

Transgenic Rice Seeds Accumulating Recombinant Hypoallergenic Birch Pollen Allergen Bet v 1 Generate Giant Protein Bodies

Shuyi Wang¹, Hideyuki Takahashi¹, Hiroyuki Kajiura², Taiji Kawakatsu¹, Kazuhito Fujiyama² and Fumio Takaiwa^{1,*}

¹Functional Transgenic Crops Research Unit, Genetically Modified Organism Research Center, National Institute of Agrobiological Sciences, Kannondai 2-1-2, Tsukuba, Ibaraki, 305-8602 Japan

²International Center for Biotechnology, Osaka University, 2-1 Yamada-oka, Suita, Osaka, 564-0871 Japan

*Corresponding author: E-mail, takaiwa@nias.affrc.go.jp; Fax, +81-29-838-8397.

(Received September 18, 2012; Accepted March 10, 2013)

A versatile hypoallergenic allergen derivative against multiple allergens is an ideal tolerogen for allergen-specific immunotherapy. Such a tolerogen should exhibit high efficacy, without side effects, when administered at high doses and should be applicable to several allergens. Tree pollen chimera 7 (TPC7), a hypoallergenic Bet v 1 tolerogen against birch pollen allergy, was previously selected by DNA shuffling of 14 types of Fagales tree pollen allergens. In this study, transgenic rice seed accumulating TPC7 was generated as an oral vaccine against birch pollen allergy by expressing this protein as a secretory protein using the N-terminal signal peptide and the C-terminal KDEL tag under the control of an endosperm-specific glutelin promoter. The highest level of TPC7 accumulation was approximately 207 $\mu\text{g grain}^{-1}$. Recombinant TPC7 is a glycoprotein with high mannose-type N-glycan, but without β 1,2-xylose or α 1,3-fucose, suggesting that TPC7 is retained in the endoplasmic reticulum (ER). TPC7 is deposited as a novel, giant spherical ER-derived protein body, $>20\ \mu\text{m}$ in diameter, which is referred to as the TPC7 body. Removal of the KDEL retention signal or mutation of a cysteine residue resulted in an alteration of TPC7 body morphology, and deletion of the signal peptide prevented the accumulation of TPC7 in rice seeds. Therefore, the novel TPC7 bodies may have formed aggregates within the ER lumen, primarily due to the intrinsic physicochemical properties of the protein.

Keywords: Bet v 1 • Birch pollen allergy • Hypoallergen • Protein body • Rice seed • Transgenic rice.

Abbreviations: allergen-SIT, allergen-specific immunotherapy; BiP, binding protein; CBB, Coomassie Brilliant Blue; DAF, days after flowering; ER, endoplasmic reticulum; GluB, rice glutelin B; IL-10, interleukin-10; PA, 2-aminopyridine; PB, protein body; PB-I/II, protein body type-I/II; PSV, protein storage vacuole; RT-PCR, reverse transcription-PCR; SSP,

seed storage protein; TEM, transmission electron microscopy; TPC, tree pollen chimera.

Introduction

Type 1 allergy represents a significant health problem in industrialized countries and affects approximately 15–22% of the European population (Canonica et al. 2007). Allergic symptoms, such as rhino-conjunctivitis, asthma and atopic dermatitis, are elicited by environmental aeroallergens such as pollen or house dust mites. Pollinosis caused by seasonal tree or grass pollen allergens is representative of IgE-mediated type 1 allergy disease. Birch pollen is widely distributed in Europe, North America, Russia and northern Japan. The major pollen allergens of Fagales trees belonging to the same order as birch are known as the Bet v 1 family because of the strong IgE cross-reactivity between Bet v 1, a major birch pollen allergen, and homologous allergens observed in alder (Aln g 1), hazelnut (Cor a 1), hornbeam (Car b 1) and oak (Que a 1) pollen (Breiteneder et al. 1989, Gajhede et al. 1996, Niederberger et al. 1998, Weber 2003, Radauer and Breiteneder 2006). Bet v 1 is a glycoprotein, with a molecular weight of approximately 20 kDa, that belongs to the PR10 protein group. More than 90% of birch pollen allergy patients are sensitized to Bet v 1-specific IgE. Furthermore, Bet v 1 allergy is also known to cause oral allergy syndrome, as the following food allergens are cross-reactive to Bet v 1-specific IgE: apple Mal d 1, soybean Gly m 4, carrot Dau c 1 and peanut Aha h 8, all of which belong to the same PR10 group in fruits and vegetables (Vieths et al. 2002).

Allergen-specific immunotherapy (SIT) is the only treatment of type 1 allergy that leads to long-lasting relief of clinical symptoms (Frew 2003, Larche et al. 2006). Allergen-SIT is performed by repeated administration of allergens by subcutaneous injection, or via mucosal routes to desensitize allergens.

Plant Cell Physiol. 54(6): 917–933 (2013) doi:10.1093/pcp/pct043, available online at www.pcp.oxfordjournals.org

© The Author 2013. Published by Oxford University Press on behalf of Japanese Society of Plant Physiologists.

All rights reserved. For permissions, please email: journals.permissions@oup.com

T cells are key regulators of allergy disease. Therefore, the induction of peripheral tolerance against T cells specific to the responsible allergen is the main target in allergen-SIT (Akids and Akids 2007). The main drawbacks of conventional allergen-SIT include systemic allergic side effects such as anaphylactic shock, as well as the long duration of treatment. To solve these problems, it would be highly desirable to develop safer, more effective and more convenient allergen-SIT. To achieve this goal, tolerogens used for desensitization to allergens must be changed from crude allergen extracts to recombinant hypoallergenic tolerogens with reduced IgE binding activity, without the loss of T-cell epitopes (Linhart and Valenta 2005, Valenta et al. 2010, Cromwell et al. 2011). To modify the three-dimensional structure required for recognition by the allergen-specific IgE and to reduce IgE binding capacity, fragmentation and oligomeric formation of Bet v 1 were achieved to generate hypoallergenic Bet v 1 derivatives, and the clinical efficacy of this tolerogen was examined (Vrtala et al. 2001, Niederberger et al. 2004).

The creation of versatile recombinant hypoallergenic allergens retaining all T-cell epitopes and immunogenicity should provide an ideal tolerogen for allergen-SIT. Hypoallergenic Bet v 1 derivatives against multiple Fagales pollen allergens were previously generated through in vitro random recombination by means of DNA shuffling. Importantly, not only did tree pollen chimera 7 (TPC7) and TPC9 exhibit lower allergenicity than native Bet v 1, as determined by testing IgE reactivity and using basophil activation assays, but this chimera also displayed higher immunogenicity than native Bet v 1 (Wallner et al. 2007). Therefore, this TPC7/TPC9 chimera should be a suitable tolerogen for specific immunotherapy not only against birch pollen allergy, but also against allergies caused by other cross-reactive tree pollen.

Transgenic plants are ideal platforms for the production of high-value recombinant proteins, such as pharmaceuticals, vaccines and antibodies (Daniell et al. 2009, Sharma and Sharma 2009). Production in cereal crop seed has already been demonstrated to give rise to a high yield of recombinant proteins based on the established expression system using a strong seed-specific promoter, codon optimization, and a specific host reducing storage proteins and depositing into protein bodies (PBs) (Stoger et al. 2005, Lau and Sun 2009, Kawakatsu and Takaiwa 2010). Furthermore, protein production in rice seeds has the additional benefit of high biomass yield with a minimal risk of outcrossing. In particular, when codon-optimized foreign recombinant protein genes designed for expression in rice seeds are produced in rice endosperm tissue, the proteins are accumulated safely and in large quantities without degradation, even if the proteins are not produced in other vegetative tissues such as leaves or roots, thus providing a superior production platform for recombinant proteins (Takaiwa et al. 2007). Furthermore, recombinant proteins produced in the seed are stable for several years at ambient temperatures without the loss of activity, thereby eliminating the need for cold-chain transportation and storage (Nochi et al. 2007).

In this study, a codon-optimized TPC7 synthetic gene, in which the rice glutelin B1 (GluB-1) signal peptide and the KDEL endoplasmic reticulum (ER) retention signal were fused to the N- and C- termini, respectively, was expressed in stable transgenic rice seeds under control of the endosperm-specific, 2.3 kb glutelin *GluB-1* promoter. TPC7 strongly accumulated in the endosperm tissue, leading to suppression of cysteine-poor 13 kDa prolamin and 26 kDa globulin production. Remarkably, when the intracellular location of TPC7 was examined, it was found to be deposited in an extraordinarily large organelle called the TPC7 body, which was distinct from PB-I and PB-II. The biochemical properties of TPC7 were characterized to examine the origin of this unique TPC7 body.

Results

Generation of transgenic rice seeds containing TPC7 and expression of TPC7 and SSPs in rice seeds

Based on the reported amino acid sequence of TPC7 (Supplementary Fig. S1), the DNA sequence coding for the TPC7 antigen was synthesized using codons that are preferentially used in many rice seed storage protein (SSP) genes. To maximize the production yield of TPC7 in rice seeds, the glutelin GluB-1 signal peptide and the KDEL ER retention signal were linked to the N- and C- termini of TPC7, respectively, and this gene construct was expressed under the control of the rice major SSP glutelin 2.3 kb *GluB-1* promoter (Fig. 1A). Accumulation levels of TPC7 in individual mature dry grains (at least four grains per independent transgenic line) were examined by Western blot analysis using an anti-TPC7 antibody. The line with the highest level of expression was advanced to the next generation by self-pollination, and the resulting homozygous seeds were further analyzed by SDS-PAGE. As shown in Fig. 1B, compared with wild-type seeds, TPC7 that accumulated in mature seeds was detected as additional bands on the SDS-polyacrylamide gel with molecular masses of 17 and 20 kDa. These bands specifically reacted with the anti-TPC7 antibody, as seen in the Western blot analysis (Fig. 1B). The estimated molecular mass of the lower band was consistent with the expected molecular mass calculated from the amino acid sequence of TPC7. The highest level of TPC7 accumulation was approximately $207 \mu\text{g grain}^{-1}$, accounting for 17% of total seed protein (Supplementary Fig. S2).

N-Glycan analysis of TPC7

TPC7 has a possible N-glycosylation site (Asn-Phe-Thr), between positions 79 and 81. To determine whether the upper band observed on the SDS-polyacrylamide gels was glycosylated TPC7, total seed proteins were digested by either endoglycosidase H (Endo H) or peptide-N-glycosidase F (PNGase F). As shown in Fig. 1C, the upper band, produced by either digestion, shifted to the 17 kDa region, at the same position as

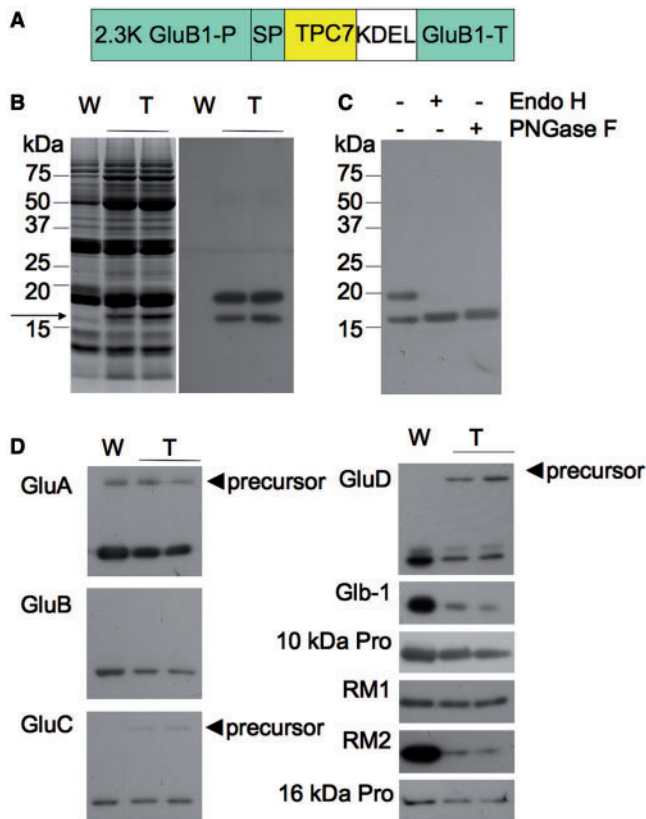


Fig. 1 Expression of TPC7 in rice seeds and analysis of SSPs in transgenic seeds. (A) Schematic representation of the gene construct used for expression of the TPC7 peptide in transgenic rice seeds. The TPC7 gene was under the control of the 2.3 kb glutelin B1 promoter (GluB1-P). SP and GluB1-T represent the glutelin B1 signal peptide sequence and 0.65 kb 3'-untranslated region, respectively. KDEL, tetrapeptide endoplasmic reticulum (ER) retention signal. (B) SDS-PAGE and Western blot analysis of TPC7 transgenic dry seeds. Total proteins were extracted from seeds of the wild type (W) and two independent homozygous TPC7 (T) transgenic lines, and separated by 12% SDS-PAGE. Coomassie Brilliant Blue (CBB) staining (left) and immunodetection with anti-TPC7 antibody (right). Molecular size markers are indicated on the left. The unglycosylated TPC7 polypeptide is indicated by an arrow. (C) N-Glycosylation analysis of TPC7 peptide accumulated in rice seeds. Total proteins from TPC7 seeds were digested in the absence (–) or presence (+) of either Endo H or PNGase F. Reaction mixtures were analyzed by Western blot analysis using anti-TPC7 antibody. (D) Western blot analysis of SSPs in TPC7 transgenic seeds. Immunodetection with anti-GluA, B, C, D, anti-Glob, anti-10 kDa, 13 kDa cysteine-rich (RM1), 13 kDa cysteine-poor (RM2) and 16 kDa prolamins.

the lower band, indicating that more than half of the expressed TPC7 was glycosylated with high mannose-type *N*-glycan.

