactant (Promega, Madison, WI) at 37 °C for 3 h.

The chymotrypsinized samples containing 5 % acetonitrile were analyzed using an Agilent Technologies 1200 series nanoLC system (Agilent Technologies, Santa Clara, CA) equipped with microTOF-QII (BRUKER DALTONICS, Bremen, Germany). For the liquid chromatography portion of the analysis, a ZORBAX 300SB-C18 (5 μm , 0.3 mm \times 5 mm) and ZORBAX 300SB-C18 (3.5 mm, 75 μm \times 150 mm) (Agilent Technologies) were used as a trapping column and an analytical column, respectively. The mobile phase was composed of 0.1 % formic acid (solvent A) and acetonitrile containing 0.1 % formic acid (solvent B) for nanoLC and 0.1 % trifluoroacetic acid (solvent C) for peptide trapping in the column. Following injection, the flow was directed to the trapping column in solvent C at a flow rate 10 $\mu\text{l}/\text{min}$. The CTB peptides were separated by linearly increasing the solvent B concentration from 8 to 30 % over 30 min at a flow rate of 300 nL/min, followed by washing with 95 % solvent B for 5 min, and equilibration for 18 min with initial flow. In the mass spectrometry portion of the analysis, the MS/MS parameters were as follows: scan range m/z 50–4500, nebulizer flow of 1.0 bar, dry gas flow rate of 5.0 l/min, and dry temperature of 180 °C in the positive-ion mode. The MS data were analyzed using DataAnalysis 4.0 software (BRUKER DALTONICS).

Preparation of N-Glycans and Their Structural Analysis by HPLC and LC–MS/MS

The N-glycopeptide from CTB was obtained by in-gel digestion as described above and completely dried out. The N-glycans were released from the in-gel-digested CTB by hydrazinolysis at 100 °C for 10 h. The released N-glycans were 2-aminopyridine (PA)-labeled as described previously [36]. The reaction products of PA-N-glycans were purified by size fractionation (SF)-high-performance liquid chromatography (HPLC) using a Shodex Asahipak NH2P-50 4E column (4.6 mm ID \times 250 mm; SHOWA DENKO Co., Ltd., Tokyo, Japan). The N-glycans were detected by reverse phase (RP)-HPLC. The mobile phase was composed of 0.02 % trifluoroacetic acid (solvent A) and acetonitrile/0.02 % trifluoroacetic acid (solvent B) (20/80, v/v). RP-HPLC was performed using a Cosmosil 5C₁₈-AR-II column (4.6 \times 250 mm; Nacalai Tesque, Kyoto, Japan) with an HITACHI LaChrom HPLC System by linearly increasing the solvent B concentration from 0 to 20 % over 35 min at a flow rate of 0.7 ml/min. The eluted fractions were monitored by measuring the fluorescence intensity using excitation and emission wavelengths of 310 and 380 nm, respectively.

The molecular masses of the PA-N-glycans and the number of their sugar moieties were estimated by LC–MS/MS using an Agilent Technologies 1200 series instrument (Agilent Technologies, Santa Clara, CA) equipped with

HCT plus software (BRUKER DALTONICS). For the LC, the mobile phase was composed of acetonitrile/acetic acid (solvent A: 98/2, v/v) and water/acetic acid/triethylamine (solvent B: 92/5/3, v/v/v). The PA-sugar chain was separated using a Shodex Asahipak NH2P-50 2D column (2.0 mm ID × 150 mm; SHOWA DENKO Co., Ltd.) by linearly increasing the solvent B concentration from 20 to 55 % over 35 min at a flow rate of 0.2 ml/min. The MS/MS parameters were as follows: scan range m/z 350–2750, nebulizer flow of 5.0 psi, dry gas flow rate of 3.0 l/min, dry temperature of 300 °C, target count of 200,000, and MS/MS Frag. Ampl. of 1.0 V in the positive-ion mode. The relative amount of *N*-glycan detected in each plant was calculated on the basis of the peak area of the LC.

To determine the PA-*N*-glycan structures, PA-*N*-glycans were further separated by SF-HPLC under the following conditions; the mobile phase was composed of 80 % acetonitrile (solvent A) and 20 % acetonitrile (solvent B) (20/80, v/v) using a Shodex Asahipak NH2P-50 4E column with the an HITACHI LaChrom HPLC System by linearly increasing the solvent B concentration from 10 to 60 % over 25 min at a flow rate of 0.7 ml/min. Each structure of PA-*N*-glycans was compared with authentic PA-*N*-glycans purchased from TaKaRa or prepared from *Arabidopsis thaliana* Col-0 WT as previously reported [37].

Results

CTB was Expressed in Rice

A binary vector, pTL-CTB130Hm (Fig. 1), was introduced into rice plants (*Oryza sativa* L.cv. Nipponbare) mediated by agrobacterium to produce transgenic rice plants expressing CTB. The insertion of the expression cassette into genome DNA and CTB expression was analyzed by PCR using genomic DNA as a template and western blotting using anti-CTB antibody, respectively (data not shown). The transgenic

rice plants expressing the highest level of CTB were grown in a closed green house for about 4 months, and mature seeds that accumulated CTB 3.37 mg/g rice seeds were harvested.

CTB was Expressed in Two-Forms: An *N*-Glycosylated Form and Non-*N*-glycosylated Form

CTB contains a potential *N*-glycosylation site, Asn32, in its expressed sequence. This indicates that CTB might have been expressed as an *N*-glycoprotein in rice. Furthermore, western blotting analysis of transgenic rice using an anti-CTB antibody demonstrated that CTB was expressed in two forms in rice, which led us to hypothesize that one of the CTBs, i.e., the one-detected by the upper band (band 1), had been *N*-glycosylated in rice (Fig. 2a). To ascertain whether these two forms of CTB were derived from *N*-glycosylated and/or non-*N*-glycosylated CTB and to further characterize them, rice-derived CTB was first purified using a Co^{2+} column. A previous study demonstrated that CTB could be purified with Co^{2+} column, based on the formation of pentameric CTB [13]. In this study, CBB staining and western blotting using anti-CTB antibody and anti-KDEL antibody showed that the CTB expressed in rice carried KDEL tag at their C-terminus and successfully purified (Fig. 2b, c). Furthermore, western blotting analysis using an anti-HRP antibody that recognizes plant-specific *N*-glycoproteins carrying β 1,2-xylose and/or α 1,3-fucose residue(s) revealed that CTB band 1 was immune-sensitive against anti-HRP antibody. The lower band (band 2), on the other hand, was not immune-insensitive (Fig. 2d). These results indicated that one of the forms of CTB expressed in rice was *N*-glycosylated with plant-specific sugar residue(s) and that the other CTB forms were not *N*-glycosylated. Furthermore, CTB was expressed as pentameric protein with *N*-glycosylated and/or non-*N*-glycosylated CTBs in rice seeds (Fig. 2e). This agreed well with the result of western blotting which indicated that CTBs could be purified using Co^{2+} column by the formation of pentameric complex.

