



Specific region affects the difference in accumulation levels between apple food allergen Mal d 1 and birch pollen allergen Bet v 1 which are expressed in vegetative tissues of transgenic rice

Fumio Takaiwa¹ · Yuko Ogo^{1,2} · Yuhya Wakasa¹

Received: 7 July 2018 / Accepted: 14 October 2018
© Springer Nature B.V. 2018

Abstract

Key message Specific domain of the Mal d 1 was identified to be mainly involved in higher accumulation level in vegetative tissues of transgenic rice than the Bet v 1.

Abstract Apple food allergen Mal d 1 and birch pollen allergen Bet v 1 belong to the same pathogen related protein 10 (PR10) family. When green fluorescent protein (GFP) fused to either of these allergens was expressed as a secretory protein in transgenic rice by ligating an N terminal signal peptide and a C terminal KDEL ER retention signal under the control of the maize ubiquitin constitutive promoter, the GFP:Mald1 highly accumulated in various tissues, whereas accumulation level of the GFP:Betv1 was remarkably reduced in vegetative tissues except for seed. Analysis by RT-PCR exhibited that there was little difference in their transcript levels, indicating the involvement of post-transcriptional regulation. To investigate the cause of such difference in accumulation levels, deletion analysis of the Mal d 1 and domain swapping between them were carried out in transgenic rice. The results showed that the region between positions 41–90 in the Mal d 1 is predominantly implicated in higher level accumulation in vegetative tissues as well as seed as compared with the Bet v 1. The GFP:Mald1 was localized in oligomeric form within ER lumen or ER-derived particles in vegetative tissues, whereas in seed mainly deposited into novel huge ER-derived protein bodies with the size of 5–10 µm in aleurone cells.

Keywords Apple allergen · Bet v 1 · PR10 · Protein body · Mal d 1 · Transgenic rice

Abbreviations

Cys	Cysteine
BiP	Binding protein
DTT	Dithiothreitol
ER	Endoplasmic reticulum
ELP	Elastin-like polypeptide
GFP	Green fluorescence protein
OAS	Oral allergy syndrome

PB	Protein body
2-MER	2-Mercaptoethanol
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
RT-PCR	Reverse transcription-polymerase chain reactions
PR10	Pathogen related protein 10
PDIL	Protein disulfide isomerase like

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s11103-018-0789-7>) contains supplementary material, which is available to authorized users.

✉ Fumio Takaiwa
takaiwa@affrc.go.jp

¹ Plant Molecular Farming Unit, Division of Biotechnology, Institute of Agrobiological Sciences, National Agriculture and Food Research Organization, Owashi 1-2, Tsukuba, Ibaraki 305-8634, Japan

² Institute of Crop Science, National Agriculture and Food Research Organization, Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan

Introduction

Mal d 1 is the major allergen of apple fruit eliciting oral allergy syndrome (OAS) such as itching and swelling of lips, tongue and throat, which belongs to the family of pathogen-related 10 (PR10) proteins (Fernandes et al. 2013; Mari et al. 2005). It has been reported that more than 70% of birch pollen-allergic patients sensitizing to the major birch pollen allergen Bet v 1 display OAS after apple consumption (Geroldinger-Simic et al. 2011). Namely, the OAS by apple is provoked by the allergic cross-reaction between the Mal

d 1 and the homologous Bet v 1 specific IgE, indicating the presence of common IgE epitopes and tertiary structure required for its binding to the Bet v 1 specific IgE (Ebner et al. 1991; Geroldinger-Simic et al. 2011).

Mal d 1 is composed of 159 amino acids with the estimated molecular mass of 17.5 kDa. Its subcellular localization is estimated to be intracellular and cytoplasmic like most of PR10 proteins, since there is no signal peptide sequence directing to the endomembrane system as a secretory protein (Fernandes et al. 2013). Mal d 1 is mainly present in the pulp and peel of apple, but is not only expressed in apple fruit but also in various vegetative tissues and seed (Beuning et al. 2004; Marzban et al. 2005). Although the biological function of Mal d 1 remains less known, its expression has been studied to respond to fungal or bacterial infection and also to be activated by stress and storage, suggesting that it participates in defense mechanism against pathogens or abiotic stress like other PR10 proteins (Agarwal and Agarwal 2014; Fernandes et al. 2013).

Mal d 1 is encoded by a multigene family composed of 31 members, which are divided into four groups based on amino acid sequence similarities, Mal d 1.01, Mal d 1.02, Mal d 1.03 and Mal d 1.04 (Mal d 1 a-d) (Pagliarani et al. 2013). Mal d 1.01 and Mal d 1.02 are predominantly expressed in apple fruit (peel and pulp) (Marzban et al. 2005). Amino acid sequence identities between Mal d 1 isoforms and Bet v 1 range from 55 to 68% (Ma et al. 2006).

Rice seed provides good production platform for high-value recombinant proteins in terms of scalability, safety, stability and cost-effectiveness (Stoger et al. 2005; Takaiwa et al. 2015). Especially, when recombinant proteins are produced as secretory protein and then deposited into protein-bodies (PBs) in the endosperm cells, they are highly and stably accumulated (Takaiwa et al. 2017). Moreover, when orally administered, it has been known that recombinant proteins bio-encapsulated in PBs are effectively delivered to small intestine without degradation with digestive enzymes in gastrointestinal tract, providing an ideal delivery system to gut-associated lymphoid tissues (GALT) as oral mucosal vaccines (Hofbauer and Stoger 2013; Takaiwa et al. 2015).

We previously generated transgenic rice seeds containing versatile hypo-allergenic derivatives of major birch pollen allergen Bet v 1 referred to as tree pollen chimera7 (TPC7) or TPC9, which have been developed as allergy vaccines (tolerogens) against the major allergens of trees belonging to the Fagales order Bet v 1 family for allergen-specific immunotherapy (Wang et al. 2013). When these native Bet v 1 and its derivatives (TPC7 and TPC9) were expressed as secretory protein in rice seeds under the control of the endosperm-specific glutelin *GluB-1* promoter by ligating the glutelin *GluB-1* signal peptide and the KDEL ER retention signal to their N- and C-termini, they were deposited into remarkably huge spherical ER-derived PBs designated as TPC7 bodies

with the size of > 10 μm (Ogo et al. 2014; Wang et al. 2013). Furthermore, their accumulation levels reached about 23% of total seed proteins (Ogo et al. 2014). Such ability to form huge PBs is expected to have a potential to produce high amounts of foreign recombinant proteins by fusion to them. Moreover, it is interesting to examine whether properties yielding such huge PB formation and high productivity are retained in vegetative tissues, when expressed in vegetative tissues.