To investigate the glycans linked to TPC7 in more detail, sugar chains were fractionated and characterized by both reversed phase- and size fractionation-HPLC, and a total of 12 fractions were obtained. The molecular masses of compounds in the fractions were analyzed by MALDI-TOF-MS (matrix-assisted laser desorption ionization-time of flight-mass

Table 1 Relative amounts of PA-glycan structures detected in TPC7

N-Glycan structure		Ratio (%)
Core mannose type	M3	1.4
	M4	1.3
	M5	1
	Total	3.7
High mannose type	M7A	3.2
	M7B	17.1
	M8A	47.3
	M8B	3.9
	M8C	1.1
	M9	13.9
	Total	86.5
α 1,3-glucose linked	GlcM7B	2.6
	GlcM8A	4.9
	GlcM9	2.3
	Total	9.8

The relative ratio of structures was calculated on the basis of the peak area in the size fractionation-HPLC profile. Peaks indicate the PA-glycans collected by reverse phase- and size fractionation-HPLCs.

spectrometry). The elution points of each glycan in the reversed phase-HPLC profile were compared with authentic 2-aminopyridine (PA)-glycans, and the glycan structures were determined (**Supplementary Fig. S3**). The *N*-glycans were exclusively of the high mannose type, with 7–9 mannose residues (**Table 1**). These high mannose-type glycans comprised 86.5% of the total glycans released from TPC7. Core mannose types with 3–5 mannose residues and the α 1,3-glucose-linked type were also detected as minor glycans, accounting for 3.7% and 9.8% of total glycans, respectively. The generation of core mannose-type glycan suggests processing by vacuole-localizing mannosidase. α 1,3-Glucose-linked glycosylation may be generated by re-glycosylation with UGGT (UDP-glucose:glycoprotein glucosyltransferase), after retrieval to the ER from the Golgi. High mannose glycans are added to accessible asparagines in proteins containing the Asn-X-Ser/Thr sequence when secretory proteins are translocated into the ER, and more complex glycans are built on the mannose framework in the Golgi apparatus. Thus, major modification with high mannose-type glycans and the absence of complex-type *N*-glycans, such as plant-specific α 1,3-fucose and β 1,2-xylose, suggest that most of the TPC7 protein may be retained in the ER and further deposited into ER-derived PBs or retrieved from the *cis*-Golgi by the retrieval machinery. However, there also remains the possibility that protein folding limits the access of Golgi enzymes to glycans.

Effect of TPC7 accumulation on seed storage proteins and chaperone proteins

In rice, the predominant SSPs are glutelins, globulins and prolamins. We examined whether the synthesis of TPC7 protein in maturing seeds had any effect on SSP production in transgenic seeds. Total seed protein was extracted from individual mature seeds of T₃ homozygous transgenic lines and subjected to Western blot analysis using specific antibodies, after

SDS-PAGE. As shown in **Fig. 1D**, 26 kDa globulin (Glb-1) and 13 kDa cysteine-poor prolamin (RM2) were markedly suppressed in transgenic rice grains compared with wild-type rice. However, the levels of most glutelins and 13 kDa cysteine-rich (RM1), 10 kDa and 16 kDa prolamin were similar to or lower than control levels.

Next, we examined whether the molecular chaperones involved in ER quality control were induced by the strong production of TPC7 in developing endosperm tissue. Specifically, the levels of binding proteins (BiPs; BiP1, BiP4&5), protein disulfide isomerases (1-1, 2-3) and calnexin in mature seeds of transgenic rice were compared with those of wild-type seeds. We found that the expression and accumulation of these chaperones significantly increased in mature seeds accumulating TPC7 (**Fig. 2**; **Supplementary Fig. S4**). These results indicate that high accumulation levels of TPC7 may be recognized as ER stress by BiPs (especially the induction of BiP4&5 as an ER stress marker), inducing the expression of chaperone proteins to refold highly expressed protein in the ER lumen (Oono et al. 2010, Wakasa et al. 2012).

TPC7 accumulation induces novel storage organelle formation in rice endosperm

Rice endosperm cells contain two types of PBs: ER-derived PBs, referred to as PB-Is, and protein storage vacuoles (PSVs), designated PB-IIs. Major rice SSP glutelins and 26 kDa globulin accumulate in PB-IIs, whereas prolamins are deposited in PB-Is (Tanaka et al. 1980, Krishnan et al. 1986). PB-I and PB-II are easily distinguishable due to differences in morphological features and electron density. The PB-Is containing prolamins are smooth and spherical, with a diameter of approximately 1–2 μm , with concentric rings of varying electron density.

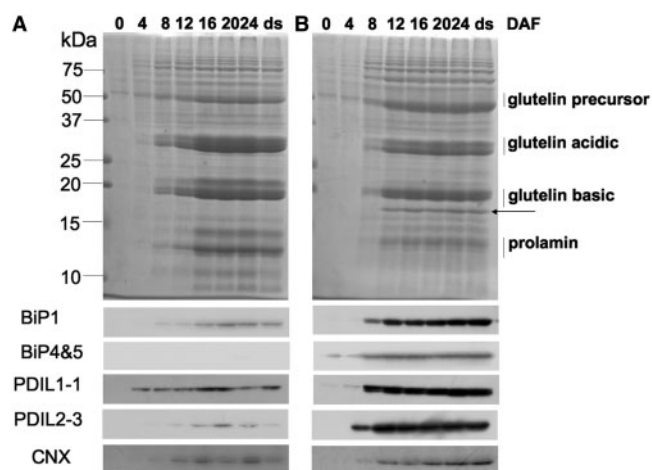


Fig. 2 Accumulation levels of chaperone proteins. Seed proteins extracted from developing (0–24 DAF) and mature (ds) seeds were separated by SDS-PAGE (top panels), and accumulation levels of chaperone proteins were analyzed by Western blot analysis (other panels). (A) Wild type. (B) TPC7 transgenic rice. The unglycosylated TPC7 band is indicated by an arrow.

In contrast, the PB-IIs containing glutelins and 26 kDa α -globulin are larger (3–4 μm) and irregularly shaped, with a highly uniform staining density. To investigate the effect of reduced SSP levels on PB formation by accumulation of TPC7 and the intracellular localization of TPC7 in subaleurone cells, the intracellular structures of endosperm cells in the rice subaleurone layer tissue were examined. First, thin sections (1 μm) of seeds at 15–20 days after flowering (DAF) from representative homozygous lines were stained with toluidine blue and observed by optical microscopy. Interestingly, extraordinarily large organelles were newly observed in transgenic rice seeds accumulating TPC7, compared with wild-type seeds (**Fig. 3A, B**). This novel organelle was spherical, with a diameter of approximately 5–20 μm (**Fig. 3B**). To observe this novel organelle in detail, 80 nm sections were observed by transmission electron microscopy (TEM). In transgenic rice seeds, the size and number of both PB-Is and PB-IIs drastically diminished (**Fig. 3C, D**), supporting the notion that levels of some SSPs were remarkably reduced. The novel organelle had a smooth, round shape and an electron-lucent structure (**Fig. 3D**). The electron density of this organelle was lower than that of PB-Is. Next, immune-cytochemical studies were carried out in developing endosperm, at 18 DAF, using several antibodies. When an anti-RM1 antibody was used, immuno-gold particles were only detected in PB-Is, but not in the novel organelle (**Fig. 3E, F**). However, PB-Is were rarely incorporated into the novel organelle (**Fig. 3F**). Furthermore, the smooth surface of PB-Is observed in non-transgenic rice became wavy or rugged, with radial stripes, although the spherical shape and electron-lucent characteristic of PB-Is in the wild type was retained (**Fig. 3E, F**). When the intracellular localization of GluA was examined, gold particles labeled with the anti-GluA antibody were observed in PB-IIs but not in the novel organelle (**Fig. 3H**). These results indicate that this novel organelle, detected in transgenic rice seeds, is a unique organelle that is different from PB-Is and PB-IIs. In transgenic rice seeds, gold particles labeled with an anti-TPC7 antibody were specifically localized in the novel organelle (**Fig. 3G**), which we designated the TPC7 body.

Next, to reconfirm the novelty of organelle-accumulating TPC7, thick sections (150 μm) of developing seeds, at 15–20 DAF, were stained with rhodamine B to localize ER-derived PB-Is (red), and immunostained with anti-TPC7 (green) antibody to visualize the TPC7 body. The TPC7 signal was detected within the extraordinarily large organelle in transgenic endosperm cells, but not in wild-type seeds (**Fig. 4A, B**), confirming the generation of the novel TPC7 body. In addition, the number and size of PB-Is stained with rhodamine were clearly smaller than those of non-transgenic rice as a result of TPC7 body formation.

BiP is the key chaperone involved in the folding of secretory proteins, such as SSPs, in the ER lumen. Thus, major BiP1 interacts with nascent immature secretory protein synthesized from membrane-bound polysomes on the ER. The recombinant TPC7 had a KDEL ER retention signal at the C-terminus and

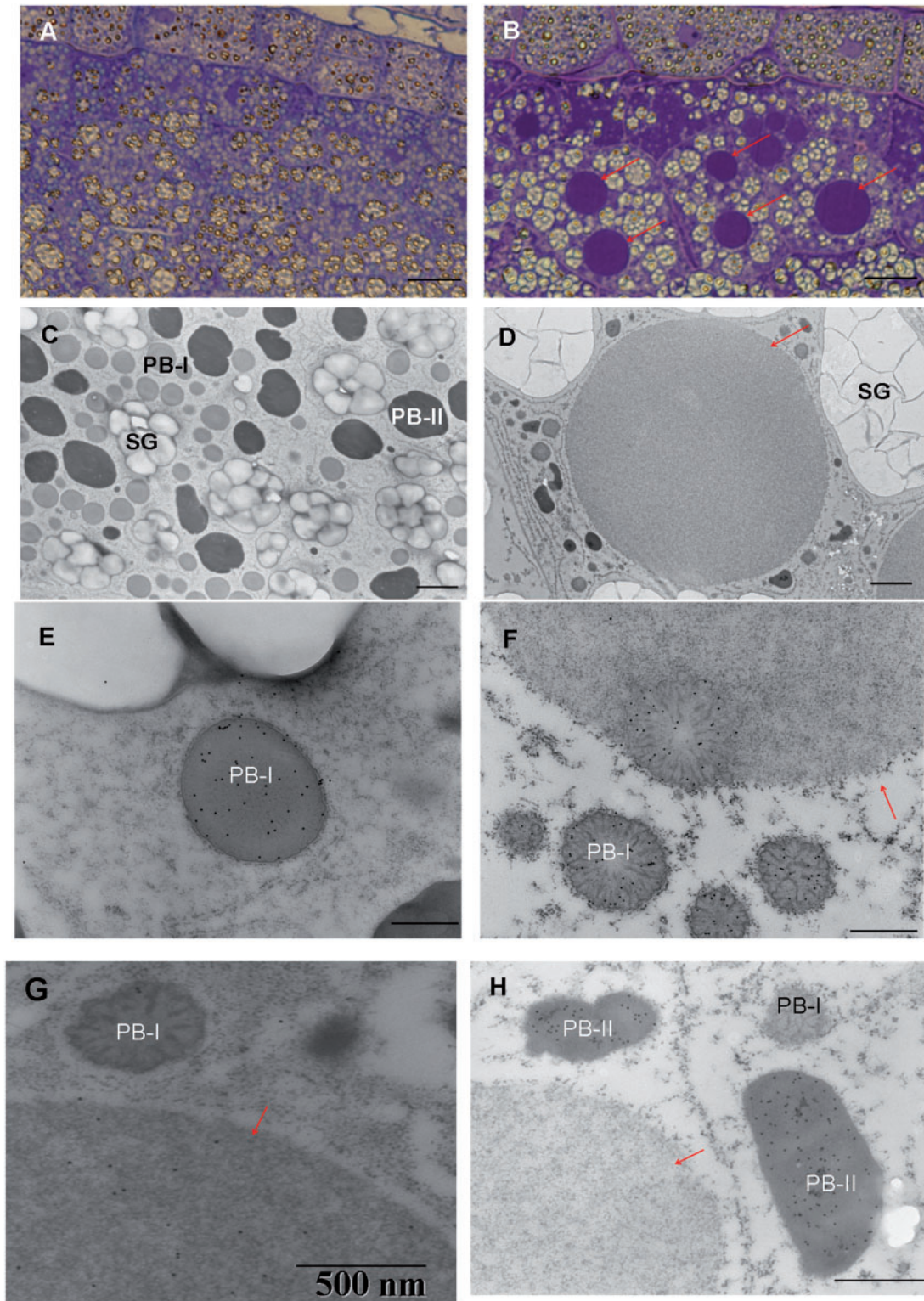


Fig. 3 Optical microscopy and transmission electron microscopy (TEM) of a subaleurone layer cell in developing seeds of wild-type and TPC7 transgenic plants. (A, C, E) Wild type. (B, D, F–H) TPC7 transformant. (A, B) Sections stained with toluidine blue and observed by optical microscopy. Bars = 20 μ m. (C, D) Sections were observed by TEM. Bars = 2 μ m. (E–H) Immunoelectron micrographs. (E and F) Anti-RM1 antibody. (G) Anti-TPC7 antibody. (H) Anti-GluA. Bars = 500 nm. Red arrows indicate the novel organelle the TPC7 body. SG, starch granule; PB-I, protein body I; PB-II, protein body II.