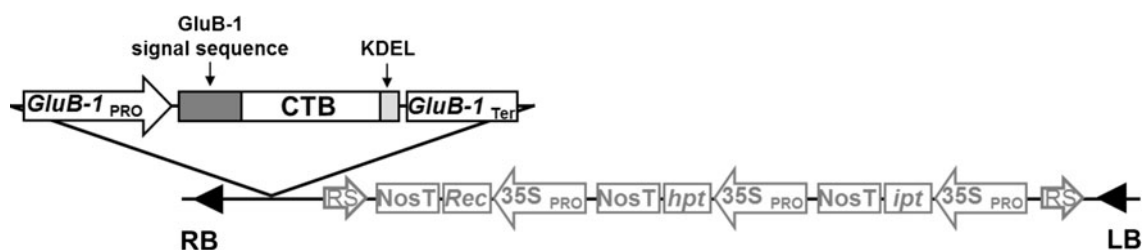
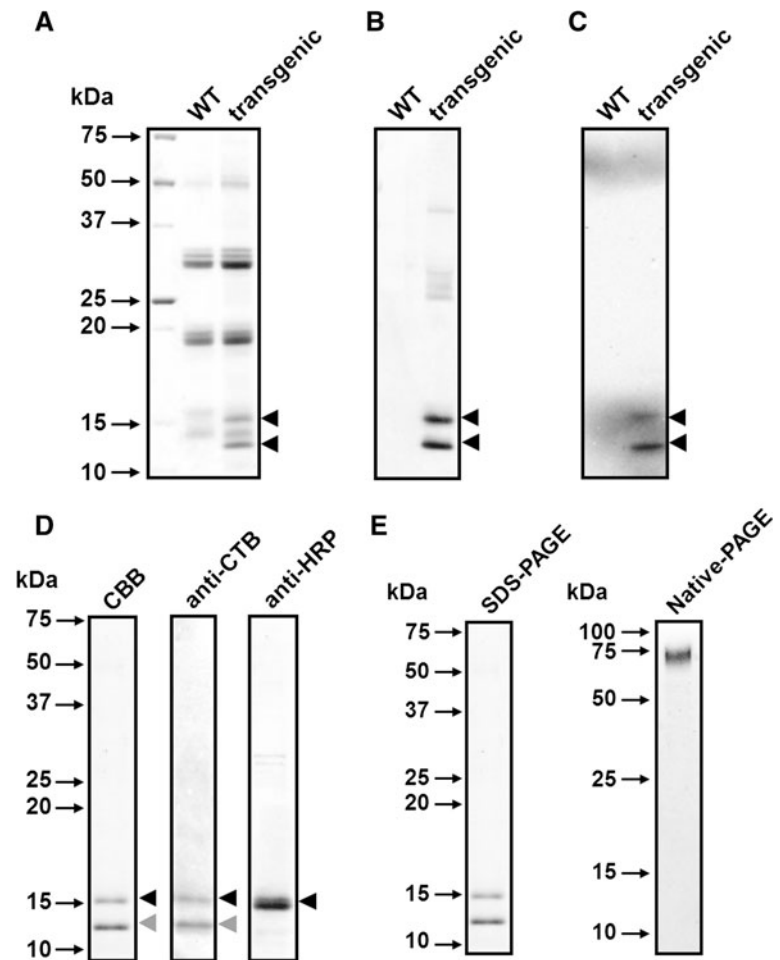


Fig. 1 Schematic representation of the plant expression vector, pTL-CTB130Hm. The original vector used for expression of CTB was the MAT vector (indicated in gray) [32]. CTB was expressed under the control of the glutelin B-1 promoter (GluB-1_{PRO}) and the glutelin B-1 terminator (GluB-1_T) with the signal peptide of GluB-1 at the N-terminus and KDEL at the C-terminus, respectively, for enhancement

of the stable expression and the accumulation of CTBs. This construct also contains the hygromycin phosphotransferase gene, *hpt*, as a selection marker and isopentenyl transferase gene, *ipt*, and recombinase gene, *Rec*, for the production of marker-free transgenic rice [32]. *RS* recognition sequence; *35S_{PRO}* cauliflower mosaic virus 35S promoter, *NosT* nopaline synthase terminator, *RB* and *LB* right and left borders, respectively

Fig. 2 Detection and analysis of the post-translational modification of CTB. Total proteins from rice seeds of WT and transgenic rice were extracted and analyzed by CBB staining (a) or western blotting using anti-CTB antibody (b) and anti-KDEL antibody (c). *Black triangles* indicate the CTB. **d** Detection of the *N*-glycosylated-form of CTB. Purified CTB was used for detection of the immunoreactivity against the anti-CTB antibody and anti-HRP antibody. *Black and gray triangles* indicate the *N*-glycosylated form and non-*N*-glycosylated form of CTB, respectively. **e** Pentameric forms of CTB. Purified CTB was analyzed by SDS-PAGE and native-PAGE.