In this study, we examined whether Mal d 1, Bet v 1 and TPC7 within the same PR10 family are deposited into huge ER-derived PBs in various tissues of transgenic rice. When expressed under the control of the maize ubiquitin constitutive promoter, the Mal d 1 stably and highly accumulated in vegetative tissues as well as seed, whereas the accumulation levels of the Bet v 1 and TPC7 were very low in vegetative tissues except for seed. The difference in accumulation levels in vegetative tissues was shown to be caused by post-translational regulation. Notably, this was mainly attributed to the region between positions 41–90 based on the results of deletions of Mal d 1 and domain-swapping between the Mal d 1 and Bet v 1. Taken together, it is the first example showing that the specific region of the expressed product is participated in the post-translation regulation, although there are many reports on post-translational regulation through subcellular localization, protein stabilization or protein modification (such as glycosylation and phosphorylation etc.) of foreign recombinant proteins (Egelkroun et al. 2012).

Materials and methods

Construction of expression vectors and generation of transgenic rice

The DNA sequences encoding Mal d 1, Bet v 1 and TPC7 were optimized for translation based on the codons frequently used in rice seed storage protein genes and their DNA sequences were synthesized by GenScript Corporation (NJ, USA). The N terminal signal peptide of the glutelin *GluB-1* that is required for translocation into the ER and the KDEL ER retention signals were attached to their N and C termini, respectively. The gene encoding the green fluorescence protein (GFP) reporter was inserted in frame between the *GluB-1* signal and these PR10 related genes. In order to express specifically in the endosperm or constitutively in various vegetative tissues, 2.3 kb glutelin *GluB-1* promoter or maize ubiquitin promoter was ligated upstream of these three fusion constructs, which was followed by the 0.65 kb of glutelin *GluB-1* terminator (Fig. 1b).

These gene expression cassettes inserted into entry vector were transferred into destination binary vector (p35SHPTAg7-GW) harboring the CaMV35S

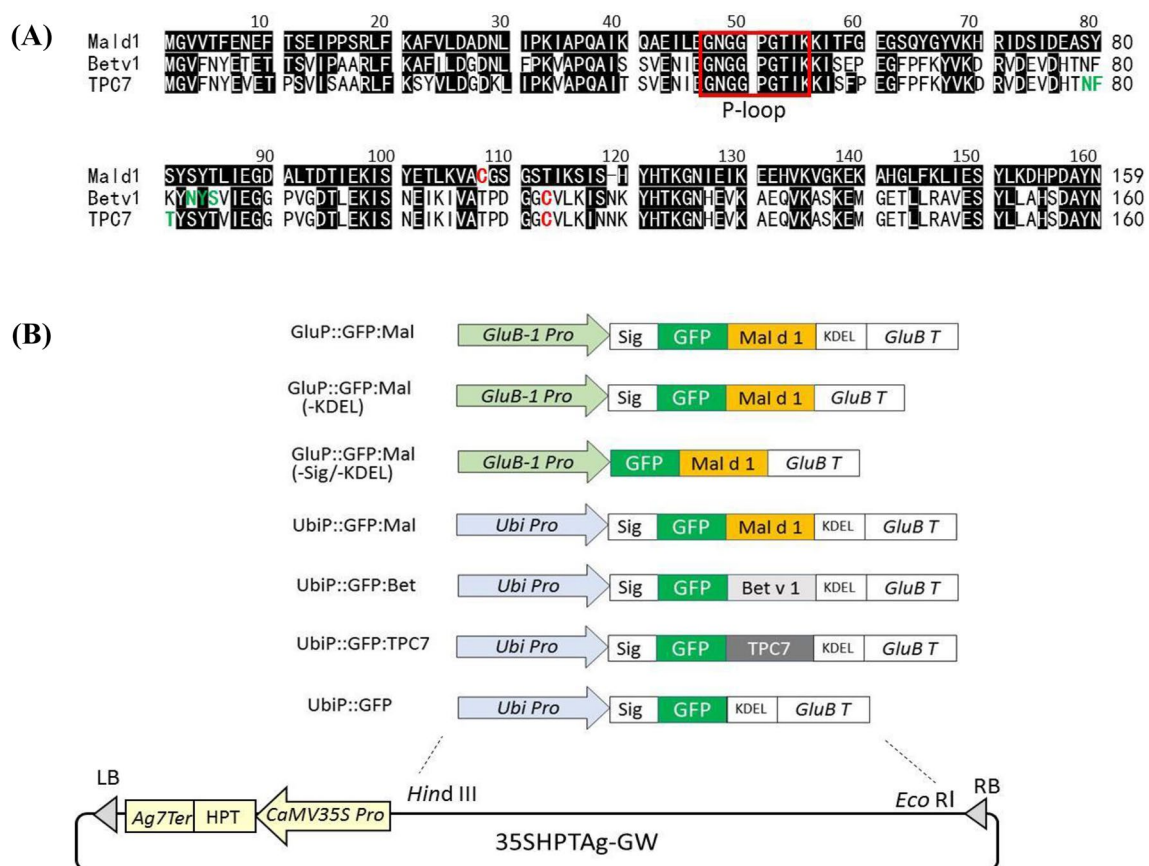


Fig. 1 Gene constructs used for expression of Mal d 1, Bet v 1 and TPC7 in transgenic rice. **a** Alignment of Mal d 1, Bet v 1 and TPC7 amino acid sequences. Cysteine residue and glycosylation sites are indicated by red and green colors. Red square represents P loop motif. **b** Schematic representation of the gene constructs used for expression of GFP, GFP:Mal d 1, GFP:Bet v 1 and GFP:TPC7 in transgenic rice. Mal d 1, Bet v 1 and TPC7 were expressed as fusion

proteins with GFP under the control of the 2.3 kb glutelin *GluB-1* promoter and maize ubiquitin promoter. SP and *GluB-1T* represent the glutelin B1 signal peptide sequence and 0.65 kb 3' untranslated region of *GluB-1*, respectively. *KDEL* endoplasmic reticulum retention signal, *HPT* hygromycin phosphotransferase gene, *CaMV35S P* cauliflower mosaic virus 35S promoter, *UbiP* maize ubiquitin gene promoter, *RB* right border, *LB* left border

promoter::hygromycin phosphotransferase gene (*HPT*) as selectable marker gene via the MultiSite Gateway LR Clonase reaction (Invitrogen) (Wakasa et al. 2006). These binary vector plasmids containing various expression constructs were introduced into rice genome via *Agrobacterium*-mediated transformation (Goto et al. 1999).

For each construct, more than 16 independent stable transgenic rice lines (16–30 lines) were generated. Transgenic rice plants were selected by hygromycin resistance and grown in the closed control greenhouse (33 °C, 16 h light/8 h dark cycle).