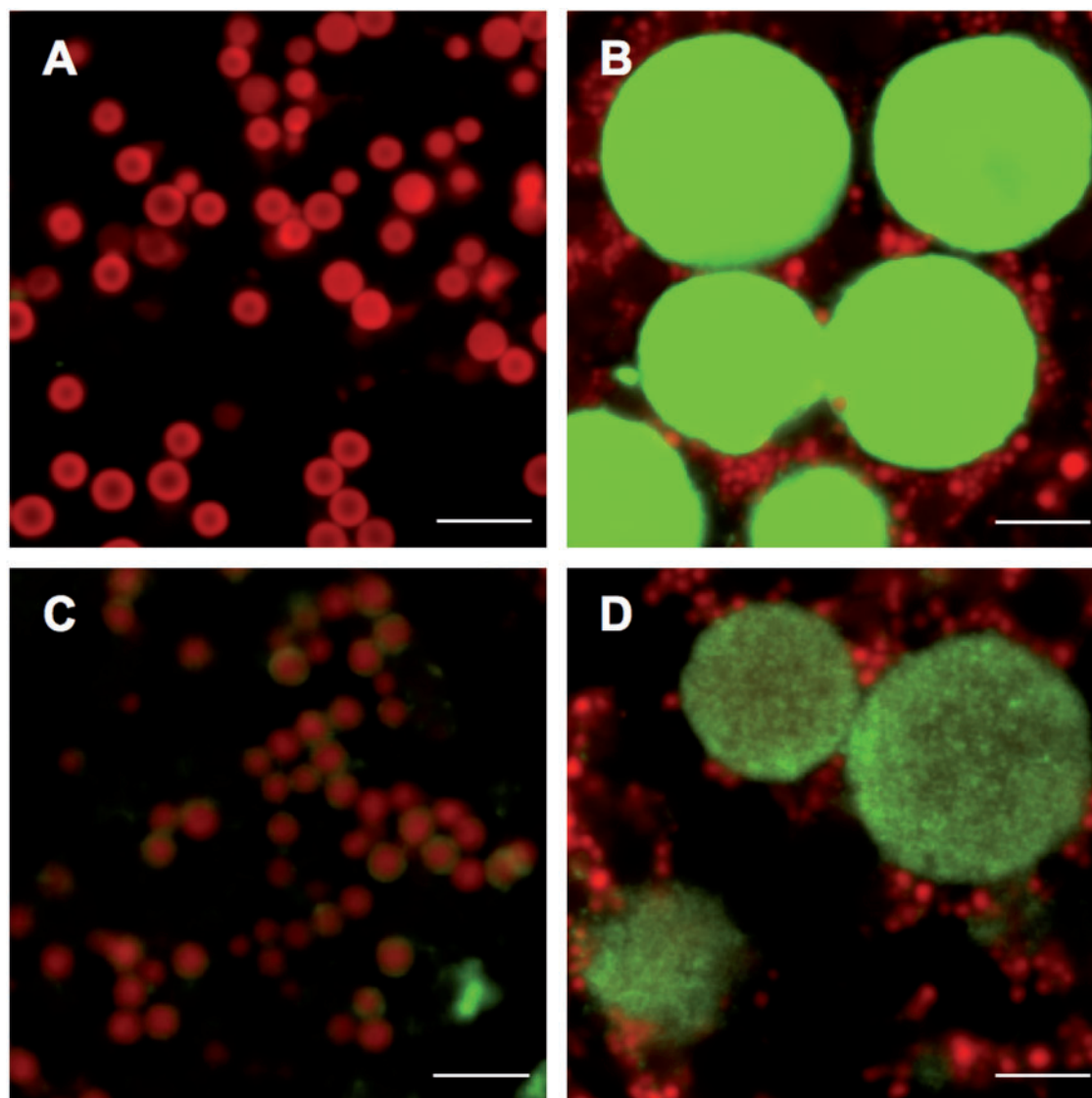


Fig. 4 Indirect immunohistochemical analysis of TPC7 in subaleurone layer cells in developing seeds of the wild type (A, C) and TPC7 transformant (B, D). (A, B) Red signal indicates PB-Is labeled with rhodamine, and green signal indicates immunofluorescence of TPC7 immunostained with anti-TPC7 antibody. (C, D) Red signal indicates PB-Is labeled with rhodamine, and green signal indicates immunofluorescence of BiP-1 immunostained with anti-BiP antibody. Bars = 5 μ m.

was expected to be retained in the ER lumen. BiP1 was found to be predominantly localized in the periphery of PB-Is in the wild type (**Fig. 4C**). In transgenic rice, BiP1 localized not only at the periphery of PB-Is, but also more strongly in the TPC7 body (**Fig. 4D**). In light of the finding that BiP1 was localized to the TPC7 body (**Fig. 4D**), the TPC7 body may be derived from the ER.

Characterization of TPC7 protein

During rice seed development, endosperm cells synthesize large amounts of proteins, salt-soluble globulins, alcohol-soluble prolamins, and acid- and alkaline-soluble glutelins. A sequential extraction experiment was carried out to examine the solubility of TPC7 protein. Rice seed proteins were

sequentially extracted with a salt solution (globulin fraction), an alcohol solution containing reducing agent (prolamin fraction) and an acid solution (glutelin fraction). As shown in **Fig. 5A**, TPC7 expressed in seeds could be extracted in the acid- and alkaline-soluble glutelin fractions, but not in the prolamin fraction. The removal of cysteine-rich prolamins was not necessary for the extraction of TPC7 because TPC7 could be efficiently extracted with 1% lactic acid without extraction of prolamins (**Fig. 5B**). It was notable that TPC7 was not eluted in the glutelin fraction before extraction with salt solution (**Fig. 5C**). Removal of the globulin fraction was required as a pre-treatment for the extraction of TPC7. These results suggest that aggregates in the TPC7 body also include salt-soluble endogenous rice proteins.

PB-Is and PB-IIs can easily be distinguished by differences in the sensitivity of these structures to pepsin digestion (Ogawa et al. 1987). Prolamin in PB-Is is resistant to pepsin digestion, whereas glutelin in PB-IIs is readily digested by this protease in wild-type rice (Takagi et al. 2010). To examine the physical properties of the TPC7 body, we performed *in vitro* digestion assays. As shown in Fig. 5D, pepsin digested GluA and TPC7 completely within 30 min, whereas RM1 cysteine-rich prolamin was partially resistant even after digestion

for 120 min with pepsin. These results indicate that the digestibility of TPC7 is different from that of prolamins deposited in PB-Is, although the TPC7 body and PB-Is are derived from the ER in a similar manner. The lower resistance of TPC7 vs. prolamin in PB-Is to gastrointestinal digestion may be explained by the fact that the TPC7 body is predominantly composed of TPC7 alone, whereas PB-Is are tightly packaged with several different types of prolamins, such as cysteine-rich and cysteine-poor 13 kDa prolamins, 16 kDa and 10 kDa prolamins, that

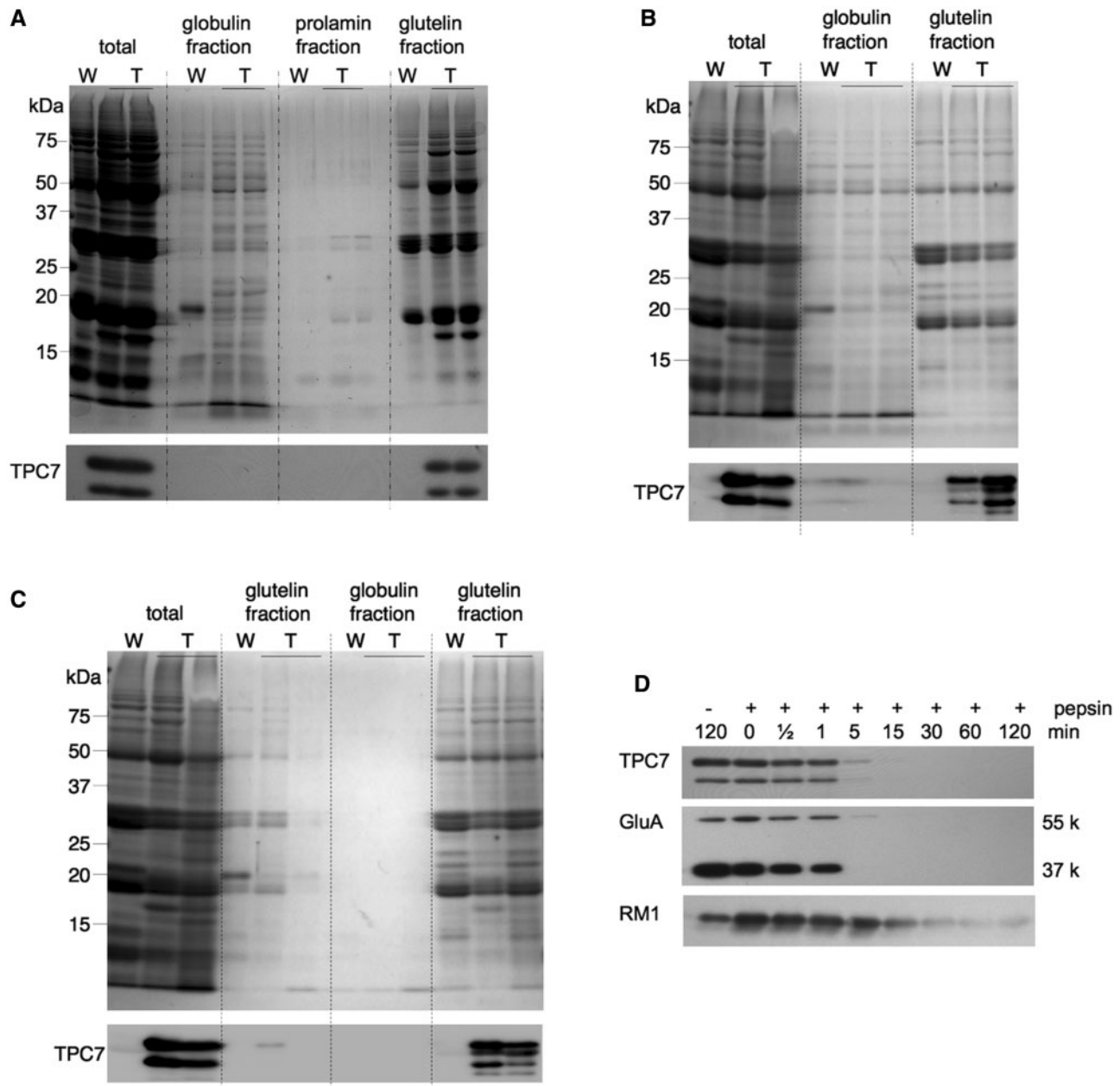


Fig. 5 Characterization of TPC7 protein. (A–C) Sequential extraction of seed proteins from the wild type (W) and TPC7 transformant (T). Each extracted protein fraction was subjected to SDS–PAGE (upper) and Western blot analysis using anti-TPC7 antibody (lower). (D) *In vitro* digestion of transgenic rice seeds expressing TPC7. A homogenate of transgenic rice seeds was mixed with pepsin. These reaction mixtures were incubated for up to 120 min at 37°C and examined by Western blot analysis using anti-TPC7, GluA and RM1 antibodies.

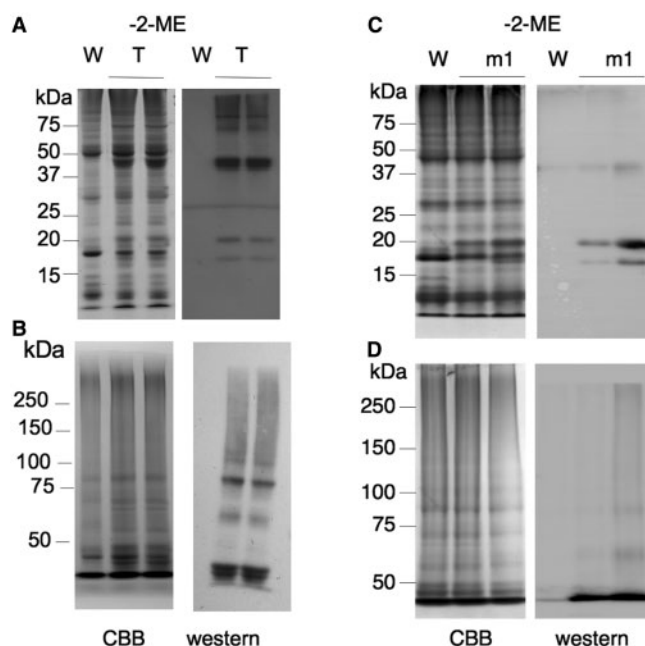


Fig. 6 TPC7 protein forms high aggregates via disulfide bonds. Proteins were extracted with buffer in the absence (–) of 5% 2-mercaptoethanol from mature seeds of the wild type (W), TPC7 (T) and TPC m1 (m1) and then subjected to 12% (A, C) and 6% (B, D) SDS–PAGE under non-reducing conditions (in the absence of 5% 2-mercaptoethanol). Analysis was performed using CBB staining (left) and Western blot analysis using anti-TPC7 antibody (right).

are spatially and systematically arranged during seed maturation (Nagamine et al. 2011).

To investigate the structural status of TPC7 deposited in the TPC7 body, TPC7 was extracted from mature seeds with buffer lacking 2-mercaptoethanol. When subjected to 12% and 6% PAGE and analyzed by Western blot analysis, the major 17 and 20 kDa bands that were observed after 2-mercaptoethanol extraction drastically decreased. Instead, very broad smear bands with molecular masses of >120 kDa were detected, along with major bands of approximately 45 kDa and a band of approximately 90 kDa (Fig. 6A, B). The latter bands may have been generated by dimer formation or large aggregate formation via inter-disulfide bonds and hydrophobic interactions characteristic of Bet v 1 protein; a trimer of Bet v 1 was previously shown (by gel filtration analysis) to form high molecular weight aggregates (>600 kDa) (Campana et al. 2011). The propensity of TPC7 to form large aggregates may be implicated in the formation of TPC7 bodies.

Signals implicated in TPC7 body formation

We have previously reported that when some recombinant proteins are expressed as secretory proteins by adding an N-terminal signal peptide to transport the proteins into the ER lumen, the proteins accumulated to higher levels than proteins lacking a signal peptide (Takagi et al. 2005b). The ligation of the KDEL ER retention signal to the C-terminus of

recombinant proteins is also an effective way to obtain higher levels of accumulation than recombinant proteins lacking a KDEL (Wandelt et al. 1992). Furthermore, TPC7 contains one free cysteine residue at position 113. This cysteine residue may be required for high levels of protein accumulation by the inter-molecular interaction between free cysteine residues via disulfide bonds. Therefore, to investigate the roles of signal peptide, the KDEL ER retention signal and free cysteine residues in the accumulation of TPC7 and the generation of TPC7 bodies in rice seeds, we transformed rice plants using five distinct constructs (Fig. 7A). In the m1 construct, the cysteine residue was changed to alanine. In the m2 construct, the KDEL ER retention signal was removed. In the m3 construct, the cysteine residue was changed to alanine and the KDEL ER retention signal was removed. In the m4 and m5 constructs, both the GluB1 signal peptide and the KDEL ER retention signal were removed. In addition, the cysteine residue was replaced by alanine in construct m5.