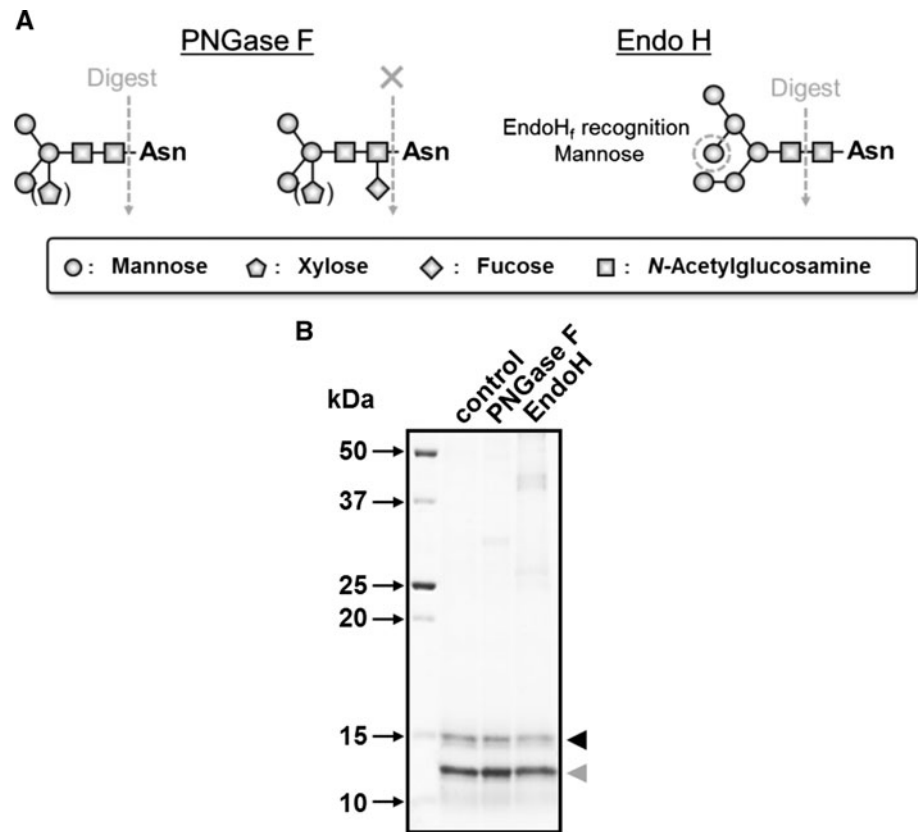
For further characterization of the *N*-glycan moieties of the CTBs expressed in rice, purified CTB was digested with two *N*-glycan-specific endoglycosidases, peptide:*N*-glycosidase F (PNGase F) and endoglycosidase H (Endo H). PNGase F cleaves *N*-glycans off at the GlcNAc-Asn linkage of non- α 1,3-fucosylated *N*-glycoprotein, but does not cleave *N*-glycan off if the *N*-glycan contains α 1,3-fucose residue (Fig. 3a). Endo H recognizes the terminal mannose residue of Man α 1,3-Man α 1,6-Man β 1,4-GlcNAc β 1,4-GlcNAc of high-mannose-type and/or hybrid-type *N*-glycans and cleaves the acetylchitobiose core of oligosaccharide (Fig. 3a). Therefore, we can easily evaluate the *N*-glycan structure(s) of CTB by PNGase F and Endo H digestion. PNGase F or Endo H digestion of CTB resulted in no obvious band shift and no accumulation of the non-glycosylated CTB form, respectively (Fig. 3b). These results indicate that most of the CTB carried α 1,3-fucosylated *N*-glycan(s) but did not *N*-glycosylate with high-mannose-type *N*-glycan(s).

Asn32 of CTB was *N*-Glycosylated

As mentioned above, the CTB expression sequence used in this study contained a consensus *N*-glycosylation sequence,

Asn-Xxx-Ser/Thr, at Asn32 to Thr34. This sequence and the results of western blotting using anti-HRP antibody (Fig. 2d) suggested that CTB was expressed as an *N*-glycoprotein. To determine the *N*-glycosylation site of CTB, in-gel digestion of purified CTB using chymotrypsin was performed for de novo sequencing using nanoLC-MS/MS analysis. A MASCOT search (http://www.matrixscience.com/search_form_select.html) revealed that both of the two forms of CTB expressed in rice corresponded to the database sequence at a sequence coverage of 51 % (data not shown). Furthermore, deconvoluted MS/MS signals showed two fragmentation patterns derived from free *N*-glycan and *N*-glycopeptide from CTB, respectively (Fig. 4). Fragmentation analysis of free *N*-glycan exhibited that the *N*-glycan of CTB mainly consisted of five sugar residues—three hexoses, one deoxyhexose, and one pentose in the core structure of GlcNAc₂-Asn. This indicates that CTB was considered to be *N*-glycosylated with xylose and fucose residues. In addition, fragmentation analysis of *N*-glycopeptide showed that the fragmented *m/z* 3395.51 agreed well with the calculated molecular mass of weight of amino acids Met27 to Tyr55 with single methylation. In this amino acid sequence, Asn32 was a good candidate for

Fig. 3 *N*-Glycosylation analysis of CTB. **a** Schematic representation of the substrate specificity and recognition site of endoglycosidases. **b** Endoglycosidase digestion of CTB. Reaction products of endoglycosidase digestion were detected by CBB staining. *Black* and *gray triangles* indicate the *N*-glycosylated form and non-*N*-glycosylated form of CTB, respectively



N-glycosylation. Therefore, these results strongly suggested that CTB was *N*-glycosylated at Asn32 with the plant-specific sugar residues, xylose and fucose.

CTB Expressed in Rice was Mainly Modified with Plant-Specific *N*-Glycans

The results of both western blotting using anti-HRP antibody and de novo sequencing suggested that one of the forms of CTB (detected as band 1) was *N*-glycosylated. To determine the CTB profile in greater detail, the *N*-glycan structures attached at Asn32 were determined by hydrazinolysis of *N*-glycopeptide obtained by in-gel digestion of both bands detected in CBB staining, followed by PA-labeling, RP-HPLC and SF-HPLC analysis, and LC-MS/MS analysis (Fig. 5a, b; Table 1). RP-HPLC analysis of PA-*N*-glycan showed that band 1 (the upper band) contained three peaks, but band 2 (the lower band) contained no obvious peaks, indicating that band 2 was not *N*-glycosylated. This result further supported the results of the western blotting using anti-HRP antibody, which demonstrated that band 2 was the non-*N*-glycosylated form of CTB. The major *N*-glycan in the band 1 was M3FX, which is generally the most abundant plant-specific *N*-glycan (Fig. 5c, d). Furthermore, M3X was also detected as a plant-specific *N*-glycan. The total quantity of *N*-glycans carrying plant-specific sugar residues, β 1,2-xylose and/or α 1,3-fucose,

accounted for approximately 70 % of the total *N*-glycans detected in CTB. In general, these β 1,2-xylosylations and α 1,3-fucosylations took place in the Golgi. Therefore, most of the CTB was *N*-glycosylated via Golgi-*N*-glycosylation machinery. These findings also suggested that the KDEL sequence at the C-terminus was not functional in the CTB retrieval for the ER. Interestingly, in addition to the plant-specific *N*-glycan structures, CTB also contained terminal-mannosylated-*N*-glycans, M3, M4, M5, and M6B. These structures were synthesized by the Golgi- or vacuole-localized glycosylhydrolases, α 1,2-mannosidase I and/or α -mannosidase. Together, these results strongly supported the idea that CTB was *N*-glycosylated through the Golgi and/or vacuole *N*-glycosylation machineries against the ER retrieval mediated by the C-terminal KDEL sequence.