SDS-PAGE and immunoblot analysis

Individual mature grain (four grains per independent transgenic line) and leaves and roots frozen in liquid nitrogen were ground into a fine powder using a Multibeads shocker (Yasui Kikai, Osaka, Japan). Total proteins were extracted

from the powder of a mature grain, leaves and roots with 400 µl (grain) or 200 µl (leaves and roots) of urea-SDS buffer (50 mM Tris-HCl, pH 6.8, 8 M urea, 4% SDS, 5% 2-mercaptoethanol (2-MER), 20% glycerol) as described previously (Tada et al. 2003). After separation of total protein extract (2–5 µl) by electrophoresis on 13% SDS-PAGE, proteins were visualized by Coomassie Brilliant Blue (CBB)-R250 staining or were transferred to PVDF membranes (Millipore, Billerica, MA, USA) for immunodetection with rabbit anti-GFP or appropriate antibodies. Rabbit antibodies against TPC7, rice glutelin A (GluA), rice 14 kDa Cys-rich prolamin (RM1), rice 13 kDa Cys-poor prolamin (RM4), rice chaperones of binding protein 1 and 4 and 5 (BiP1 and BiP4/5), protein disulfide isomerase-like 1–1 and 2–3 (PDIL1-1 and PDIL2-3), and calnexin (CNX) were prepared as described previously (Oono et al. 2010; Wakasa et al. 2011; Wang et al. 2013; Yasuda et al. 2009). For immunodetection, membranes were incubated with rabbit antibodies,

followed by horseradish peroxidase-conjugated anti-rabbit secondary IgG antibodies (dilution 1:5000) (Cell Signaling Technology, Danvers, MA, USA). Bands were visualized using Clarity Western ECL detection kit (Bio-Rad).

Confocal laser scanning microscopy

Developing seeds were harvested 15–20 days after flowering (DAF) from various transgenic lines and then cross-sectioned with a DTK-1000 Microslicer (Dosaka EM) to approximately 150 μm in thickness for indirect immune-histochemical analysis (Yasuda et al. 2006). Mouse anti-GFP monoclonal antibody, rabbit anti-CNX, anti-PDIL1-1, anti-PDIL2-3, anti-BiP1, anti-BiP4/5, anti-TiP3 and anti-RM1 (Cys-rich 14 kDa prolamins) polyclonal antibodies were used to investigate their intracellular localization by Alexa-488 conjugated anti-mouse IgG antibody and Alexa-546 conjugated anti-rabbit IgG antibody. After immunostaining of samples, rhodamine B was used to stain ER-derived protein bodies (PB-Is). Samples cross-sections were spread on glass slides, sealed with a cover glass, and observed with a confocal laser scanning microscope (FLUOVEIW; OLYMPUS, Tokyo, Japan).

RNA extraction and RT-PCR analysis

Total RNA was extracted from leaves and roots as previously described (Takaiwa et al. 1987). RT-PCR analysis was carried out using ReverTraAce qPCR RT Master Mix with gDNA Remover (TOYOBO, Osaka, Japan) with gene-specific primers for GFP and 17S rRNA.

Extraction of recombinant proteins from mature seeds

Leaves and roots were collected from each transgenic rice plant and frozen instantly in liquid nitrogen and then powdered using a Multibeads shaker. Pulverized seed, leaves and roots were treated with 10 volume of various extraction buffer for 20 min by vortex, and then centrifuged at 15,000 rpm for 10 min. The supernatant was transferred to new tube and then equal volume of 2 $\times$ sample buffer (125 mM Tris-HCl (pH6.8), 20% glycerol, 4% SDS) without 2-MER was added. Some aliquot (5 μl) was applied and subjected to electrophoresis on 6 or 8% SDS-PAGE.

Step-wise extraction of seed proteins

Rice seed storage proteins were extracted from pulverized mature seeds (10 mg) containing GluP::GFP:Mald1 using the stepwise extraction method described previously (Takaiwa et al. 2009). Extraction buffer containing 0.5 M NaCl and 10 mM Tris-HCl (pH 7.5) was used to extract

albumin and globulin. The residues were then treated with either 60% (v/v) *n*-propanol for extraction of Cys-poor prolamins or 60% *n*-propanol containing 10 mM DTT for extraction of all prolamins containing Cys-rich prolamins. Finally, residues were treated with 1% (v/v) lactic acid for extraction of glutelins. For all steps, extraction was repeated twice with the same solution by sonication for 2 min on ice. Proteins extracted from seed was evaluated by SDS-PAGE and immunoblotting using anti-GFP, anti-RM4 (Cys-poor 13 kDa prolamins) and anti-RM1 (Cys-rich 14 kDa prolamins) antibodies.

Results

Endosperm-specific expression of Mal d 1 in transgenic rice

When GFP:Mald1 fusion protein was expressed as the secretory protein by ligating an N terminal glutelin signal peptide and a C terminal KDEL ER retention signal under the control of the endosperm-specific 2.3 kb glutelin *GluB-1* promoter, it accumulated at high levels in the endosperm of transgenic rice. As shown in Fig. S1, it was detected as a visible band with the molecular mass of 50 kDa on CBB staining gel (Fig. S1). On the other hand, when the KDEL ER retention signal was removed, accumulation level of the GFP:Mald1(-KDEL) was significantly decreased. Furthermore, when produced in the cytoplasm by the absence of both the signal peptide and KDEL tag, its accumulation level (GFP:Mald1(-sig/-KDEL)) was remarkably reduced to undetectable level (Fig. S1). These results are in remarkable contrast with the expression of the inherent Mal d 1 in apple fruit, since it accumulates at relatively high level (1–20 $\mu\text{g/g}$ FW) irrespective of the absence of the signal peptide (Matthes and Schmitz-Eiberger 2009).

Next, since GFP:Mald1 was expressed as secretory protein in seeds, it was examined whether ER stress was induced by high accumulation of GFP:Mald1. When analyzed by immunoblotting using several anti-chaperon specific antibodies, it was shown that BiPs (BiP1 and BiP4/5) and PDIL2-3 were significantly increased in these transgenic seeds containing the GFP:Mald1 and GFP:Mald1(-KDEL) constructs, indicating that ER stress was induced by their high production (Fig. S1).