The expression levels of the *TPC7* gene in maturing seeds were first analyzed by reverse transcription–PCR (RT–PCR) in the transgenic plants. The accumulation levels of TPC7 in mature, dry grains were then examined by Western blot analysis with an anti-TPC7 antibody. Although the *TPC7* transcript levels were not affected by the removal of either the KDEL retention signal or the N-terminal signal peptide, the lack of both signals resulted in a remarkable reduction in the *TPC7* transcript level in transgenic rice containing m4 or m5 (Fig. 7B). In fact, TPC7 accumulation was not detectable in transgenic rice containing either m4 or m5 (Fig. 7C). These results indicate that the glutelin signal sequence might be required for the stable expression and accumulation of TPC7 in rice seeds.

The morphology of the TPC7 body in transgenic rice containing construct m1 or m2 was slightly distorted from that of the native body, and the size of the TPC7 body decreased (Figs. 4B, 7D, E). The surfaces and structures of the TPC7 body were not smooth or spherical in transgenic rice containing either m1 or m2 (Fig. 7G, H). In transgenic rice containing construct m3, the TPC7 body was rarely detected. The TPC7 bodies, when present, were spherical, with non-smooth surfaces (Fig. 7J). Most of the TPC bodies were <0.2 µm in size (Fig. 7I). Remarkably, they were located in the nucleus. This was never observed in wild-type rice or transgenic rice containing other constructs for TPC7 expression (Fig. 7F, I–L). We have no clear explanation for why the double mutations in the m3 construct favor nuclear localization, although conversion of cysteine to alanine and removal of KDEL from TPC7 resulted in nuclear localization in the transgenic rice. The presence or absence of glycosylation may be related to the localization at the nucleus. However, the ratio of glycosylated TPC7 to unglycosylated TPC7 was not appreciably affected by the different mutations (Fig. 7C). To examine whether the glycosylated and unglycosylated forms of TPC7 were co-localized in TPC7 bodies, we tried to carry out subcellular fractionation. TPC7 bodies were collected by centrifugation and then subjected to Western blot

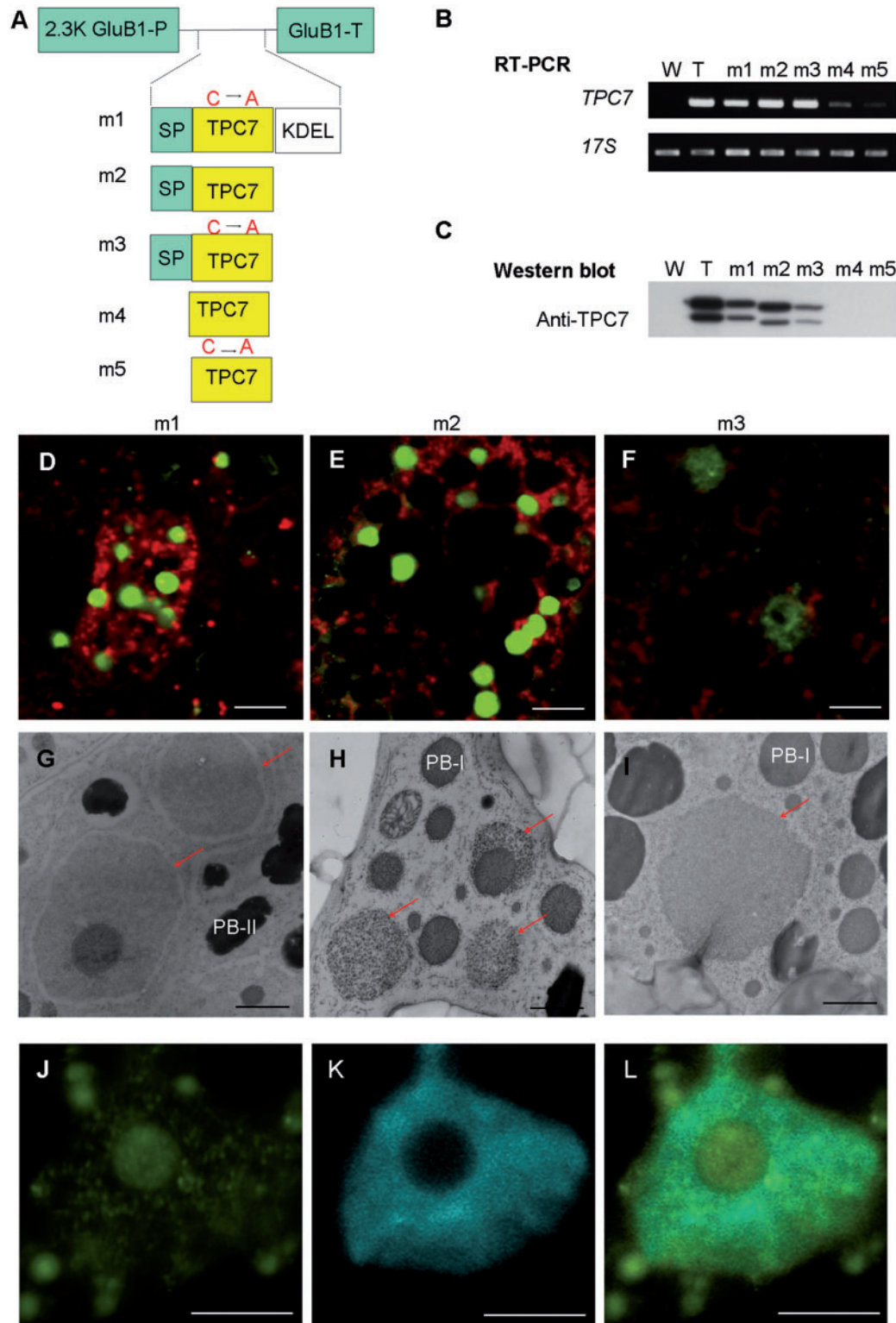


Fig. 7 Comparison of TPC7 peptide accumulation and TPC7 body localization using different constructs in rice seeds. (A) Schematic representation of five constructs compared with the TPC7 construct. C→A indicates cysteine changed to alanine in the TPC7 amino acid sequence. (B) Expression level of *TPC7* genes in developing seeds. Expression of the *17S* gene was evaluated as an internal control for RT-PCR. (C) Accumulation level of TPC7. (D–F) Merged image for indirect immunohistochemical analysis of the TPC7 body in subaleurone cells of m1, m2 and m3 transgenic rice seeds. Green and red signals indicate the localization of TPC7 and PBs, respectively. Bars = 10 μm. (G–I) Sections of m1, m2 and m3 were observed by TEM. Red arrows indicate the TPC7 body. Bars = 1 μm. (J–L) Immunostaining of TPC7 (green) in transgenic rice containing the m3 construct. DAPI (4',6-diamidino-2-phenylindole)-stained (blue) nucleus. Bars = 5 μm.

analysis using an anti-TPC7 antibody. As shown in **Supplementary Fig. S5**, both the glycosylated and unglycosylated forms of TPC7 were detected in the same fraction. These results suggest that the glycosylated and unglycosylated forms of TPC7 were co-localized in TPC7 bodies. Furthermore, the TPC7 gene containing alanine instead of asparagine (the N-glycosylation site) at position 79 was introduced into the rice genome. **Supplementary Fig. S6** shows that distinct TPC7 bodies similar to those of native TPC7 bodies were present in the transgenic rice seed. These findings indicate that there was little correlation between glycosylation and trafficking to the nucleus. Further work will be required to understand why double mutations in m3 direct nuclear localization.

Role of the cysteine residue in the TPC7

When the m1 mutant lacking the single cysteine residue was analyzed by SDS-PAGE under non-reducing conditions, the amount of dimers or multimers formed via interdisulfide bonds was significantly lower than with native TPC7 protein (**Fig. 6C, D**). This result indicates that disulfide bond formation was involved in the formation of the high molecular weight aggregates. Multimers larger than dimers could also have been the result of interchain disulfide bond formation between TPC7 and endogenous rice proteins, resulting in heteromer formation *in vivo* or during the extraction procedure, because native TPC7 has a free cysteine residue. However, it is notable that TPC7 bodies made up of TPC7 multimers were detected even in transgenic rice seed containing the m1 construct, although the size of such bodies was lower than in rice seeds containing wild-type TPC7. These results suggest that self-aggregation via hydrophobic interactions may also participate in TPC7 body formation.

We first examined whether m1 protein could be extracted by using neutral buffer containing a non-ionic detergent (1% Triton X-100) to solubilize the membrane. As shown in **Supplementary Fig. S7**, mutagenized TPC7 lacking the cysteine residue could not be extracted by neutral buffer containing the non-ionic detergent. On the other hand, extraction with neutral buffer and saline solution containing 0.15% SDS resulted in the extraction of TPC7 as well as m1 (**Fig. 8**). Extraction with saline solution required a higher concentration of SDS than extraction without salt, both for TPC7 and for m1. This may be due to a decrease in stability of SDS by ion exchange between SDS and saline. In light of the results showing that glutelins and prolamins cannot be extracted with these solutions, high molecular weight aggregates of TPC derivatives deposited in TPC7 bodies may be formed via hydrophobic interactions in a manner similar to that of Bet v 1.

Furthermore, the interaction of TPC7 protein with endogenous seed proteins via the cysteine residue was examined by comparing the sequential extraction patterns of TPC7 seeds with those of m1 seed. **Supplementary Fig. S8** shows there was no difference in the sequential extraction pattern between TPC7 seeds and m1 seeds. Although we would not exclude totally the possibility that heterotypic interactions are also

responsible for the formation of the large polymers, these results suggest that self-aggregation of TPC7 via disulfide bonds and hydrophobic interactions mainly contributes to the formation of large TPC7 bodies.

Discussion

Plants offer an attractive platform for the development of affordable vaccines (Tiware et al. 2009). The advantages of plants include the low cost of production, scalability, low health risks due to little or no contamination with human pathogens, and the potential for the vaccines to be used in the unprocessed or partially processed form. In particular, plant-based vaccines provide a means to deliver large quantities of a designated antigen in an encapsulated form. Notably, when transgenic plants containing tolerogens are orally administered, not only systemic (parental), but also mucosal immune tolerance can be induced. Furthermore, the bio-encapsulation of tolerogens in plant tissue protects the material from digestive enzymes and harsh conditions in the gastrointestinal tract, thus providing an ideal system for delivery to the mucosal immune system in gut-associated lymphoid tissue (GALT), due to the presence of cell compartments, consisting of cell walls and protein bodies, when antigens are orally administered (Streatfield 2006, Takaiwa 2011). Stability at ambient temperatures for several years and needle-free administration also make oral immunotherapy an attractive alternative to subcutaneous injection (Nochi et al. 2007).

The production of Bet v 1-related antigens in plants has not been reported, although recombinant Bet v 1 and hypoallergenic Bet v 1 derivatives, such as Bet v 1 trimer and Bet v 1 fragments, have been produced in *Escherichia coli* as tolerogens for allergen-specific immunotherapy (Gronlund and Gafvelin 2010). These recombinant Bet v 1 derivatives have been subjected to clinical trials and have proceeded to Phase III.

In this study, the TPC7 hypoallergenic tolerogen was highly and specifically expressed in the endosperm tissue of stably transgenic rice seeds. The highest level of TPC7 accumulation was approximately 207 $\mu\text{g grain}^{-1}$, accounting for 17% of total seed protein. Interestingly, TPC7 was deposited in extraordinarily large, probably ER-derived bodies that we designated TPC7 bodies. The morphology of the TPC7 body is spherical, and its surface is smooth. The diameter of the TPC7 body sometimes reaches $>20 \mu\text{m}$, with an average diameter of 5–10 μm . More than half of the accumulated TPC7 was found to be glycosylated, with one N-glycosylation site (Asn-X-Ser/Thr) between positions 79 and 81. Glycosylation was mainly limited to high mannose-type N-glycans, with a lack of complex-type N-glycans containing plant-specific α 1,3-fucose and β 1,2-xylose, suggesting that TPC7 may be completely retrieved to the ER from the Golgi and may be not trafficked to the Golgi apparatus. As a result, TPC7 efficiently accumulates in the ER lumen, and TPC7 bodies may be generated by specific aggregation due to hydrophobic properties. The ER origin of the TPC7 body was also