Discussion

CTB is an ideal and a good-candidate molecule as not only carrier protein, adjuvant, but also the elicitor of immune response. Thus, the large-scale and low-cost production of CTB or its fusion proteins for use in medical applications is a challenging task. Previously, the CTB fusion protein of MPR_{649–684}, which is membrane proximal region of gp41 from HIV virus, was successfully expressed in *N. benthamiana* plants and its *N*-glycosylation and *N*-glycans were analyzed

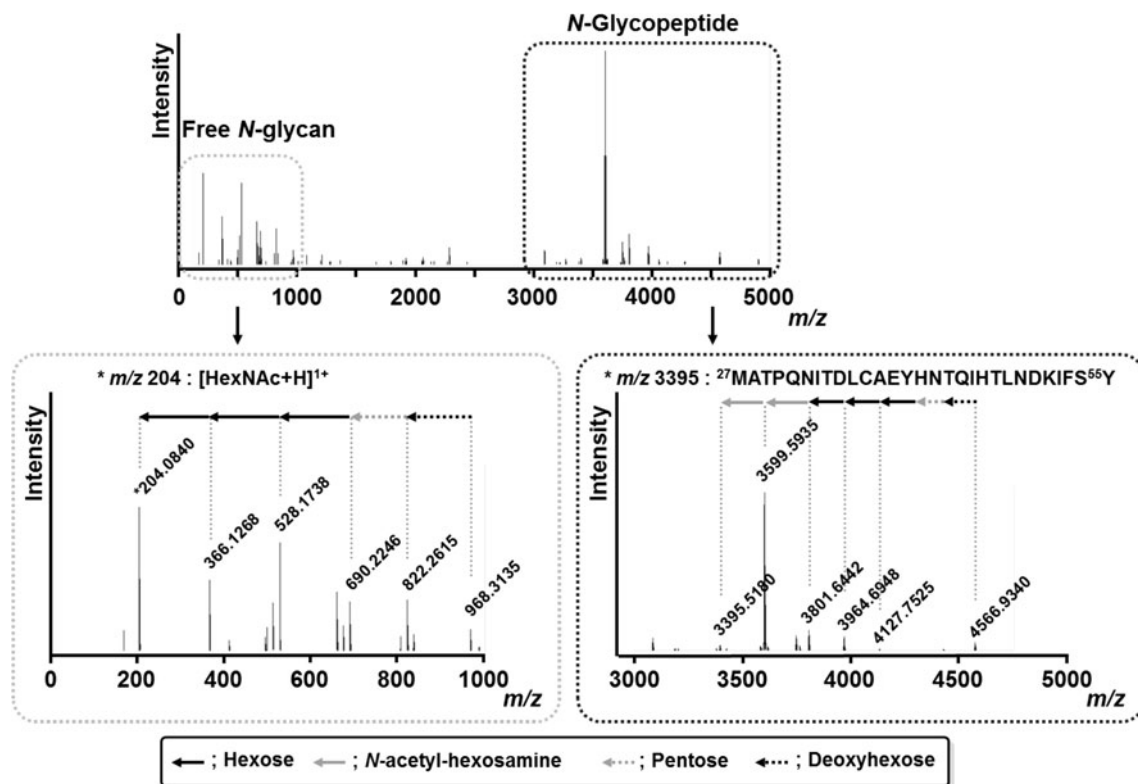


Fig. 4 Determination of the *N*-glycosylation site and the components of sugar residues in CTB *N*-glycan. Chymotrypsin-digested *N*-glycosylated CTB was subjected to nanoLC–MS/MS analysis. The upper panel shows deconvoluted MS/MS spectrums of free-sugar residues from the CTB *N*-glycan and *N*-glycosylated peptide of CTB. The lower panels show the enlarged spectrums of free-sugar residues

and *N*-glycosylated peptide, respectively. The black arrow, black dotted arrow, gray arrow, and gray dotted arrow indicate hexose, deoxyhexose, *N*-acetyl-hexosamine, and pentose, respectively. *m/z* 204 and 3395 are the calculated masses of the $[\text{HexNAc} + \text{H}]^{1+}$ and peptide sequence of Met27 to Tyr55

[13]. However, the *N*-glycosylation and *N*-glycans of plant-derived CTB itself remain to be clarified. Here, focusing on the large-scale protein production in plants and its *N*-glycosylation, we expressed CTB in rice seeds and analyzed to evaluate whether the *N*-glycosylation occurred at the potential *N*-glycosylation site, and whether the plant-derived CTB contains high-mannose-type structures like those detected in CTB–MPR_{694–684} or immunogenic β 1,2-xylose and/or α 1,3-fucose residue(s). Recombinant CTB was *N*-glycosylated at Asn32 with a plant-specific *N*-glycan carrying β 1,2-xylose and/or α 1,3-fucose residue(s). The presence of β 1,2-xylose and/or α 1,3-fucose residue(s) on *N*-glycans strongly indicated that CTB was *N*-glycosylated with Golgi *N*-glycosylation machinery, because XYLT and FUCTs are considered to localize at medial-Golgi [38–40]. Furthermore, high-mannose-type *N*-glycan was not observed in RP-HPLC analysis of PA-sugar chain, providing evidence that the CTB expressed in rice seeds escaped from ER-retrieval, even though this CTB had a KDEL sequence at its C-terminal, and that the CTB was transferred to the Golgi and vacuole *N*-glycosylation machineries.