Furthermore, subcellular localization of GFP:Mald1 with or without KDEL signal was investigated in maturing transgenic seeds by immuno-histological analysis using confocal laser scanning microscopy. As shown in Fig. 2a, the GFP:Mald1 with the KDEL signal was deposited in a spherical huge ER-derived protein body called as TPC7 body (Wang et al. 2013) with a diameter of 5–10 μm in maturing endosperm cells. By contrast, when the KDEL tag was

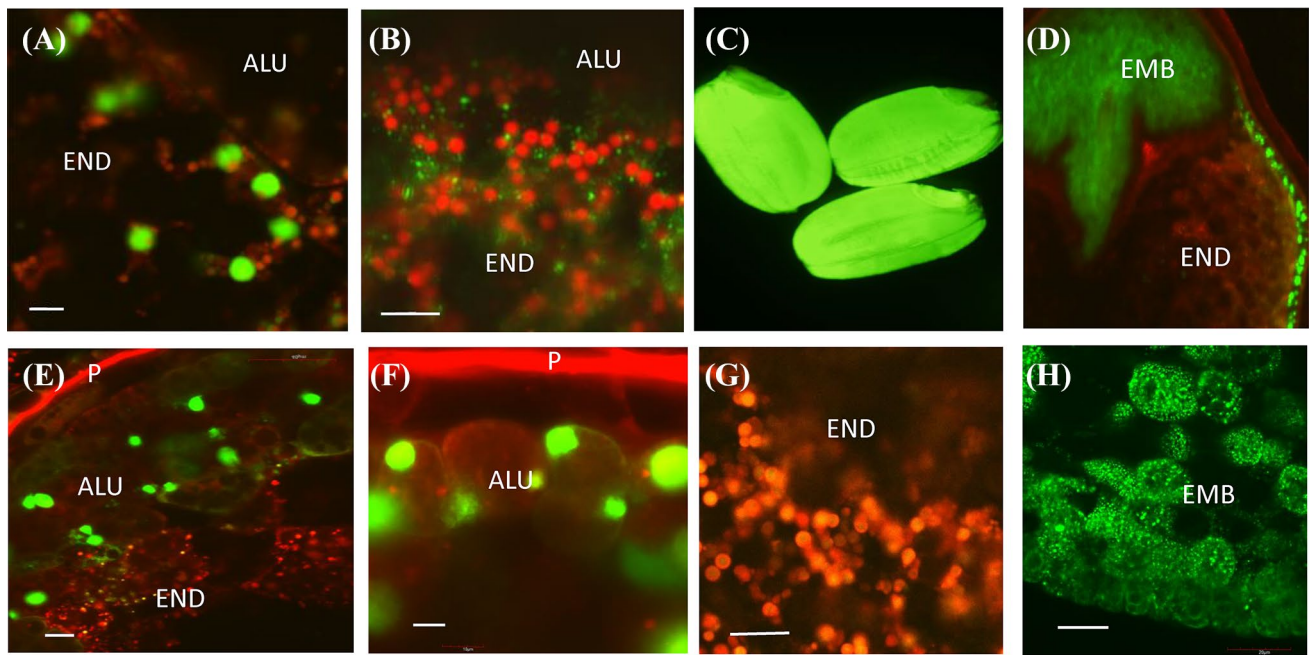


Fig. 2 Expression pattern of GFP:Mald1 in transgenic rice seed. **a** and **b** Intracellular localization of GFP:Mald1 in transgenic rice seed harboring the GluP::GFP:Mal or GluP::GFP:Mal(-KDEL). **c–h** Expression pattern of GFP:Mald1 of transgenic rice seeds har-

boring UbiP::GFP:Mal and its intracellular localization in aleurone layers, starchy endosperm and embryo. ALU aleurone layer, END endosperm, EMB embryo, *p* pericarp thin bars = 5 μ m. Thick bars = 100 μ m

removed, generation of giant PBs was abolished and it was sequestered into small granules within ER lumen (Fig. 2b).

Constitutive expression of Mal d 1, Bet v 1 and TPC7 in transgenic rice

The GFP:Mald1, GFP:Betv1 and GFP:TPC7 attaching the GluB-1 signal peptide and KDEL ER retention signal at their N and C termini were expressed under the control of the maize ubiquitin promoter. The GFP:Mald1 accumulated in mature seeds at higher levels than the GFP:Betv1 and GFP:TPC7 (Fig. 3a). Next, individual dehulled whole grain was divided into the embryo and endosperm and their accumulation levels were examined by immunoblotting using anti-GFP antibody. As shown in Fig. 3c, accumulation levels of the GFP:Betv1 and GFP:TPC7 were under detectable level in most of their transgenic rice embryos, whereas the GFP:Mald1 highly accumulated in its embryos. These results were in contrast with that they accumulated at similar expression levels in their endosperm (Fig. 3b). Thus, difference in accumulation levels of whole grains may be largely accounted for by the expression levels in embryos. On the other hand, when the GFP reporter containing the N terminal GluB-1 signal and C terminal KDEL tag was directly expressed without fusion, it accumulated in both the embryo and endosperm like the GFP:Mald1 construct (Fig. S2). These findings suggested that fusion to the Bet v 1 or

TPC7 may give rise to instability of their fusion products in embryos.

When accumulation levels of the GFP:Mald1 in mature seeds that were directed by the endosperm-specific glutelin *GluB-1* promoter were compared with those that were driven by the ubiquitin promoter, the endosperm-specific promoter led to higher levels of its accumulation (Fig. 3a). This result may reflect the difference in promoter activities in maturing seeds between the *GluB-1* promoter and maize ubiquitin promoter (Qu and Takaiwa 2004).

Next, accumulation levels of the GFP:Mald1, GFP:Betv1 and GFP:TPC7 in leaves and roots were examined by immunoblotting using anti-GFP antibody. It should be noted that these GFP fusion products with the size of 50 kDa is processed (degraded) into 37–39 kDa in vegetative tissues. As shown in Fig. 4a, the GFP:Mald1 highly accumulated in leaves and roots, whereas the accumulation levels of GFP:Betv1 and GFP:TPC7 were very low. Especially, their accumulation levels in roots were much lower than those of the GFP:Mald1. These results suggested that the GFP:Betv1 and GFP:TPC7 products are unstable in vegetative tissues in a similar manner to the expression in their embryos. Therefore, the Mal d 1 may act as stabilizing sequence by protecting fusion proteins from degradation or the Bet v 1 and TPC7 may cause instability in vegetative tissues.