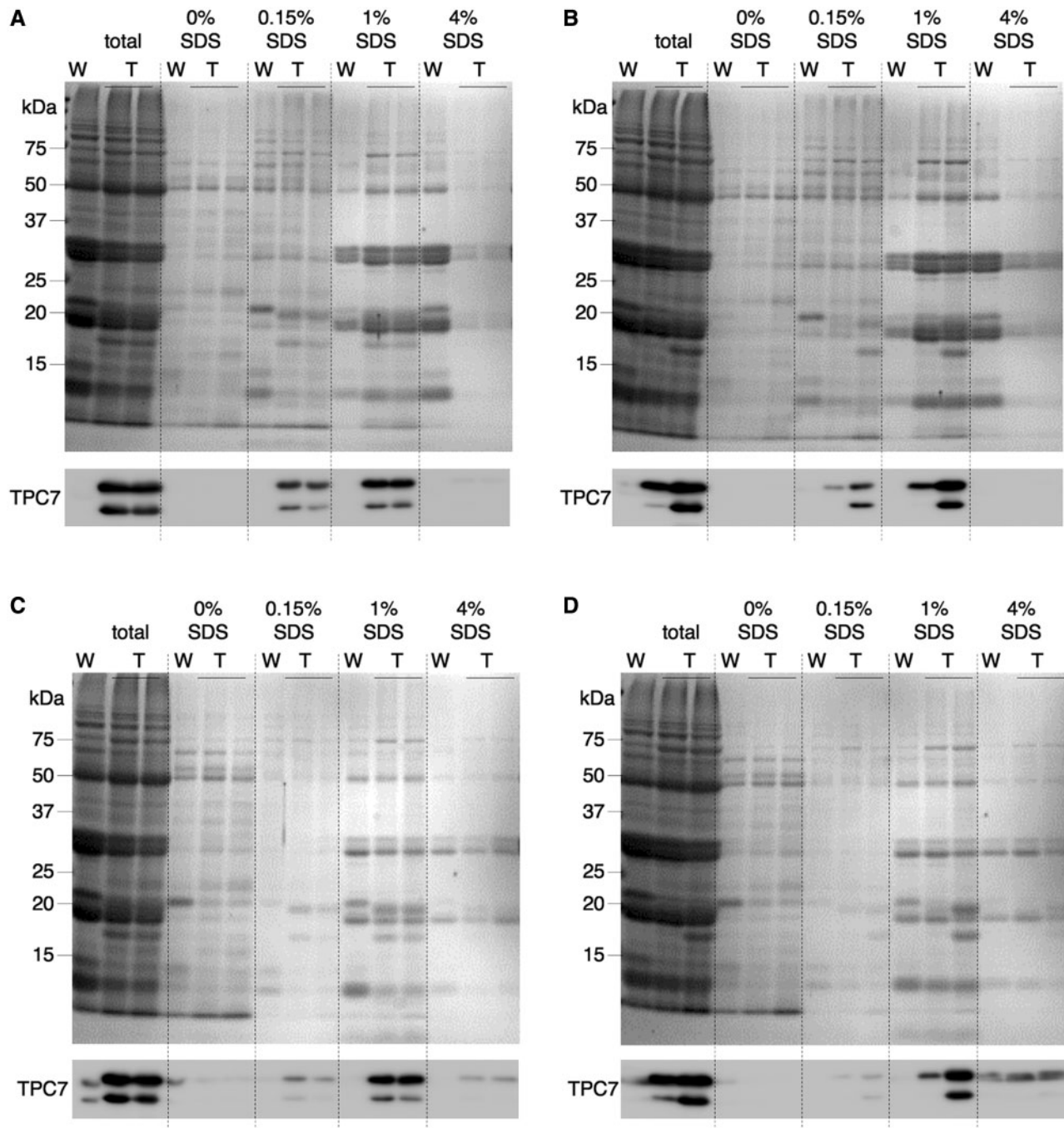


Fig. 8 Solubility of TPC7 and TPC7m1 proteins in neutral buffer containing different concentrations of SDS. Sequential extraction of TPC7 (A and C) and TPC7m1 (B and D) proteins from transgenic seeds using neutral buffer (10 mM Tris-HCl, pH 7.5) containing 0, 0.15, 1 and 4% SDS (A and B) and saline buffer (10 mM Tris-HCl, pH 6.8, 0.5 M NaCl) containing 0, 0.15, 1 and 4% SDS (C and D). Each extracted protein fraction was subjected to SDS-PAGE (upper) and Western blot analysis using anti-TPC7 antibody (lower).

supported by the finding that the major BiP1 characteristic of ER-derived PB-Is was localized to the TPC7 body.

When recombinant proteins are produced in seeds as secretory proteins via the endomembrane system, the intracellular deposition site of the recombinant protein is closely related to production yield, which is determined by targeting signals,

inherent physical properties and interaction with other proteins in the ER lumen. In particular, stability and post-translational modification such as glycosylation are affected by intracellular deposition sites and trafficking routes. When KDEL-tagged recombinant proteins are produced in seeds, production yields are often enhanced, and most of the

recombinant proteins accumulate primarily, or exclusively, in ER-derived PBs. However, the subcellular localization of recombinant proteins sometimes depends on the intrinsic properties and the assembly state of the protein (Stoger et al. 2005). Notably, the deposition of recombinant protein is affected by highly synthesized storage proteins. Free cysteine residues in the recombinant protein may be involved in the interaction of the protein with cysteine-rich SSPs such as cysteine-rich prolamins, resulting in the deposition of the protein into ER-derived PBs by co-localization with prolamins (Kawagoe et al. 2005). On the other hand, some recombinant proteins are prone to form aggregates in the ER lumen due to physical properties, such as high hydrophobic domains within a molecule. Generally, aggregated proteins are sensed as ER stress and are therefore degraded by the unfolded protein response that serves as an ER quality control mechanism. However, it is important to note that the products of high levels of aggregation in the ER lumen can accumulate without degradation when expressed in cereal endosperm tissue. Of course, ER stress is activated when the ER is overloaded with recombinant proteins. As shown in Fig. 2, the expression of high levels of TPC7 in rice endosperm tissue elicited ER stress, leading to the distortion of native storage protein trafficking and PB formation. In particular, the synthesis of prolamins that are sorted into ER-derived PB-Is is severely suppressed, compared with that of glutelins deposited in PSVs (PB-IIs). When a large TPC7 body was generated in the ER lumen and budded off, the morphology of the PBs (especially PB-Is) was distorted, and the number and sizes of PB-Is and PB-IIs were remarkably reduced.

Despite bearing a KDEL retention signal, the phenomena in which recombinant proteins are targeted to other locations,

such as PSVs and the apoplast, have been reported in several studies of proteins expressed in seeds (Table 2). Human serum albumin and *Aspergillus niger* phytase were delivered to PSVs, along with SSPs, in wheat endosperm. Several ScFVs (scFvT84.66, HA78 scFv-Fcs and 2G-12 scFv-Fcs) and the 14D9 monoclonal antibody have been targeted to PSVs as well as the ER (ER bodies) in rice, Arabidopsis and tobacco seeds (Torres et al. 2001, Petruccielli et al. 2006, Loos et al. 2011). GAD67/65 and proinsulin were mainly targeted to PSVs in the cotyledons of Arabidopsis seeds (Morandini et al. 2011). Notably, when multimeric antibody subunits in IgA were simultaneously expressed in rice seeds, the assembled antibodies were delivered to PSVs, whereas non-assembled individual antibody subunits were retained in ER-derived PBs (Nicholson et al. 2005).

When several recombinant proteins with KDEL tags were produced in rice endosperm, most of the proteins were trafficked into typical ER-derived PB-Is. Recombinant proteins with free cysteine residues (7Crp, Der p 1 and Der f 2) were found to interact with various types of cysteine-rich prolamins, including 13, 16 and 10 kDa prolamins, which facilitated the deposition of the proteins into PB-Is by the formation of aggregates between the recombinant protein and cysteine-rich prolamins (Yang et al. 2008, Takaiwa et al. 2009, Yang et al. 2012). Such aggregates are retained in the ER, resulting in the formation of ER-derived PBs that contain polymerized aggregates and are sequestered into PB-Is. When interleukin-10 (IL-10) with a KDEL tag was produced in rice seeds, IL-10 itself aggregated in the ER lumen and formed unique, small (mini) ER bodies, which then assembled with prolamins, resulting in the formation of an IL-10–prolamin assembly. In rice endosperm,

Table 2 Subcellular targetting of foreign (recombinant) proteins containing a C-terminal KDEL tag in transgenic plants

Host plant	Recombinant protein	Promoter	Tissue	Localization	Reference
Rice	ScFv T84.66	Ubiquitin-1	Endosperm	PB-I, PB-II (PSV)	Torres et al. (2001)
			Leaf, callus	ER vesicle	
	7Crp	Glutelin <i>GluB-1</i>	Endosperm	PB-I	Takaiwa et al. (2009)
	Der p 1, Der p 1(45–145)	Glutelin <i>GluB-1</i>	Endosperm	PB-I	Yang et al. (2008), Suzuki et al. (2011)
	2 × Amyloid (1–42)	Glutelin <i>GluB-1</i>	Endosperm	PB-I	Oono et al. (2010)
	Schuffe Cry j 2	Glutelin <i>GluB-1</i>	Endosperm	PB-I	Wakasa et al. (2012)
	Modified Der f 2	Glutelin <i>GluB-1</i>	Endosperm	ER body (Derf 2 body)	Yang et al. (2011)
	mIL-10	Glutelin <i>GluB-1</i>	Endosperm	ER body (IL-10 body)	Yang et al. (2012)
	TPC7	Glutelin <i>GluB-1</i>	Endosperm	ER body (TPC7 body)	This study
Wheat	Secretory IgA	Ubiquitin-1	Endosperm	PB-I (subunit) PB-II (PSV) (assembly)	Nicholson et al. (2005)
	Human serum albumin	Ubiquitin-1	Endosperm Leaf	Prolamin PB ER	Arcalis et al. (2004)
Arabidopsis	ScFv MBP10	β-Phaseolin	Cotyledon	Periplasmic space, ER	Van Droogenbroeck et al. (2007)
	ScFv 2G12	β-Phaseolin	Cotyledon	ER-derived vesicle	Loos et al. (2011)
	ScFv HA78	β-Phaseolin	Cotyledon	ER-derived vesicle	Loos et al. (2011)
	GAD67/65mut	β-Phaseolin	Cotyledon	PSV	Morandini et al. (2011)
	Proinsulin	β-Phaseolin	Cotyledon	PSV	Morandini et al. (2011)
	mIL-10	β-Phaseolin	Cotyledon	ER body, apoplast	Morandini et al. (2011)
Tobacco	Antibody 14D9	α' -βconglycinin	Cotyledon	PSV	Petruccielli et al. (2006)
		CaMV 35S	Leaf	ER	
Medicago	Phytase	CaMV 35S	Cotyledon	ER	Abranches et al. (2008)

numerous tiny IL-10 granules and small PB-Is were loosely gathered together, like ER-derived PB structures, or occasionally constituted huge, unique IL-10–PB-I assemblies. In a similar way, hIL-10, produced in *Arabidopsis* seeds, was deposited in ER-derived membrane compartments that were similar to mini ER bodies found in cereal seeds (Morandini et al. 2011). A similar intracellular localization in ER vesicles was also observed when granulocyte–macrophage colony-stimulating factor (GM-CSF), lacking the KDEL tag, was expressed as a secretory protein with a signal peptide in rice endosperm (Ning et al. 2008).

On the other hand, modified mite allergen Der f 2 derivatives aggregated in the ER lumen and were deposited in unique, PB-like structures, referred to as Der f 2 bodies. The Der f 2 body differs in terms of intracellular localization and physicochemical properties from PB-Is and PB-IIs (Yang et al. 2012).

The TPC7 body is different from PB-Is and PB-IIs. Biogenesis of this PB may be essentially identical to that of the Der f 2 body. Formation of the Der f 2 body may be attributed to aggregation properties based on the uniquely enhanced hydrophobic domains within the molecule; the ER-derived Der f 2 body can be formed even from a modified Der f 2 derivative lacking the cysteine residue. However, the critical difference between the TPC7 body and the Der f 2 body is that the former can only form large organelles, due to aggregation properties.

The driving forces for formation of large PBs were examined through the self-assembly of TPC7. A mutation of only the cysteine residue in TPC7 altered the morphology of the TPC7 body (rugged surface). Because the free cysteine residue of TPC7 is not involved in the interaction with cysteine-rich prolamins, this residue may be responsible for the generation of TPC7 bodies by self-assembly or aggregate formation via disulfide bond formation and hydrophobic interactions. Actually, multimer formation by the m1 mutant lacking the single cysteine residue was significantly lower than multimer formation by the native TPC7 protein when the proteins were analyzed by SDS–PAGE under non-reducing conditions. However, multimers of TPC7 could have resulted from the formation of heteromers between endogenous rice proteins such as glutelins or cysteine-rich prolamins and TPC7 via disulfide bond formation. This possibility was excluded by the results showing the absence of immuno-gold-labeled anti-glutelin or anti-prolamin antibody reactivity in TPC7 bodies (Fig. 3). Sequential extraction experiments showed that the extraction efficiency was not affected by the presence or absence of the cysteine residue. Furthermore, not only m1 but also TPC7 could be efficiently extracted with neutral buffer containing 0.15% SDS, but prolamins and glutelins could not. These findings suggest that not only disulfide bond formation but also self-aggregation via hydrophobic interactions between TPC7 molecules may participate in TPC7 body formation, and that aggregation between TPC7 and endogenous seed proteins does not occur in TPC7 bodies.

It is important to examine whether transgenic rice seeds containing TPC7 are effective birch pollen allergens that can

serve as immune tolerogens by oral administration to mice. Indeed, it will be interesting to determine how large TPC7 bodies that accumulate TPC7 are delivered to GALT when they are actually fed to mice.

Materials and Methods

Production of transgenic rice

DNA encoding TPC7 was synthesized using optimized codons, preferentially used in genes encoding many major rice SSPs (Takaiwa et al. 2008). The modified TPC7 gene, with or without a *GluB-1* signal peptide (72 bp) sequence and ER retention signal (KDEL), attached to the N- and C-termini, respectively, was ligated downstream from the endosperm-specific 2.3 kb *GluB-1* promoter, and upstream of a 0.65 kb *GluB-1* terminator (Fig. 1A). In the m1, m3 and m5 constructs, point mutation was carried out, leading to the substitution of cysteine (TGC) with alanine (GCC) in the TPC7 amino acid sequence (Supplementary Fig. S1). The resulting constructs were cloned into the *Hind*III and *Eco*RI sites of the binary vector pGTV-35S-HPT (Takagi et al. 2005a). Rice plants (*Oryza sativa* cv. Kita-ake) were transformed with the plasmids described above by *Agrobacterium*-mediated transformation, as previously described (Qu and Takaiwa 2004). The following primer sets were employed: the TPC7 gene was amplified with TGCCATGGGTGTTTCAATTACGAAG and CGAGCTCTTAG TTGTAAGCATCGGAGTGTC. The point mutation primer set was GGCAACCCCTGATGGAGGAGCCGTATTGAAGATCAA CAAC and GTTGTGATCTTCAATACGGCTCCTCCATCAGG GGTGCC.

Expression of TPC7 in *Escherichia coli* and anti-TPC7 antibody preparation

The coding region of TPC7 was amplified by PCR using the primer set ATCATATGGGTGTTTCAATTACGAAG and ATG GATCCTTAAAGCTCATCCTTGTTG, which contain *Nde*I and *Bam*HI restriction sites, respectively. The PCR product was inserted into the pET-15b expression vector (Novagen). Expression of the TPC7 peptide tagged with 6× His residues in *E. coli* strain Rosetta-gami 2(DE3) (Novagen) was carried out according to the manufacturer's protocol. His-tagged TPC7 peptide was purified using Ni Sepharose 6 Fast Flow (GE Healthcare) and immunized against a rabbit to raise anti-TPC7 antibodies.