Why did the CTB expressed in rice seeds escape from ER-retrieval? In general, the subcellular localization of heterologous proteins expressed in plants is determined by

appropriate signal tag or endogenous targeting and/or sorting sequences. Recombinant proteins expressed by the driven with the GluB-1 promoter and the GluB-1 signal peptide were accumulated in protein bodies (PBs), especially in glutelin-containing protein storage vacuoles (PSVs) in rice seeds [32, 41]. Another well-known PBs is the prolamin-containing PBs. Prolamins accumulate in PBs inside the rough ER without the typical ER-retrieval signal, whereas glutelins are transported from the cisternal ER through the Golgi to the PSVs. In PSVs, just as in glutelins, proteins are considered to be accumulated with the processing of macromolecular complexes formation. *Escherichia coli* heat-labile enterotoxin B (LTB) and immunoglobulin, for instance, required the multimeric protein formations and assembly for localization [42–44]. Thus, correct folding and assembly as multimeric proteins are indispensable for the protein subcellular destination and localization. The localization of CTB expressed in this study was not clear, but one possible hypothesis was considered: CTB was accumulated in PSVs through glutelin-like processing. As showed in Fig. 2e, CTB seemed to be pentameric with *N*-glycosylated and/or non-*N*-glycosylated CTBs in rice seeds. This suggests that CTBs were assembled before they reached PSVs, resulting in CTB

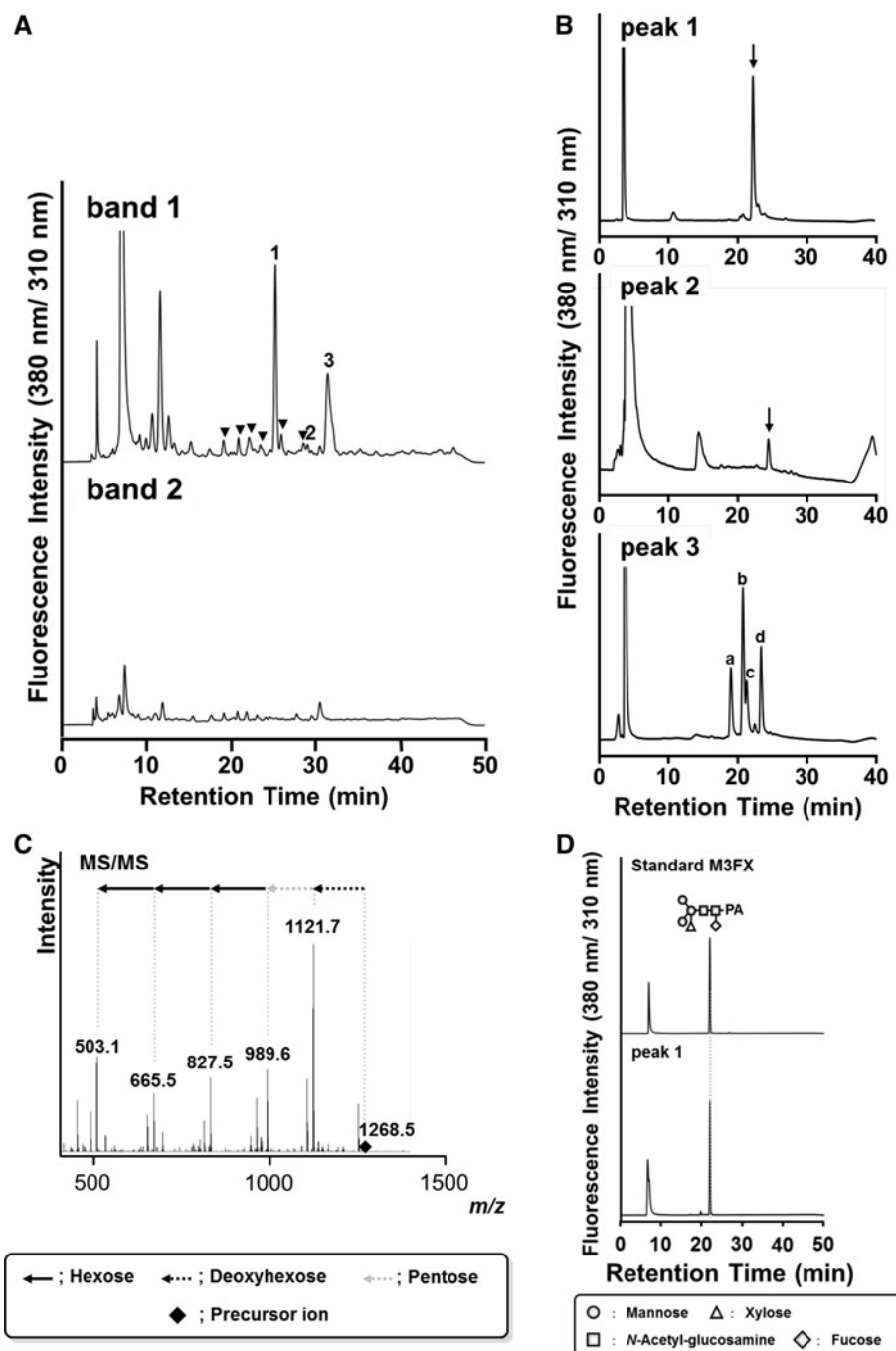


Fig. 5 *N*-Glycan analysis of CTB. *N*-Glycosylated CTB was excised from a gel and chymotrypsinized. The *N*-glycan was prepared by hydrazinolysis of the *N*-glycopeptide, followed by labeling with PA, and analyzed. **a** RP-HPLC analysis of the PA-sugar chain using a C_{18} column. **a** RP-HPLC analysis of the PA-sugar chain using a C_{18} column. *Band 1* was prepared from *N*-glycosylated CTB and *band 2* was prepared from non-*N*-glycosylated CTB. *Black triangles* indicate non-PA-sugar derivatives. **b** Secondary chromatographic separation of PA-derivatives from RP-HPLC using an amide column. *Peak 1*, *peak 2*, and four significant peaks in *peak 3* were collected, followed by further structural analysis by LC-MS/MS and comparison with authentic PA-sugar chains. *Arrows* indicate the collected peak, followed by LC-MS/MS analysis and comparison with authentic