When the transcript (mRNA) levels in leaves and roots between the Mal d 1 and Bet v 1 were examined by RT-PCR,

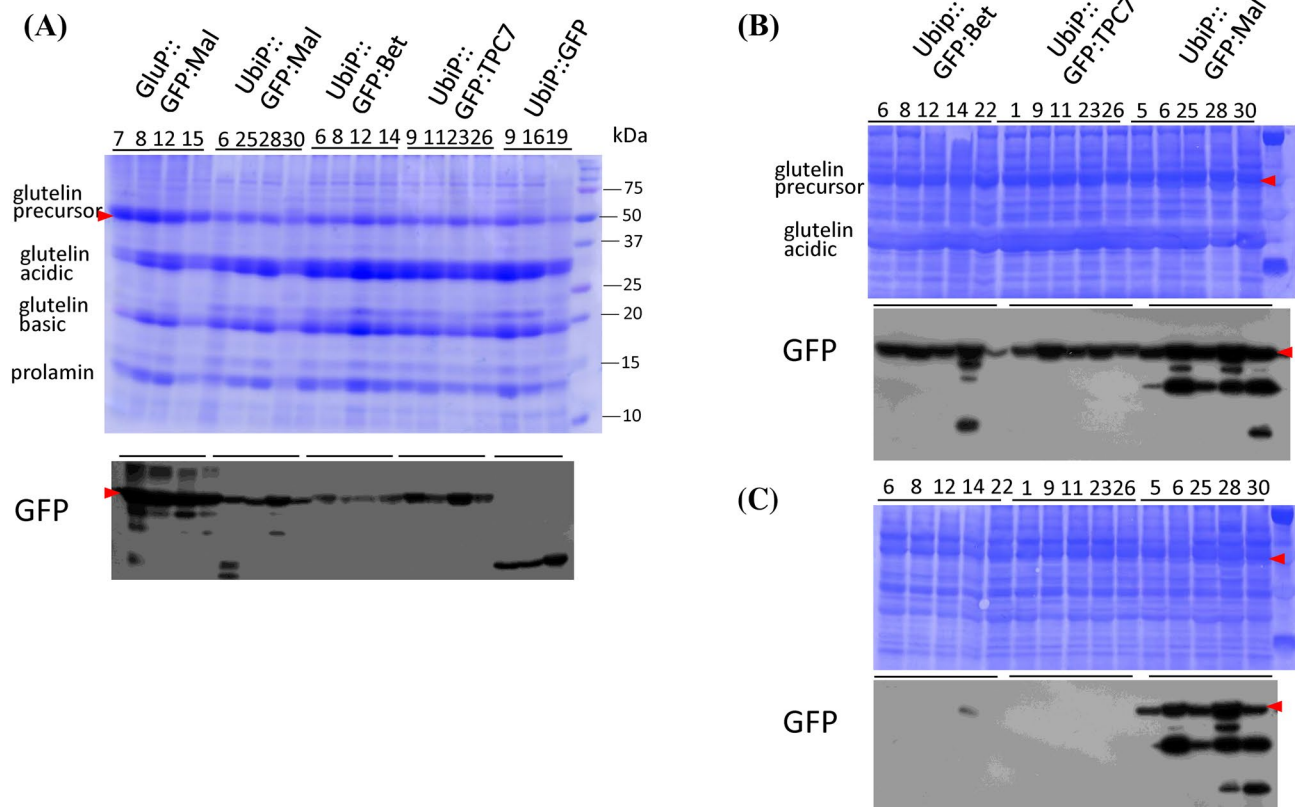


Fig. 3 Immunoblot analysis of GFP:Mald1, GFP:Betv1 and GFP:TPC7 expressed in transgenic rice seeds under the control of the maize ubiquitin promoter. GFP:Mald1 was also expressed under the control of the glutelin *GluB-1* promoter for comparison. Total proteins were extracted from the whole seed (a), endosperm (b) and

embryo (c) with SDS-urea buffer, and subsequently separated by 13% SDS-PAGE. Accumulation levels were examined by immunoblotting with anti-GFP antibody. GFP fusion products with the predicted size of 50 kDa are indicated by arrow head. Degradation products of about 37 kDa and 20 kDa are detected in high expression lines

there was little difference in their transcript levels between them, indicating that their accumulation levels are determined by post-transcriptional level (Fig. 4b).

Subcellular localization of GFP:Mald1

When the GFP:Mald1 was expressed under the control of the maize ubiquitin promoter, GFP fluorescence was detected in the whole grain (Fig. 2c). When this transgenic rice seed was vertically sectioned to examine the localization site of GFP:Mald1 using the confocal laser scanning microscopy, high fluorescence signal was detected in aleurone layer and embryo (Fig. 2d). Then, the subcellular localization of GFP:Mald1 in various parts of transgenic rice seed was examined in detail not only by GFP fluorescence (Fig. 2e–h), but also by immuno-histochemical analysis (Fig. S3). In the aleurone cells, the GFP:Mald1 was shown to be predominantly deposited in huge spherical protein bodies (PBs) with a diameter of $> 5 \mu\text{m}$ (Fig. 2e–f), whereas in the starchy endosperm it was localized in ER-derived PBs called PB-Is containing prolamins and small granules with a size

of $\leq 1 \mu\text{m}$ (Fig. 2f, Fig. S3L). In the embryo of maturing seed, it was predominantly localized in ER lumen (Fig. 2h, Fig. S3H, K). To investigate the participation of chaperon proteins in huge PB formation, intracellular localization of several chaperon proteins was analyzed. As shown in Fig. S3 (A, B, D, I), BiP1, BiP4/5, PDIL1-1 and PDIL2-3 were co-localized in huge peripheral PBs in the aleurone cells, whereas they are also localized in ER lumen and PB-Is in the starchy endosperm. When antibodies against CNX (Fig. S3F), OsTip3 (Fig. S3G) and Cys-rich 14 kDa prolamin (RM1) (Fig. S3L) were used as markers for subcellular compartments of ER, protein storage vacuole (PB-II) and ER-derived PBs (PB-Is), respectively (Takaiwa et al. 2016; Wakasa et al. 2012), they were observed not to be overlapped with huge PBs. Taken together, these results indicate that huge PBs accumulating GFP:Mald1 are formed by aid of BiPs and PDILs chaperon proteins.

Next, when intracellular localization of the GFP:Mald1 was examined in root, it was shown to be mainly localized as small granules in ER lumen (Fig. S4), since CNX, PDIL1-1, BIP1 were co-localized with the GFP:Mald1 at