Protein extraction and Western blot analysis

Mature seeds from transgenic rice lines were individually ground into a fine powder with a multi-bead shocker (Yasui Kikai). For total protein extraction, extraction buffer [50 mM Tris–HCl (pH 6.8), 8 M urea, 4% (w/v) SDS, 20% (v/v) glycerol, 5% (v/v) 2-mercaptoethanol] was added to the seed powder and vortexed for 1 h at room temperature. The mixture was centrifuged at 13,000×g for 10 min (Tada et al. 2003).

Total seed protein was separated by 12% (w/v) SDS-PAGE and transferred to polyvinylidene fluoride (PVDF) membranes (Millipore) for Western blot analysis. The expression of the TPC7 peptide in mature seeds was detected immunologically, using the rabbit anti-TPC7 antibody described above. Anti-glutelin A, B, C, D, 26 kDa globulin (Glb-1), and 13 kDa cysteine-rich prolamin (RM1), 13 kDa cysteine-poor prolamin (RM2), 10 kDa and 16 kDa prolamin rabbit antibodies were also used, followed by the use of secondary anti-rabbit IgG horseradish peroxidase (HRP) conjugated antibody. Signals were detected with ECL-Western blotting detection reagent (GE Healthcare). SDS-polyacrylamide gels were stained with Coomassie Brilliant Blue (CBB) G-250 (Nacalai Tesque).

The recombinant protein yields were estimated from the intensity of the bands that reacted with the TPC7 antibody, using purified TPC7 produced from *E. coli* as a standard. The band images were scanned and quantified using ImageQuant TL (GE Healthcare).

Glycosylation analysis

Total seed proteins were extracted as described above and precipitated with chloroform/methanol. The proteins were then resuspended in reaction buffer and incubated in the presence or absence of Endo H or PNGase F at 37°C for 1 h, according to the manufacturer's protocol (New England Biolabs). The reaction products were used for SDS-PAGE and Western blot analysis.

Purified TPC7 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. TPC7 was digested using trypsin (Sigma) in ProteaseMAX Surfactant (Promega) at 37°C for 3 h. The N-glycopeptide from TPC7 was completely dried out. The N-glycans were released from the in-gel-digested TPC7 N-glycopeptide, by hydrazinolysis at 100°C for 10 h. The released N-glycans were PA labeled as previously described (1). The reaction products of PA-N-glycans were purified by size fractionation-HPLC using a Shodex Asahipak NH2P-50 4E column (4.6 mm ID $\times$ 250 mm; Showa Denko). The N-glycans were detected by reverse phase-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). Reverse phase-HPLC was performed using a Cosmosil 5C18-AR-II column (4.6 $\times$ 250 mm; Nacalai Tesque) with a 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⁻¹. 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 sugar moieties, were estimated by liquid chromatography-tandem mass spectrometry (LC-MS/MS) using an Agilent Technologies 1200 series instrument equipped with HCT plus software (Bruker Daltonics). For 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, by vol.). The PA-sugar chain was separated using a Shodex Asahipak NH2P-50 2D column (2.0 mm ID $\times$ 150 mm; Showa Denko), by linearly increasing the solvent B concentration from 20% to 55% over 35 min at a flow rate of 0.2 ml min⁻¹. The MS/MS parameters were as follows: scan range m/z 350–2,750; nebulizer flow of 5.0 p.s.i.; dry gas flow rate of 3.0 l min⁻¹; 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 quantity 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 size fractionation-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 a 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⁻¹. Each PA-N-glycan structure was compared with authentic PA-N-glycans purchased from TAKARA and Masuda Chemical Industries or prepared from the enzymatic partial digestion of authentic PA-sugar GlcMan8-PA using Jack Bean α -mannosidase (Sigma).

Sequential protein extraction

Rice SSP was extracted from pulverized mature seeds using a sequential extraction method. Globulins were extracted with 500 μ l of globulin extraction buffer (0.5 M NaCl, 10 mM Tris-HCl, pH 6.8) from 25 mg of seed powder. After removal of the globulins, prolamins were extracted from residual proteins with 500 μ l of prolamin extraction buffer (60% *n*-propanol, 5% 2-mercaptoethanol). Each extraction step was repeated five times and performed by 2 min of sonication followed by centrifugation at 13,000 $\times g$ for 10 min. After the stepwise removal of the globulin and prolamin fractions, glutelins were extracted from residual proteins with 500 μ l of glutelin extraction buffer (1% lactic acid, 1 mM EDTA). The supernatant, containing glutelins, was neutralized with 1 N NaOH and centrifuged to produce the glutelin fraction. The extracted fractions were used for SDS-PAGE and Western blot analysis.

TPC7 and TPC7m1 proteins were also extracted from 25 mg of seed powder with 500 μ l of neutral buffer [10 mM Tris-HCl (pH 7.5), 1% (v/v) Triton X-100], and then sequentially extracted with neutral buffer (10 mM Tris-HCl, pH 7.5) or saline buffer (0.5 M NaCl, 10 mM Tris-HCl, pH 6.8) containing 0.15, 1 and 4% SDS.

In vitro pepsin digestion

A 5 mg aliquot of rice seed powder was dissolved in hydrochloric acid (HCl, pH 1.2) and 30 mM sodium chloride (NaCl) with 0.16% (w/v) pepsin (Sigma). Digestion reactions were incubated for 0, 0.5, 1, 5, 15, 30, 60 or 120 min at 37°C (Takagi et al. 2010) and terminated with the addition of protein extraction buffer, as described above. Reaction mixtures were used for SDS-PAGE and Western blot analysis.

Indirect immunohistochemical analysis by confocal laser scanning microscopy

Developing seeds (15–20 DAF) of transformants and wild-type plants were cross-sectioned with a DTK-1000 Microslicer (Dosaka EM) to approximately 150 μm in thickness for indirect immunohistochemical analysis, as previously described (Yasuda et al. 2006). The anti-TPC7 or maize BiP antibody was used at a 1:300 or 1:200 dilution, respectively, as primary antibody. Rhodamine B was used to stain PB-Is. The specimens were observed with a confocal laser scanning microscope (Fv10i, Olympus) using green fluorescent protein (GFP) and rhodamine filter sets.

Indirect immunohistochemical analysis by TEM

Developing seeds (15–20 DAF) of transformant and wild-type plants were fixed with fixation buffer (4% paraformaldehyde and 1% glutaraldehyde in 50 mM phosphate buffer, pH 7.4) for 2 h at room temperature. The fixed specimens were washed with phosphate buffer and dehydrated in a graded ethanol series. The specimens were embedded in LR White resin (London Resin Company). The embedded specimens were cut with a glass knife using an ultramicrotome (Utracut-R, Leica) and mounted on nickel grids. The section samples were reacted with primary antibody (TPC7 antibody, GluA antibody, RM1 antibody, maize BiP antibody), followed by incubation in a 1:30 dilution of AuroProbe EM Goat Anti-Rabbit IgG (H + L; GE Healthcare). The grids were observed with a transmission electron microscope (JEM-1010, JEOL) running at 80 kV.

RNA preparation and RT-PCR

Total RNA was prepared from developing seeds (15 DAF) of wild-type and transgenic rice plants as previously described (Takaiwa et al. 1987). Aliquots containing 1 μg of total RNA were subjected to cDNA synthesis using mixed primers of random hexamer and oligo(dT) and SuperScript III reverse transcriptase (Invitrogen), according to the manufacturer's protocol. Using the TPC7 gene primer set (described above) and the synthesized cDNAs as templates, PCR amplifications were performed using KOD FX (Toyobo) as follows: 24 cycles of 95°C for 30 s, 60°C for 30 s and 68°C for 40 s.

Supplementary data

Supplementary data are available at PCP online.

Funding

The Ministry of Agriculture, Forestry and Fisheries of Japan [research grant Genomics and Agricultural Innovation, GMC0009 to F.T.] the Ministry of Education, Culture, Sports, Science and Technology of Japan [Grant-in-Aid for Young Scientists, 22688001 to T.K.].

Acknowledgments

The authors thank Ms. M. Utsuno, Ms. Y. Ikemoto, Ms. H. Yajima and Ms. Y. Suzuki, for technical assistance.

References

- Abranches, R., Arcalis, E., Marcel, S., Riberiro-Pedro, M., Rodriguez, J. and Stager, E. (2008) Functional specialization of *Medicago truncatula* leaves and seeds does not affect the subcellular localization of a recombinant protein. *Planta* 227: 649–658.
- Akids, M. and Akids, C.A. (2007) Mechanisms of allergen-specific immunotherapy. *J. Allergy Clin. Immunol.* 119: 780–789.
- Arcalis, E., Marcel, S., Altmann, F., Kolarich, D., Drakakaki, G., Fischer, R. et al. (2004) Unexpected deposition patterns of recombinant proteins in post-endospermic reticulum compartments of wheat endosperm. *Plant Physiol.* 136: 3457–3466.
- Bousquet, P.J., Chinn, S., Janson, C., Kogevinas, M., Burney, P. and Jarvis, D. (2007) Geographical variation in the prevalence of positive skin tests to environmental aeroallergens in the European Community Respiratory Health Survey I. *Allergy* 62: 301–309.
- Breiteneder, H., Pettenburger, K., Bito, A., Valenta, R., Kraft, D., Rumpold, H. et al. (1989) The gene coding for the major birch pollen allergen Betv1, is highly homologous to a pea disease resistance response gene. *EMBO J.* 8: 1935–1938.
- Campana, R., Vrtala, S., Maderegger, B., Dall'Antonia, Y., Zafred, D., Blatt, K. et al. (2011) Altered IgE epitope presentation: a model for hypoallergenic activity revealed for Bet v 1 trimer. *Mol. Immunol.* 48: 431–441.
- Canonica, G.W., Bousquet, J., Mullol, J., Scadding, G.K. and Virchow, J.C. (2007) A survey of the burden of allergic rhinitis in Europe. *Allergy* 62(Suppl. 85), 17–25.
- Cromwell, O., Hafner, D. and Nandy, A. (2011) Recombinant allergens for specific immunotherapy. *J. Allergy Clin. Immunol.* 127: 865–872.
- Daniell, H., Singh, N.D., Mason, H. and Streatfield, S.J. (2009) Plant-made vaccine antigens and biopharmaceuticals. *Trends Plant Sci.* 14: 669–679.
- Frew, A.J. (2003) Immunotherapy of allergic disease. *J. Allergy Clin. Immunol.* 111: S712–S719.
- Gajhede, M., Osmark, P., Poulsen, FM., Ipsen, H., Larsen, J.N., Joost van Neerven, R.J. et al. (1996) X-ray and NMR structure of Bet v 1, the origin of birch pollen allergy. *Nat. Struct. Biol.* 3: 1040–1045.
- Giddings, G., Allison, G., Brooks, D. and Carter, A. (2000) Transgenic plants as factories for biopharmaceuticals. *Nat. Biotechnol.* 18: 1151–1155.
- Gronlund, H. and Gafvelin, G. (2010) Recombinant Bet v 1 vaccine for treatment of allergy to birch pollen. *Hum. Vaccin.* 6: 970–977.
- Kawagoe, Y., Suzuki, K., Tasaki, M., Yasuda, H., Akagi, K., Katoh, E. et al. (2005) The critical role of disulfide bond formation in protein sorting in the endosperm of rice. *Plant Cell* 17: 1141–1153.
- Kawakatsu, T. and Takaiwa, F. (2010) Cereal seed storage protein synthesis: fundamental processes for recombinant protein production in cereal grains. *Plant Biotechnol. J.* 8: 939–953.
- Krishnan, H.B., Franceschi, V.R. and Okita, T.W. (1986) Immunohistochemical studies on the role of the Golgi complex in protein body formation in rice seeds. *Planta* 169: 471–480.
- Larche, M., Akdis, C. and Valenta, R. (2006) Immunological mechanisms of allergen-specific immunotherapy. *Nat. Rev. Immunol.* 6: 761–771.