PA-sugar chains using RP-HPLC. **c** LC-MS/MS analysis of the predominant PA-sugar chain from *peak 1*. The *labeled signal* represents $(M + H)^+$ ions. The value of m/z 503 detected in the MS/MS spectra agrees well with the calculated mass of HexNAc₂-PA, and the mass of the precursor ion, m/z 1268.5, was considered to correspond to Hex₃DeoxyHex₁Pentose₁HexNAc₂-PA. The *black arrow*, *black dotted arrow*, and *gray dotted arrow* indicate hexose, deoxyhexose, and pentose, respectively. A *black diamond* indicates the precursor ion of MS/MS fragmentation. **d** RP-HPLC profiles of Hex₃DeoxyHex₁Pentose₁HexNAc₂-PA. The PA-sugar chain from *peak 1* corresponded to an authentic sugar chain M3FX prepared from *A. thaliana*

Table 1 Relative amounts of *N*-glycan structures detected in CTB

Peak	Structure	Abbreviation	Ratio (%)
3-a		M3	9.1
3-b		M3X	18.2
1		M3FX	52.8
3-c		M4	6.5
3-d		M5	11.3
2		M6B	2.1

The relative ratio of structures was calculated on the basis of peak area of HPLC

transition through ER-Golgi processing to PSVs. In addition to this pentameric assembly, accumulation of CTB in PSVs was considered to be enhanced by the GluB-1 signal sequence. GluB-1 signal sequence mediates the ER targeting of the mRNA, which activates the interactions of synthesized protein with molecular chaperons in the ER, which results in the accumulation of protein same as glutelin. This hypothesis was further supported by the results of a previous study which showed that some of CTB was accumulated in PSVs [31] and by the *N*-glycan analysis in this study. The CTB expressed in rice seeds had M3FX as a predominant structure, which is synthesized via contributions of Golgi- and vacuole-localized enzymes. The formation of this M3FX, first XYLT and FUCTs transfer β 1,2-xylose and α 1,3-fucose residues to *N*-glycan, which results in the synthesis of GN2M3FX, and further maturation occurred by the subsequent trimming of two GlcNAc residues by vacuole-localized hexosaminidases [45]. Therefore, the pentameric complexes of CTB and the presence of M3FX *N*-glycan strongly suggested that CTB expressed in rice seeds was transported through a glutelin-like post-translational protein modification, such as multimeric protein assembly and ER-Golgi transition to PSVs. To examine this hypothesis, further investigations, such as inhibition of the assembly of pentameric CTB, CTB expression with a constitutive promoter

without GluB-1 signal sequences, and localization analysis are required and in progress.

Many recombinant proteins and CTB fusion proteins have been expressed and their *N*-glycosylations analyzed. Previously, Zhang et al. [46] showed that the recombinant human transferrin (rhTF) expressed in rice seeds was not *N*-glycosylated. It was of interest that rhTF was not *N*-glycosylated, though hTF potentially has two *N*-glycosylation sites which are, in general, *N*-glycosylated in the mammalian expression system. However, the absence of *N*-glycan on rhTF was not consistent with the result which showed that human lactoferrin (rhLF) expressed in rice seeds carried *N*-glycans [47, 48], though the amino acids sequence between rhTF and rhLF is quite similar: the amino acid identity and similarity were 59 and 85 %, respectively. Unfortunately, the reasons for this discrepancy are still unclear. On the other hand, Yang et al. [35] showed that a major house dust mite allergen, Der p 1, expressed in rice seeds under the same expression construct as that of CTB was *N*-glycosylated with the predominant *N*-glycan of high-mannose-type structure, but some of Der p 1 also was not *N*-glycosylated, just as seen in the CTB in this study. However, *N*-glycan structures and subcellular localization analysis demonstrated that Der p 1 was specifically localized in prolamin-containing PBs, suggesting that

the C-terminal KDEL sequence of Der p 1 successfully worked in rice seeds, though the GluB-1 signal sequence seemed not to be functional for the targeting of Der p 1 to PSVs. Interestingly, though the CTBs expressed in rice seeds seemed to be transported through the ER–Golgi machineries by escaping from the ER retention and finally accumulated in PSVs, some of them were not *N*-glycosylated (Fig. 2d). The results in this study were thus opposed to the results of Der p 1. What was the reason for this disparity? One key factor was the 3D structure and native forms of CTB. As mentioned above, a native CTB showed pentameric complexes. In this form, the CTBs seem to be an inaccessible to KDEL receptors, whereas the Der p 1 forms to be accessible to these receptors (data not shown). Der p 1 is monomeric in acidic solution, but forms a dimer in alkaline solution and in the crystal [49, 50], which indicates that the Der p 1 expressed in rice seeds is dimeric. Despite this dimeric solution, the KDEL signals are located opposite each other and are exposed to the surface. Therefore, Der p 1 was considered to be successfully binding to the KDEL receptor and retrieved to the ER. Each C-terminus of pentameric CTB, on the other hand, projects to same sides and make it difficult to access toward KDEL receptor. Therefore, CTB was speculated to have escaped from the ER-retrieve and to have *N*-glycan structures different from those of Der p 1. Furthermore, the pentameric forms of CTB and GluB-1 signal sequence also enhanced rather the ER–Golgi transit to PSVs by glutelin-like machinery than the ER retrieval mediated by KDEL signal, as mentioned above.

In this study, we focused on the expression of CTB in rice seeds and showed the post-translational modifications of CTB, including *N*-glycosylation, *N*-glycan analysis, and multimeric assembly. These results suggested that heterologous protein productions in plants, especially in rice seeds, are influenced by not only the endogenous signal sequence, promoter-specific expression, and the ER-retrieval sequence but also the multimeric assembly of proteins. These issues may be solved by the far-reaching choice of a target protein and its expression systems using a signal sequence, the expression of appropriate fusion proteins, and protein localization in plants.