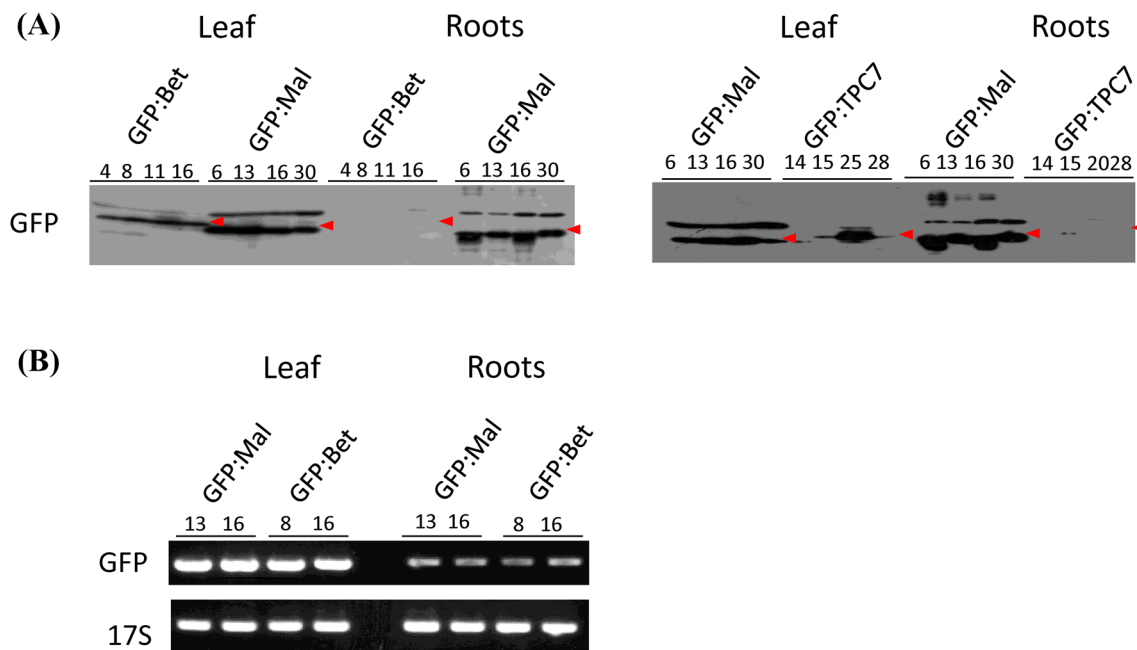


Fig. 4 Expression of GFP:Mald1, GFP:Betv1 and GFP:TPC7 in leaf and roots of transgenic rice lines. **a** Immunoblot analysis of GFP:Mald1, GFP:Betv1 and GFP:TPC7 expressed in leaf and roots of transgenic rice lines under the control of the maize ubiquitin promoter. Processed 37–38 kDa fragments are indicated by arrow head.

b RT-PCR of GFP:Mald1 and GFP:Betv1 transcripts in leaf and roots. Total RNAs were extracted from leaves and roots. The 17S rRNAs were used for normalization of starting RNA concentration in RT-PCR analysis

the peripheral sites of individual vegetative cell. Huge PBs could not be observed in root tissues. Strong fluorescence of GFP was observed at rooting site.

Identification of regions of Mal d 1 responsible for stable accumulation in vegetative tissues

To investigate which domain(s) in the Mal d 1 molecule is involved in high and stable accumulation in vegetative tissues, the Mal d 1 was progressively deleted from the N or C termini and their resultant regions were fused to the GFP reporter as shown in Fig. 5a. They were expressed under the control of the maize ubiquitin constitutive promoter.

When these fusion product levels were examined in vegetative tissues (leaf and roots) by immunoblotting using anti-GFP antibody, the regions of the Mal d 1 between positions 1–120, positions 41–159 and positions 41–120 conferred higher levels of accumulation than the regions between positions 81–120 and positions 81–159 (Fig. 5b). This result indicates that the region between positions 41–120 is critical to confer high level of accumulation in vegetative tissues. Then, to characterize the minimum region required for high expression in vegetative tissues, accumulation levels from the regions between positions 41–107, 41–90 and 41–70 were examined. As shown in Fig. 5c, the region between positions 41–90 led to the highest accumulation

level. Notably, further deletion of 20 amino acids (region between positions 41–70) or addition of 17 amino acids (region between positions 41–107) resulted in reduction of accumulation levels. These results indicate that the minimum region implicated in high accumulation in vegetative tissue is localized between positions 41–90. This region contains ‘P-loop’ (phosphate-binding loop) motif, (GxG-GxGxxK) rich in glycine residue between positions 48–54 that is highly conserved in most of PR10 proteins (Fig. 1). In order to examine its function, this sequence (GNGGPGTIK) was substituted to ANAAPATIK. However, mutation of this sequence did not have any effect on accumulation levels of the linked GFP in most of leaves as well as seeds (Fig. S5). This result indicated that P-loop motif is not responsible for enhancing the accumulation level of the Mal d 1 in transgenic rice. In light of that the amino acid sequence within this region is highly divergent from the Bet v 1 except for the P-loop motif, other amino acid sequences than the P-loop motif may be implicated in controlling the accumulation levels.

Next, it was examined whether the above deletions of the Mal d 1 have an effect on accumulation levels of the linked GFP in transgenic rice seeds (Fig. S6). Progressive deletion from the N terminus reduced the linked GFP accumulation levels (Fig. S6A). The region between positions 41–90 gave rise to higher level accumulation than the regions between