- Lau, O.S. and Sun, S.S.M. (2009) Plant seeds as bioreactors for recombinant protein production. *Biotechnol. Adv.* 27: 1015–1022.
- Linhart, B. and Valenta, R. (2005) Molecular design of allergy vaccines. *Curr. Opin. Immunol.* 17: 646–655.
- Loos, A., Von Droogenbroeck, B., Hillmer, S., Grass, J., Pabst, M., Castilho, A. et al. (2011) Expression of antibody fragments with a controlled N-glycosylation pattern and induction of endoplasmic reticulum-derived vesicles in seeds of *Arabidopsis*. *Plant Physiol.* 155: 2036–2048.
- Misaki, R., Kimura, Y., Fujiyama, K. and Seki, T. (2001) Glycoproteins secreted from suspension-cultured tobacco BY2 cells have distinct glycan structures from intracellular glycoproteins. *Biosci. Biotechnol. Biochem.* 65: 2482–2488.
- Morandini, F., Avesani, L., Bortesi, L., Van Droogenbroeck, B., De Wilde, K., Arcalis, E. et al. (2011) Non-food/feed seeds as biofactories for the high-yield production of recombinant pharmaceuticals. *Plant Biotechnol. J.* 9: 911–921.
- Nagamine, A., Matsusaka, H., Ushijima, T., Kawagoe, Y., Ogawa, M., Okita, T.W. et al. (2011) A role for the cysteine-rich 10 kDa prolamin in protein body I formation in rice. *Plant Cell Physiol.* 52: 1003–1016.
- Nicholson, L., Gonzalez-Melendi, P., von Dolleweerd, C., Tuck, H., Perrin, Y., Ma, J. et al. (2005) A recombinant multimeric immunoglobulin expressed in rice shows assembly-dependent subcellular localization in endosperm cells. *Plant Biotechnol. J.* 3: 115–127.
- Niederberger, V., Horak, F., Vrtala, S., Spitzauer, S., Krauth, M.T., Valent, P. et al. (2004) Vaccination with genetically engineered allergens prevents progression of allergic disease. *Proc. Natl Acad. Sci. USA* 101: 14677–14682.
- Ning, T., Xie, T., Qiu, Q., Yang, W., Zhou, S., Zhou, L. et al. (2008) Oral administration of recombinant human granulocyte-macrophage colony stimulating factor expressed in rice endosperm can increase leukocytes in mice. *Biotechnol. Lett.* 30: 1679–1686.
- Nochi, T., Takagi, H., Yuki, Y., Yang, L., Masumura, T., Mejuma, M. et al. (2007) Rice-based mucosal vaccine as a global strategy for cold-chain- and needle-free vaccination. *Proc. Natl Acad. Sci. USA* 104: 10986–10991.
- Ogawa, M., Kumamaru, T., Sato, H., Iwata, N., Omura, T., Kasai, Z. et al. (1987) Purification of protein body-I of rice seed and its polypeptide composition. *Plant Cell Physiol.* 28: 1517–1527.
- Oono, Y., Wakasa, Y., Hirose, S., Yang, L., Sakuta, C. and Takaiwa, F. (2010) Analysis of ER stress in developing rice endosperm accumulating β -amyloid peptide. *Plant Biotechnol. J.* 8: 691–718.
- Petrucelli, S., Otegui, M.S., Lareu, F., Dinh, O.T., Firchette, A.-C., Circosta, A. et al. (2006) A KDEL-tagged monoclonal antibody is efficiently retained in the endoplasmic reticulum in leaves, but is both partially secreted and sorted to protein storage vacuoles in seeds. *Plant Biotechnol. J.* 4: 511–527.
- Qu, L.Q. and Takaiwa, F. (2004) Evaluation of tissue specificity and expression strength of rice seed component gene promoters in transgenic rice. *Plant Biotechnol. J.* 2: 113–125.
- Radauer, C. and Breiteneder, H. (2006) Pollen allergens are restricted to few protein families and show distinct patterns of species distribution. *J. Allergy Clin. Immunol.* 117: 141–147.
- Sharma, A.K. and Sharma, M.K. (2009) Plants as bioreactors: recent developments and emerging opportunities. *Biotechnol. Adv.* 27: 811–832.
- Stoger, E., Ma, J.K., Fischer, R. and Christou, P. (2005) Sowing the seeds of success: pharmaceutical proteins from plants. *Curr. Opin. Biotechnol.* 16: 167–173.
- Streatfield, S.J. (2006) Mucosal immunization using recombinant plant-based oral vaccines. *Methods* 38: 150–157.
- Suzuki, K., Kaminuma, O., Yang, L., Takai, T., Mori, A., Umez-Goto, M. et al. (2011) Prevention of allergic asthma by vaccination with transgenic rice seed expressing mite allergen: induction of allergen-specific oral tolerance without bystander suppression. *Plant Biotechnol. J.* 9: 982–990.
- Tada, Y., Utsumi, S. and Takaiwa, F. (2003) Foreign gene products can be enhanced by introduction into storage protein mutants. *Plant Biotechnol. J.* 1: 411–422.
- Takagi, H., Hiroi, T., Hirose, S., Yang, L. and Takaiwa, F. (2010) Rice seed ER-derived protein body as an efficient delivery vehicle for oral tolerogenic peptides. *Peptide* 3: 1421–1425.
- Takagi, H., Hiroi, T., Yang, L., Tada, Y., Yuki, Y., Takamura, K. et al. (2005a) A rice-based edible vaccine expressing multiple T cell epitopes induces oral tolerance for inhibition of Th2-mediated IgE responses. *Proc. Natl Acad. Sci. USA* 102: 17525–17530.
- Takagi, H., Saito, S., Yang, L., Nagasaka, S., Nishizawa, N. and Takaiwa, F. (2005b) Oral immunotherapy against a pollen allergy using a seed based peptide vaccine. *Plant Biotechnol. J.* 3: 521–533.
- Takaiwa, F. (2011) Seed-based oral vaccines as allergen-specific immunotherapies. *Hum. Vaccin.* 7: 357–366.
- Takaiwa, F., Hirose, S., Takagi, H., Yang, L. and Wakasa, Y. (2009) Deposition of a recombinant peptide in ER-derived protein bodies by retention with cysteine-rich prolamins in transgenic rice seed. *Planta* 229: 1147–1158.
- Takaiwa, F., Kikuchi, S. and Oono, K. (1987) A rice glutelin gene family: a major type of glutelin mRNAs can be divided into two classes. *Mol. Gen. Genet.* 208: 15–22.
- Takaiwa, F., Takagi, H., Hirose, S. and Wakasa, Y. (2007) Endosperm tissue is good production platform for artificial recombinant proteins in transgenic rice. *Plant Biotechnol. J.* 5: 84–92.
- Takaiwa, F., Yang, L. and Yasuda, H. (2008) Health-promoting transgenic rice: application of rice seeds as a direct delivery system for bioactive peptides in human health. In *Biotechnology in Agriculture and Forestry*. Edited by Hirano, H.Y. pp. 357–373. Springer, Heidelberg.
- Tanaka, K., Sugimoto, T., Ogawa, M. and Kasai, Z. (1980) Isolation and characterization of two types of protein bodies in rice endosperm. *Agric. Biol. Chem.* 44: 1633–1639.
- Tiwari, S., Verma, P.C., Singh, P. and Tuli, R. (2009) Plants as bioreactors for the production of vaccine antigens. *Biotechnol. Adv.* 27: 449–467.
- Torres, E., Gonzalez-Melendi, P., Stoger, E., Shaw, P., Twyman, R.M., Nicholson, L. et al. (2001) Native and artificial reticuloplasmins co-accumulate in distinct domains of the endoplasmic reticulum and in post-endoplasmic reticulum compartments. *Plant Physiol.* 127: 1212–1223.
- Valenta, R., Ball, T., Focke, M., Linhart, B., Mothes, N., Niederberger, V. et al. (2004) Immunotherapy of allergic disease. *Adv. Immunol.* 82: 105–153.
- Valenta, R., Ferreira, F., Focke-Tajkl, M., Linhart, B., Niederberger, V., Swoboda, I. et al. (2010) From allergen genes to allergy vaccines. *Annu. Rev. Immunol.* 28: 211–241.
- Van Droogenbroeck, B., Cao, J., Stadlmann, J., Altmann, F., Colanesi, S., Hilmer, S. et al. (2007) Aberrant localization and underglycosylation of highly accumulating single-chain Fv-Fc antibodies in transgenic *Arabidopsis* seeds. *Proc. Natl Acad. Sci. USA* 104: 1430–1435.

- Vieths, S., Scheurer, S. and Ballmer-Weber, B. (2002) Current understanding of cross-reactivity of food allergens and pollen. *Ann. NY Acad. Sci.* 964: 47–68.
- Vrtala, S., Hirtenlehner, K., Susani, M., Akdis, M., Kussebi, F., Akdis, C.A. et al. (2001) Genetic engineering of a hypoallergenic trimer of the major birch pollen allergen Bet v 1. *FASEB* 15: 2045–2047.
- Wallner, M., Stöcklinger, A., Thalhamer, T., Bohle, B., Vogel, L., Briza, P. et al. (2007) Allergy multivaccines created by DNA shuffling of tree pollen allergens. *J. Allergy Clin. Immunol.* 120: 374–380.
- Wakasa, Y., Hayashi, S. and Takaiwa, F. (2012) Expression of OsBiP4 and OsBiP5 is highly correlated with the endoplasmic reticulum stress response in rice. *Planta* 236: 1519–1527.
- Wakasa, Y., Takagi, H., Hirose, S., Yang, L., Saeki, M., Nishimura, T. et al. (2013) Oral immunotherapy with transgenic rice seed containing destructed Japanese cedar pollen allergens, Cry j 1 and Cry j 2, against Japanese cedar pollinosis. *Plant Biotechnol. J.* 11: 66–67.
- Wandelt, C.I., Khan, M.R.I., Craig, S., Schroeder, H.E., Spencer, D. and Higgins, T.J.V. (1992) Vicilin with carboxy-terminal KDEL is retained in the endoplasmic reticulum and accumulates to high levels in the leaves of transgenic plants. *Plant J.* 2: 181–192.
- Weber, R.W. (2003) Patterns of pollen cross-allergenicity. *J. Allergy Clin. Immunol.* 112: 229–239.
- Yang, L., Hirose, S., Suzuki, K., Hiroi, T. and Takaiwa, F. (2012a) Expression of hypoallergenic Der f 2 derivatives with altered intramolecular disulphide bonds induces the formation of novel ER-derived protein bodies in transgenic rice seeds. *J. Exp. Bot.* 63: 2947–2959.
- Yang, L., Hirose, S., Takahashi, H., Kawakatsu, T. and Takaiwa, F. (2012b) Recombinant protein yield is enhanced by specific suppression of endogenous seed proteins at the same deposit site in transgenic rice seed. *Plant Biotechnol. J.* 10: 1035–1045.
- Yang, L., Suzuki, K., Hirose, S., Wakasa, Y. and Takaiwa, F. (2007) Development of transgenic rice seed accumulating a major Japanese cedar pollen allergen (Cry j 1) structurally disrupted for oral immunotherapy. *Plant Biotechnol. J.* 5: 815–826.
- Yasuda, H., Hayashi, Y., Jomori, T. and Takaiwa, F. (2006) The correlation between expression and localization of a foreign gene product in rice endosperm. *Plant Cell Physiol.* 47: 756–763.

MGVFNYEVETPSVISAARLFKSYVLDGDKLIPKVAPQAIT
SVENIEGNGGPGTIKKISFPEGFPFKYVKDRVDEVDHTNF
TYSYTVIEGGPVGDTLEKISNEIKIVATPDGGCVLKINNK
YHTKGNHEVKAEQVKASKEMGETLLRAVESYLLAHSDAYN

Fig. S1 Amino acid sequence of TPC7 protein.

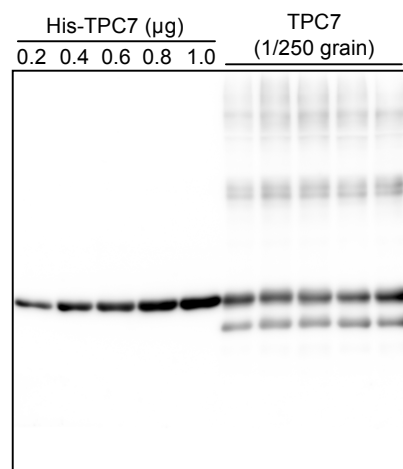


Fig. S2 Accumulation levels of TPC7 in mature seeds of transgenic rice plants. The amounts of TPC7 in mature dry seeds of the transgenic line that produced the largest amount of TPC7 (T3 generation) were determined by western blot analysis using purified His-TPC7 produced in *E. coli* as a standard. Five seeds from this transgenic line were extracted with urea/SDS buffer and the amounts of TPC7 in the mature seeds was determined the bands on the western blot image were quantified with ImageQuant TL. The highest level of TPC7 that accumulated in mature seeds was approximately 207 $\pm$ 29 μ g/grain.