References

- De Haan, L., & Hirst, T. R. (2004). Cholera toxin: A paradigm for multi-functional engagement of cellular mechanisms (review). *Molecular Membrane Biology*, 21, 77–92.
- Zhang, R. G., Scott, D. L., Westbrook, M. L., Nance, S., Spangler, B. D., Shipley, G. G., et al. (1995). The three dimensional crystal structure of cholera toxin. *Journal of Molecular Biology*, 251, 563–573.
- Zhang, R. G., Westbrook, M. L., Westbrook, E. M., Scott, D. L., Otwinowski, Z., Maulik, P. R., et al. (1995). The 2.4 Angstrom crystal structure of cholera toxin B subunit pentamer: Cholera-antigen. *Journal of Molecular Biology*, 251, 550–562.
- Sanchez, J., & Holmgren, J. (2011). Cholera toxin—a foe & a friend. *Indian Journal of Medical Research*, 133, 153–163.
- Soria-Guerra, R. E., Moreno-Fierros, L., & Rosales-Mendoza, S. (2011). Two decades of plant-based candidate vaccines: A review of the chimeric protein approaches. *Plant Cell Reports*, 30, 1367–1382.
- Chen, M., Liu, X., Wang, Z., Song, J., Qi, Q., & Wang, P. G. (2005). Modification of plant N-glycans processing: The future of producing therapeutic protein by transgenic plants. *Medicinal Research Reviews*, 25, 343–360.
- Hellwig, S., Drossard, J., Twyman, R. M., & Fischer, R. (2004). Plant cell cultures for the production of recombinant proteins. *Nature Biotechnology*, 22, 1415–1422.
- Horn, M. E., Woodard, S. L., & Howard, J. A. (2004). Plant molecular farming: Systems and products. *Plant Cell Reports*, 22, 711–720.
- Basaran, P., & Rodríguez-Cerezo, E. (2008). Plant molecular farming: Opportunities and challenges. *Critical Reviews in Biotechnology*, 28, 153–172.
- Daniell, H., Streatfield, S. J., & Wycoff, K. (2001). Medical molecular farming: Production of antibodies, biopharmaceuticals and edible vaccines in plants. *Trends in Plant Science*, 6, 219–226.
- Dove, A. (2002). Uncorking the biomanufacturing bottleneck. *Nature Biotechnology*, 20, 777–779.
- Daniell, H., Lee, S. B., Panchal, T., & Wiebe, P. O. (2001). Expression of the native cholera toxin B subunit gene and assembly as functional oligomers in transgenic tobacco chloroplasts. *Journal of Molecular Biology*, 311, 1001–1009.
- Matoba, N., Kajjura, H., Cherni, I., Doran, J. D., Bomsel, M., Fujiyama, K., et al. (2009). Biochemical and immunological characterization of the plant-derived candidate human immunodeficiency virus type 1 mucosal vaccine CTB-MPR_{649–684}. *Plant Biotechnology Journal*, 7, 129–145.
- Oszvald, M., Kang, T. J., Tomoskozi, S., Jenes, B., Kim, T. G., Cha, Y. S., et al. (2008). Expression of cholera toxin B subunit in transgenic rice endosperm. *Molecular Biotechnology*, 40, 261–268.
- Takagi, H., Hiroi, T., Yang, L., Takamura, K., Ishimitsu, R., Kawachi, H., et al. (2008). Efficient induction of oral tolerance by fusing cholera toxin B subunit with allergen-specific T-cell epitopes accumulated in rice seed. *Vaccine*, 26, 6027–6030.
- Arakawa, T., Chong, D. K., Merritt, J. L., & Langridge, W. H. (1997). Expression of cholera toxin B subunit oligomers in transgenic potato plants. *Transgenic Research*, 6, 403–413.
- Kim, T. G., & Langridge, W. H. (2003). Assembly of cholera toxin B subunit full-length rotavirus NSP4 fusion protein oligomers in transgenic potato. *Plant Cell Reports*, 21, 884–890.
- Choi, N. W., Estes, M. K., & Langridge, W. H. (2005). Synthesis and assembly of a cholera toxin B subunit-rotavirus VP7 fusion protein in transgenic potato. *Molecular Biotechnology*, 31, 193–202.
- Ruhlman, T., Ahangari, R., Devine, A., Samsam, M., & Daniell, H. (2007). Expression of cholera toxin B-proinsulin fusion protein in lettuce and tobacco chloroplasts—oral administration protects against development of insulinitis in non-obese diabetic mice. *Plant Biotechnology Journal*, 5, 495–510.
- Huy, N. X., Yang, M. S., & Kim, T. G. (2011). Expression of a cholera toxin B subunit-neutralizing epitope of the porcine epidemic diarrhea virus fusion gene in transgenic lettuce (*Lactuca sativa* L.). *Molecular Biotechnology*, 48, 201–209.
- Rayon, C., Lerouge, P., & Faye, L. (1998). The protein N-glycosylation in plant. *Journal of Experimental Botany*, 49, 1463–1472.
- Wilson, I. B. (2002). Glycosylation of proteins in plants and invertebrates. *Current Opinion in Structural Biology*, 12, 569–577.
- Brooks, S. A. (2004). Appropriate glycosylation of recombinant proteins for human use: Implications of choice of expression system. *Molecular Biotechnology*, 28, 241–255.