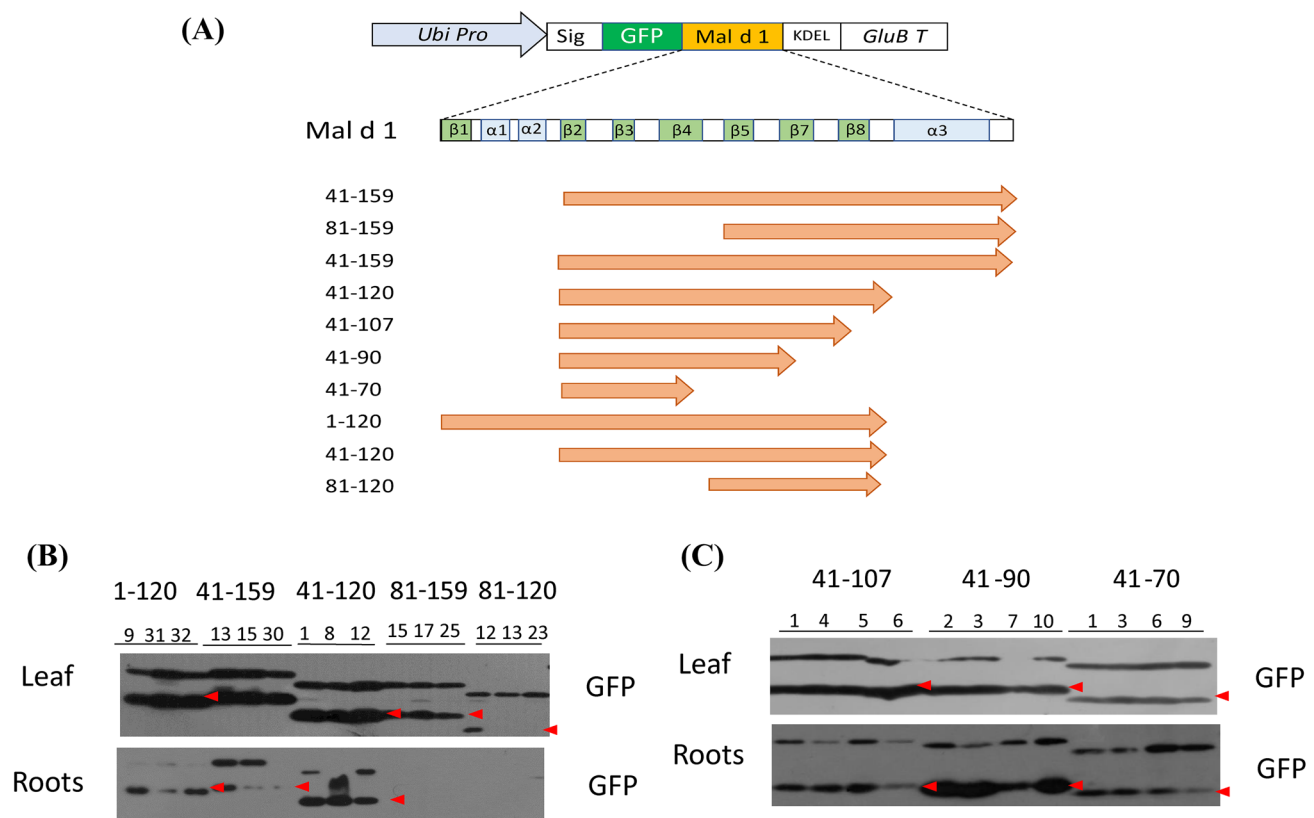


Fig. 5 Deletion analysis of Mal d 1 molecule. **a** Schematic representation of the gene constructs used for deletion analysis of Mal d 1. α , α helical structure; β , β sheet structure. **b**, **c** Effect of deleted Mal d 1 on accumulation levels of linked GFP in leaf and roots of transgenic rice.

positions 41–159, positions 41–120 and positions 41–70 (Fig. S6B). Furthermore, the region between positions 41–120 led to higher levels of accumulation than the regions between positions 81–120 and positions 1–120 (Fig. S6C). These results are fundamentally similar to those observed in vegetative tissues. Taken together, the region between positions 41–90 is involved in stable and high accumulation in seed as well as vegetative tissues.

Domain swapping between the Mal d 1 and Bet v 1

To identify the domain that is implicated in high accumulation in vegetative tissues of transgenic rice harboring GFP:Mal d 1, domain swapping was carried out between the Mal d 1 and Bet v 1 as shown in Fig. 6a. When the C terminal regions of the Bet v 1 between positions 90–160 and positions 127–160 were exchanged with the corresponding regions of the Mal d 1, accumulation levels of these chimeric products (Bet:Mal90, Bet:Mal127) were very low in leaves like the native Bet v 1. That is, accumulation levels of these chimeric products were not significantly altered by fusion with the C terminal region of Mal d 1,

Total proteins extracted from leaves and roots of 4–5 representative independent transgenic lines for each expression construct were subjected to immunoblot analysis using anti-GFP antibody

indicating that the C terminal half region of the Mal d 1 has little enhancing effect. On the other hand, when the N terminal half region of Bet v 1 between positions 1–90 was exchanged with the corresponding region of the Mal d 1, much higher level of this chimeric product (Mal:Bet90) than the native Bet v 1 was observed in leaves. These findings indicated that the N terminal half region of the Mal d 1 is responsible for high accumulation of the Mal d 1 in vegetative tissues. Therefore, in order to ascertain that the region between positions 41–90 is critical for high accumulation in vegetative tissue, the corresponding region was mutually exchanged between the Mal d 1 and Bet v 1. The Bet:Mal:Bet chimeric construct, in which the region between positions 41–90 was substituted with the corresponding region of the Mal d 1, gave rise to high levels of accumulation in leaves of most of transgenic lines, whereas the Mal:Bet:Mal chimeric construct containing the Bet v 1 sequence between positions 41–90 resulted in lower level of accumulation in most of transgenic rice leaves. Furthermore, it is important to note that accumulation levels in seeds are not so affected as in leaves by these domain swapping. These results indicated that the region

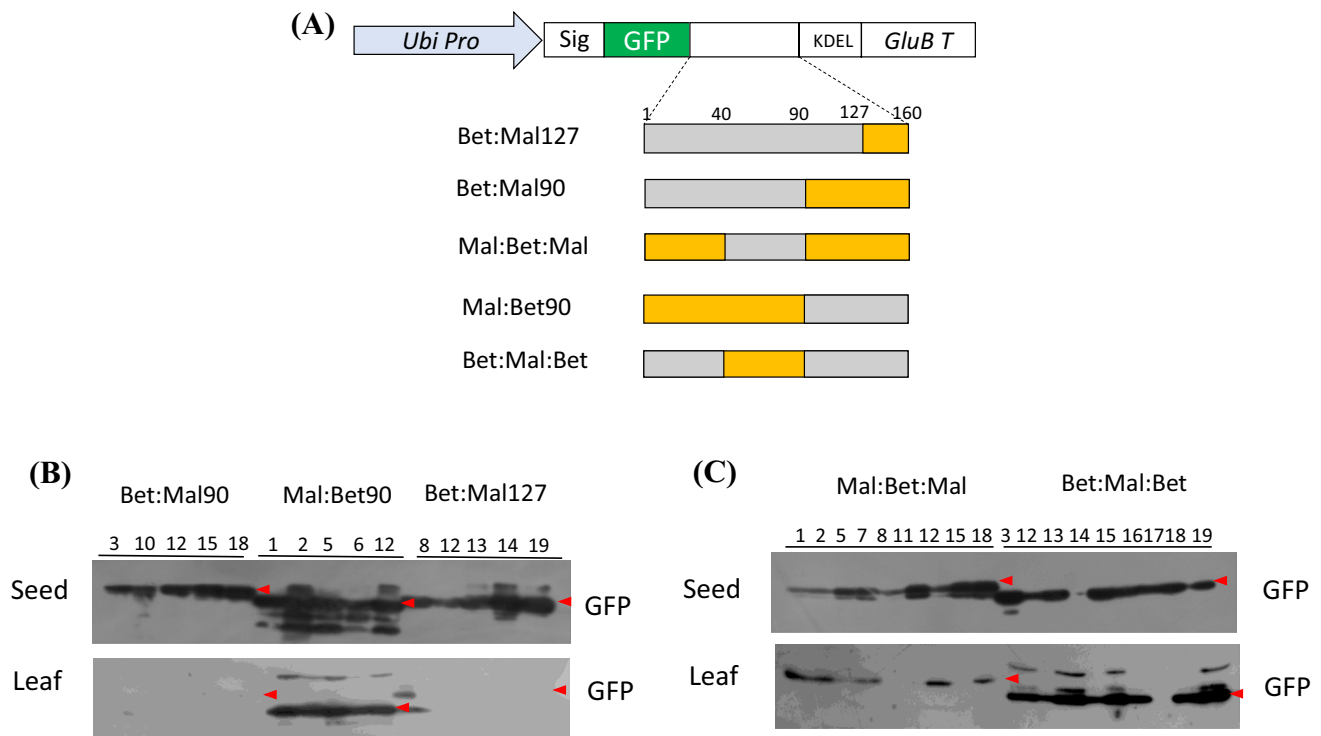


Fig. 6 Domain swapping between the Mal d 1 and Bet v 1. **a** Schematic representation of gene constructs used for domain swapping. **b, c** Immunoblot analysis of various fusion products extracted from

mature seeds and leaves of transgenic rice plants harboring various expression constructs using anti-GFP antibody

between positions 41–90 is mainly responsible for high accumulation in vegetative tissues of Mal d 1.