Core-Mannose type	$ \begin{array}{c} \text{Man}\alpha 1 \\ \quad \diagdown \quad \diagup \\ \quad \quad 6 \quad 3 \\ \quad \quad \text{Man}\beta 1 - 4\text{GlcNAc}\beta 1 - 4\text{GlcNAc} - \text{PA} \\ \quad \quad \diagup \quad \diagdown \\ \text{Man}\alpha 1 \end{array} $	M3
	$ \begin{array}{c} \text{Man}\alpha 1 \\ \quad \diagdown \quad \diagup \\ \quad \quad 6 \quad 3 \\ \quad \quad \text{Man}\beta 1 - 4\text{GlcNAc}\beta 1 - 4\text{GlcNAc} - \text{PA} \\ \quad \quad \diagup \quad \diagdown \\ \text{Man}\alpha 1 \end{array} $	M4
	$ \begin{array}{c} \text{Man}\alpha 1 \\ \quad \diagdown \quad \diagup \\ \quad \quad 6 \quad 3 \\ \quad \quad \text{Man}\beta 1 - 4\text{GlcNAc}\beta 1 - 4\text{GlcNAc} - \text{PA} \\ \quad \quad \diagup \quad \diagdown \\ \text{Man}\alpha 1 \end{array} $	M5
High-Mannose type	$ \begin{array}{c} \text{Man}\alpha 1 - 2\text{Man}\alpha 1 \\ \quad \quad \quad \diagdown \quad \diagup \\ \quad \quad \quad \quad 6 \quad 3 \\ \quad \quad \quad \quad \text{Man}\beta 1 - 4\text{GlcNAc}\beta 1 - 4\text{GlcNAc} - \text{PA} \\ \quad \quad \quad \quad \diagup \quad \diagdown \\ \text{Man}\alpha 1 - 2\text{Man}\alpha 1 \end{array} $	M7A
	$ \begin{array}{c} \text{Man}\alpha 1 - 2\text{Man}\alpha 1 \\ \quad \quad \quad \diagdown \quad \diagup \\ \quad \quad \quad \quad 6 \quad 3 \\ \quad \quad \quad \quad \text{Man}\beta 1 - 4\text{GlcNAc}\beta 1 - 4\text{GlcNAc} - \text{PA} \\ \quad \quad \quad \quad \diagup \quad \diagdown \\ \text{Man}\alpha 1 - 2\text{Man}\alpha 1 \end{array} $	M7B
	$ \begin{array}{c} \text{Man}\alpha 1 - 2\text{Man}\alpha 1 \\ \quad \quad \quad \diagdown \quad \diagup \\ \quad \quad \quad \quad 6 \quad 3 \\ \quad \quad \quad \quad \text{Man}\beta 1 - 4\text{GlcNAc}\beta 1 - 4\text{GlcNAc} - \text{PA} \\ \quad \quad \quad \quad \diagup \quad \diagdown \\ \text{Man}\alpha 1 - 2\text{Man}\alpha 1 \end{array} $	M8A
	$ \begin{array}{c} \text{Man}\alpha 1 - 2\text{Man}\alpha 1 \\ \quad \quad \quad \diagdown \quad \diagup \\ \quad \quad \quad \quad 6 \quad 3 \\ \quad \quad \quad \quad \text{Man}\beta 1 - 4\text{GlcNAc}\beta 1 - 4\text{GlcNAc} - \text{PA} \\ \quad \quad \quad \quad \diagup \quad \diagdown \\ \text{Man}\alpha 1 - 2\text{Man}\alpha 1 \end{array} $	M8B
	$ \begin{array}{c} \text{Man}\alpha 1 - 2\text{Man}\alpha 1 \\ \quad \quad \quad \diagdown \quad \diagup \\ \quad \quad \quad \quad 6 \quad 3 \\ \quad \quad \quad \quad \text{Man}\beta 1 - 4\text{GlcNAc}\beta 1 - 4\text{GlcNAc} - \text{PA} \\ \quad \quad \quad \quad \diagup \quad \diagdown \\ \text{Man}\alpha 1 - 2\text{Man}\alpha 1 \end{array} $	M8C
	$ \begin{array}{c} \text{Man}\alpha 1 - 2\text{Man}\alpha 1 \\ \quad \quad \quad \diagdown \quad \diagup \\ \quad \quad \quad \quad 6 \quad 3 \\ \quad \quad \quad \quad \text{Man}\beta 1 - 4\text{GlcNAc}\beta 1 - 4\text{GlcNAc} - \text{PA} \\ \quad \quad \quad \quad \diagup \quad \diagdown \\ \text{Man}\alpha 1 - 2\text{Man}\alpha 1 \end{array} $	M9
	$ \begin{array}{c} \text{Man}\alpha 1 - 2\text{Man}\alpha 1 \\ \quad \quad \quad \diagdown \quad \diagup \\ \quad \quad \quad \quad 6 \quad 3 \\ \quad \quad \quad \quad \text{Man}\beta 1 - 4\text{GlcNAc}\beta 1 - 4\text{GlcNAc} - \text{PA} \\ \quad \quad \quad \quad \diagup \quad \diagdown \\ \text{Man}\alpha 1 - 2\text{Man}\alpha 1 \end{array} $	
	$ \begin{array}{c} \text{Man}\alpha 1 - 2\text{Man}\alpha 1 \\ \quad \quad \quad \diagdown \quad \diagup \\ \quad \quad \quad \quad 6 \quad 3 \\ \quad \quad \quad \quad \text{Man}\beta 1 - 4\text{GlcNAc}\beta 1 - 4\text{GlcNAc} - \text{PA} \\ \quad \quad \quad \quad \diagup \quad \diagdown \\ \text{Man}\alpha 1 - 2\text{Man}\alpha 1 \end{array} $	
	$ \begin{array}{c} \text{Man}\alpha 1 - 2\text{Man}\alpha 1 \\ \quad \quad \quad \diagdown \quad \diagup \\ \quad \quad \quad \quad 6 \quad 3 \\ \quad \quad \quad \quad \text{Man}\beta 1 - 4\text{GlcNAc}\beta 1 - 4\text{GlcNAc} - \text{PA} \\ \quad \quad \quad \quad \diagup \quad \diagdown \\ \text{Man}\alpha 1 - 2\text{Man}\alpha 1 \end{array} $	
$\alpha 1,3$ -Glucose linked	$ \begin{array}{c} \text{Man}\alpha 1 \\ \quad \diagdown \quad \diagup \\ \quad \quad 6 \quad 3 \\ \quad \quad \text{Man}\beta 1 - 4\text{GlcNAc}\beta 1 - 4\text{GlcNAc} - \text{PA} \\ \quad \quad \diagup \quad \diagdown \\ \text{Glc}\alpha 1 - 3\text{Man}\alpha 1 - 2\text{Man}\alpha 1 - 2\text{Man}\alpha 1 \end{array} $	GlcM7B
	$ \begin{array}{c} \text{Man}\alpha 1 \\ \quad \diagdown \quad \diagup \\ \quad \quad 6 \quad 3 \\ \quad \quad \text{Man}\beta 1 - 4\text{GlcNAc}\beta 1 - 4\text{GlcNAc} - \text{PA} \\ \quad \quad \diagup \quad \diagdown \\ \text{Glc}\alpha 1 - 3\text{Man}\alpha 1 - 2\text{Man}\alpha 1 - 2\text{Man}\alpha 1 \end{array} $	GlcM8A
	$ \begin{array}{c} \text{Man}\alpha 1 \\ \quad \diagdown \quad \diagup \\ \quad \quad 6 \quad 3 \\ \quad \quad \text{Man}\beta 1 - 4\text{GlcNAc}\beta 1 - 4\text{GlcNAc} - \text{PA} \\ \quad \quad \diagup \quad \diagdown \\ \text{Glc}\alpha 1 - 3\text{Man}\alpha 1 - 2\text{Man}\alpha 1 - 2\text{Man}\alpha 1 \end{array} $	GlcM9

Fig. S3 Structures of the PA-sugar chains used in this study. The structures are divided into three groups, core-Mannose structure, high-Mannose structure, and $\alpha 1,3$ -Glucose-linked structure. The PA-sugar chains listed in this table were purchased from TaKaRa (Shiga, Japan) and Masuda Chemical Industries (Kagawa, Japan) or were prepared by enzymatic digestion, as described in Materials and Methods. Man, Mannose; GlcNAc, *N*-Acetylglucosamine; Glc, Glucose.

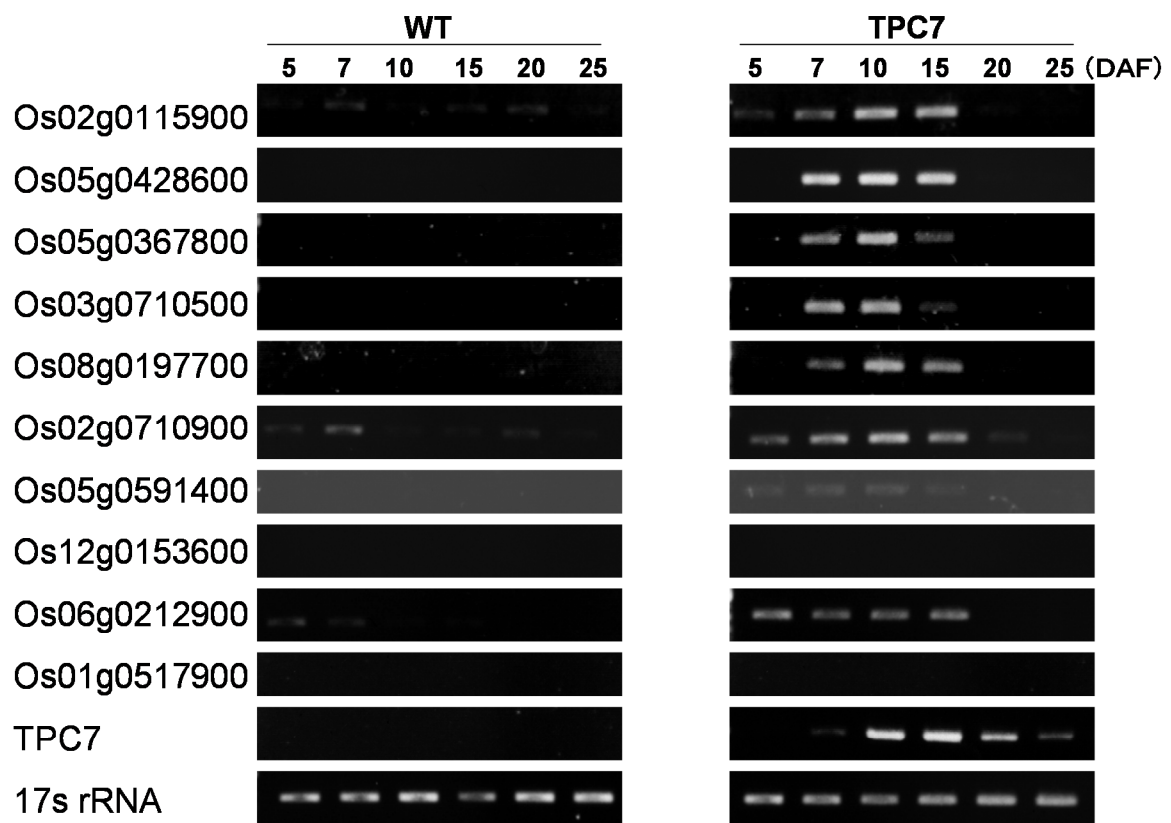


Fig. S4 Comparison of gene expression patterns of 10 BiPs during seed development by RT-PCR.

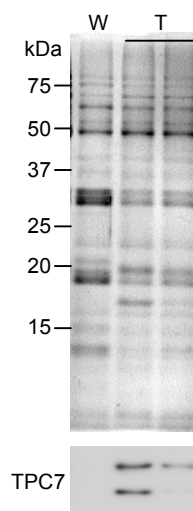


Fig. S5 Subcellular fractionation of TPC7 bodies. TPC7 seeds were pulverized and homogenized with 10 mM Tris-HCl (pH 7.5). The homogenate was centrifuged at $100 \times g$ for 10 min to remove debris. Then, the supernatant was centrifuged at $10,000 \times g$ for 10 min to collect TPC7 bodies. The precipitate containing TPC7 bodies was subjected to SDS-PAGE (upper) and western blot analysis using anti-TPC7 antibody (lower). W, wild type; T, TPC7 transformant.

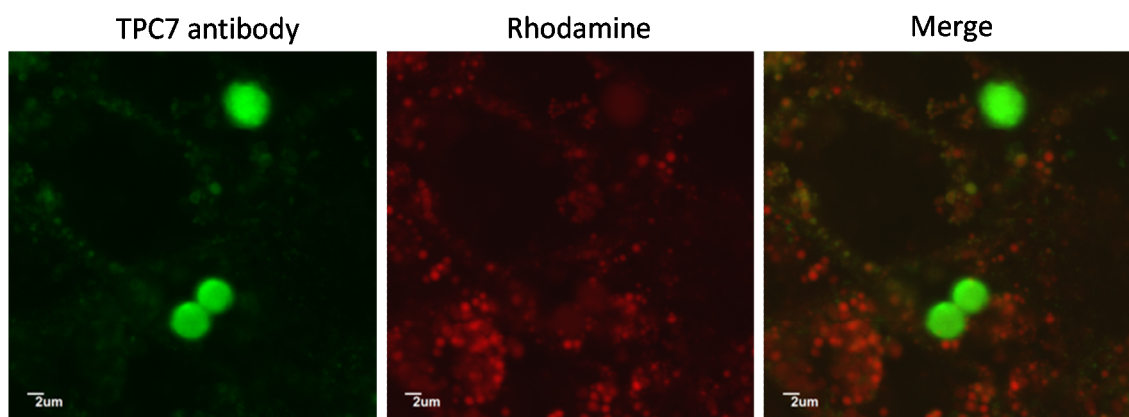


Fig. S6 Indirect immuno-histochemical analysis of unglycosylated TPC7 in subaleurone layer cells in developing seeds. The red signal indicates PB-I labeled with rhodamine and the green signal indicates immunofluorescence of TPC7 immunostained with anti-TPC7 antibody.

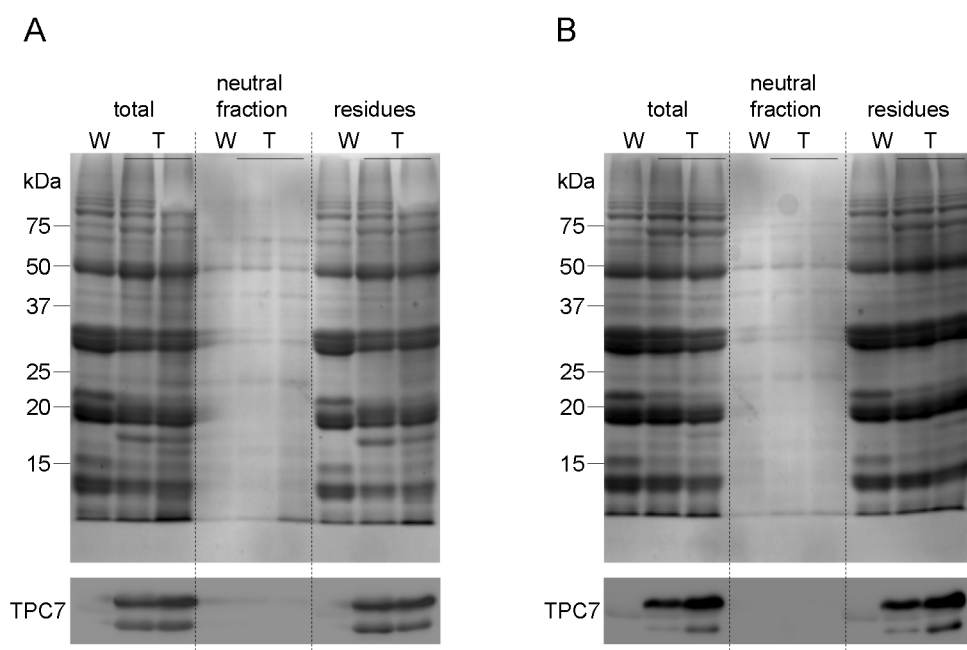


Fig. S7 Solubility of TPC7 and TPC7m1 proteins in neutral buffer. Extraction of TPC7 (A) and TPC7m1 (B) proteins from transgenic seeds using neutral buffer (10 mM Tris-HCl [pH 7.5], 1% [v/v] Triton X-100). Each extracted protein fraction was subjected to SDS-PAGE (upper) and western blot analysis using the anti-TPC7 antibody (lower).

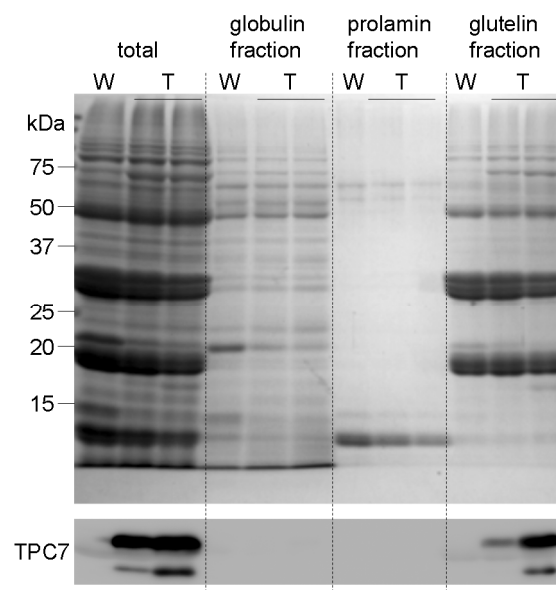


Fig. S8 Sequential extraction of seed proteins from the wild type (W) and TPC7m1 transformant (T). Each extracted protein fraction was subjected to SDS-PAGE (upper) and western blot analysis using the anti-TPC7 antibody (lower).