24. Tomiya, N., Narang, S., Lee, Y. C., & Betenbaugh, M. J. (2004). Comparing N-glycan processing in mammalian cell lines to native and engineered lepidopteran insect cell lines. *Glycoconjugate Journal*, *21*, 343–360.
25. Pattison, R. J., & Amtmann, A. (2009). N-glycan production in the endoplasmic reticulum of plants. *Trends in Plant Science*, *14*, 92–99.
26. Fitchette-Lainé, A. C., Gomord, V., Cabanes, M., Michalski, J. C., Saint, Macary, M., et al. (1997). N-Glycans harboring the Lewis a epitope are expressed at the surface of plant cells. *The Plant Journal*, *12*, 1411–1417.
27. Strasser, R., Bondili, J. S., Vavra, U., Schoberer, J., Svoboda, B., Glössl, J., et al. (2007). A unique β 1,3-galactosyltransferase is indispensable for the biosynthesis of N-glycans containing Lewis a structures in *Arabidopsis thaliana*. *The Plant Cell*, *19*, 2278–2292.
28. van Ree, R., Cabanes-Macheteau, M., Akkerdaas, J., Milazzo, J. P., Loutelier-Bourhis, C., Rayon, C., et al. (2000). β (1,2)-xylose and α (1,3)-fucose residues have a strong contribution in IgE binding to plant glycoallergens. *The Journal of Biological Chemistry*, *275*, 11451–11458.
29. Altmann, F. (2007). The role of protein glycosylation in allergy. *International Archives of Allergy and Immunology*, *142*, 99–115.
30. Jin, C., Altmann, F., Strasser, R., Mach, L., Schähls, M., Kunert, R., et al. (2008). A plant-derived human monoclonal antibody induces an anti-carbohydrate immune response in rabbits. *Glycobiology*, *18*, 235–241.
31. Nochi, T., Takagi, H., Yuki, Y., Yang, L., Masumura, T., Mejima, M., et al. (2007). Rice-based mucosal vaccine as a global strategy for cold-chain- and needle-free vaccination. *Proceedings of the National Academy of Sciences of the United States of America*, *104*, 10986–10991.
32. le Qu, Q., & Takaiwa, F. (2004). Evaluation of tissue specificity and expression strength of rice seed component gene promoters in transgenic rice. *Plant Biotechnology Journal*, *2*, 113–125.
33. Sugita, K., Kasahara, S., Ebinuma, H., Takaiwa, H., Jomori, T., Hayashi, Y., Tashita, A., & Kobara, Y. (2004). Process for producing plant storage organ with high production of recombinant protein and novel recombinant protein. Patent 371 WO 04/087910.
34. Sugita, K., Endo-Kasahara, S., Tada, Y., Lijun, Y., Yasuda, H., Hayashi, Y., et al. (2005). Genetically modified rice seeds accumulating GLP-1 analogue stimulate insulin secretion from a mouse pancreatic beta-cell line. *FEBS Letters*, *579*, 1085–1088.
35. Yang, L., Kajiura, H., Suzuki, K., Hirose, S., Fujiyama, K., & Takaiwa, F. (2008). Generation of a transgenic rice seed-based edible vaccine against house dust mite allergy. *Biochemical and Biophysical Research Communications*, *365*, 334–339.
36. Misaki, R., Kimura, Y., Fujiyama, K., & Seki, T. (2001). Glycoproteins secreted from suspension-cultured tobacco BY2 cells have distinct glycan structures from intracellular glycoproteins. *Bioscience, Biotechnology, and Biochemistry*, *65*, 2482–2488.
37. Kajiura, H., Seki, T., & Fujiyama, K. (2010). *Arabidopsis thaliana* ALG3 mutant synthesizes immature oligosaccharides in the ER and accumulates unique N-glycans. *Glycobiology*, *20*, 736–751.
38. Pagny, S., Bouissonnie, F., Sarkar, M., Follet-Gueye, M. L., Driouich, A., Schachter, H., et al. (2003). Structural requirements for *Arabidopsis* β 1,2-xylosyltransferase activity and targeting to the Golgi. *Plant Journal*, *33*, 189–203.
39. Bencúr, P., Steinkellner, H., Svoboda, B., Mucha, J., Strasser, R., Kolarich, D., et al. (2005). *Arabidopsis thaliana* β 1,2-xylosyltransferase: An unusual glycosyltransferase with the potential to act at multiple stages of the plant N-glycosylation pathway. *Biochemical Journal*, *388*, 515–525.
40. Kajiura, H., Okamoto, T., Misaki, R., Matsuura, Y., & Fujiyama, K. (2011). *Arabidopsis* β 1,2-xylosyltransferase: Substrate specificity and participation in the plant-specific N-glycosylation pathway. *Journal of Bioscience and Bioengineering*, *113*, 48–54.
41. Yang, D. C., Guo, F. L., Liu, B., Huang, N., & Watkins, S. C. (2003). Expression and localization of human lysozyme in the endosperm of transgenic rice. *Planta*, *216*, 597–603.
42. Chikwamba, R. K., Scott, M. P., Mejía, L. B., Mason, H. S., & Wang, K. (2003). Localization of a bacterial protein in starch granules of transgenic maize kernels. *Proceedings of the National Academy of Sciences of the United States of America*, *100*, 11127–11132.
43. Kim, T. G., Kim, B. G., Kim, M. Y., Choi, J. K., Jung, E. S., & Yang, M. S. (2010). Expression and immunogenicity of enterotoxigenic *Escherichia coli* heat-labile toxin B subunit in transgenic rice callus. *Molecular Biotechnology*, *44*, 14–21.
44. Nicholson, L., Gonzalez-Melendi, P., van Dolleweerd, C., Tuck, H., Perrin, Y., Ma, J. K., et al. (2005). A recombinant multimeric immunoglobulin expressed in rice shows assembly-dependent subcellular localization in endosperm cells. *Plant Biotechnology Journal*, *3*, 115–127.
45. Gomord, V., Fitchette, A. C., Menu-Bouaouiche, L., Saint-Jore-Dupas, C., Plasson, C., Michaud, D., et al. (2010). Plant-specific glycosylation patterns in the context of therapeutic protein production. *Plant Biotechnology Journal*, *8*, 564–587.
46. Zhang, D., Nandi, S., Bryan, P., Pettit, S., Nguyen, D., Santos, M. A., et al. (2010). Expression, purification, and characterization of recombinant human transferrin from rice (*Oryza sativa* L.). *Protein Expression and Purification*, *74*, 69–79.
47. Fujiyama, K., Sakai, Y., Misaki, R., Yanagihara, I., Honda, T., Anzai, H., et al. (2004). N-linked glycan structures of human lactoferrin produced by transgenic rice. *Bioscience, Biotechnology, and Biochemistry*, *68*, 2565–2570.
48. Nandi, S., Yalda, D., Lu, S., Nikolov, Z., Misaki, R., Fujiyama, K., et al. (2005). Process development and economic evaluation of recombinant human lactoferrin expressed in rice grain. *Transgenic Research*, *14*, 237–249.
49. de Halleux, S., Stura, E., VanderElst, L., Carlier, V., Jacquemin, M., & Saint-Remy, J. M. (2006). Three-dimensional structure and IgE-binding properties of mature fully active Der p 1, a clinically relevant major allergen. *Journal of Allergy and Clinical Immunology*, *117*, 571–576.
50. Chruszcz, M., Chapman, M. D., Vailes, L. D., Stura, E. A., Saint-Remy, J. M., Minor, W., et al. (2009). Crystal structures of mite allergens Der f 1 and Der p 1 reveal differences in surface-exposed residues that may influence antibody binding. *Journal of Molecular Biology*, *386*, 520–530.