Aggregation property of Mal d 1

When the GFP:Mald1 was expressed under the control of the endosperm-specific glutelin *GluB-1* promoter, it accumulated in oligomeric form (Fig. 7a) like the TPC7 and Bet v 1 that are deposited into huge PBs in endosperm (Ogo et al. 2014).

Next, to examine whether there is a difference in oligomeric formation of the GFP:Mald1 expressed in seed or leaf that was driven by the glutelin or ubiquitin promoter, it was compared by electrophoresis on 8% SDS-PAGE under non-reducing conditions. As shown in Fig. 7a, dimer of GFP:Mald1 fusion protein was detected in the leaf and seed, when extracted with non-reducing buffer. Furthermore, higher aggregates such as trimer and tetramer were also observed by separation on 6% SDS-PAGE under non-reducing condition (Fig. 7b). It is notable that oligomeric formation such as dimer or trimer was inhibited by treatment with reducing agent before loading (Fig. 7). Therefore, oligomerization was suggested to be formed by self-aggregation via disulfide bond, because there is only one free cysteine (Cys) residue at position 107 in the Mal d 1

like the Bet v 1 and TPC7. Then, the Cys residue in the Mal d 1 was substituted to the Ala residue and then produced as a fusion protein with GFP in transgenic rice. When this GFP:Mald1_{C-A} was extracted from seeds and leaves with 1% SDS and then was separated by electrophoresis on 8% SDS-PAGE, dimeric form of the modified Mal d 1 was observed irrespective of the absence of Cys residue (Fig. 7a lane 2 and 5). However, the migration of this aggregate was slightly different from that of the inherent one.

To examine whether this difference is attributed to the aggregate form, these GFP fusion proteins were extracted from transgenic seeds with different concentration of SDS (0.2%, 0.5% and 1%(w/v)). The GFP:Mald1 could be extracted with all the concentrations of SDS used here, whereas the GFP:Mald1_{C-A} could not be extracted with 0.2% or 0.5% SDS buffer (Fig. 7d). This finding suggested that Cys free modified GFP:Mald1 is aggregated in different manner from the native one.

Extraction of GFP:Mald1 from transgenic rice plants

GFP:Mald1 was extracted from mature seed and leaf of individual transgenic rice plant with various extraction buffers. Different concentrations of SDS (0.2% and 1%), 1% Tween20 and 1% Triton-X100 were used

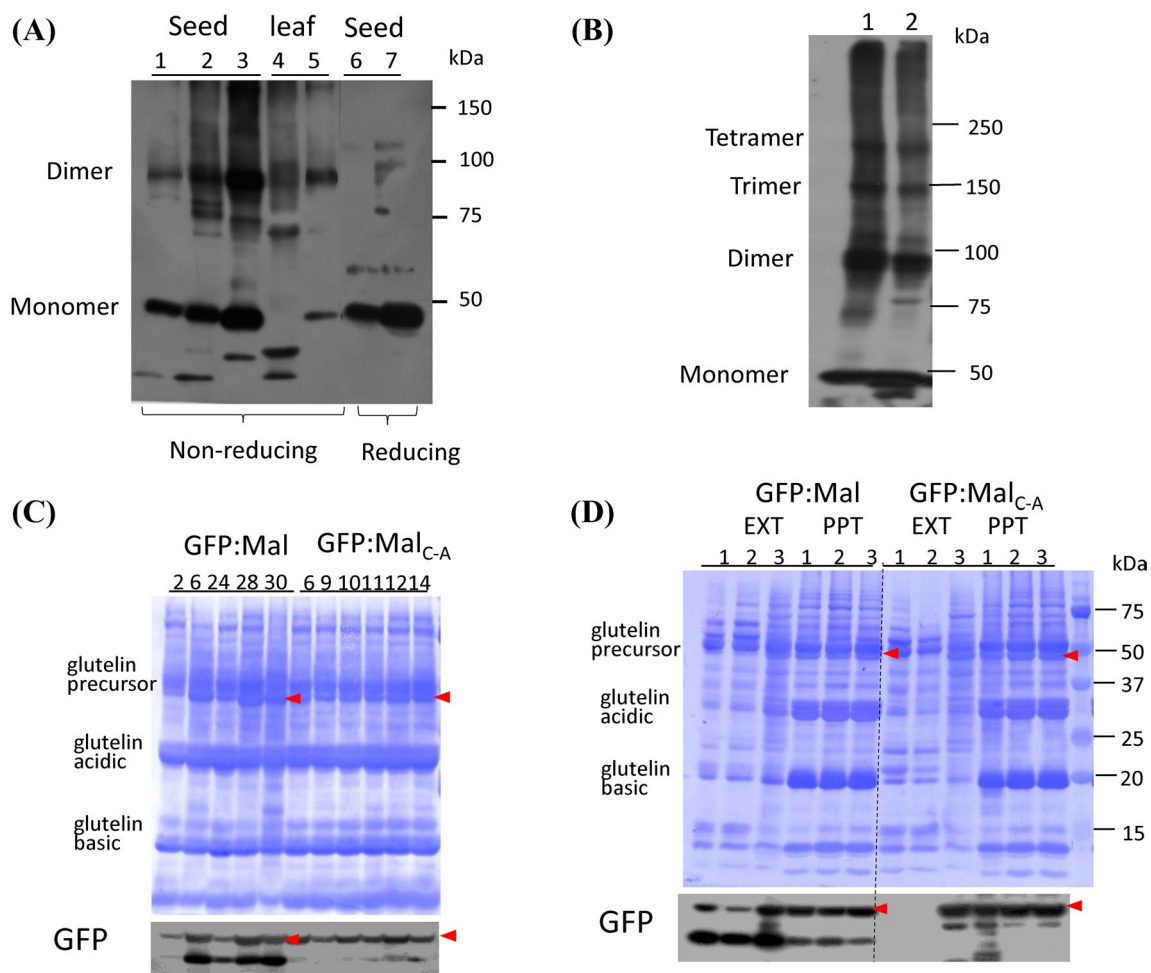


Fig. 7 Comparison of aggregation property between GFP:Mald1 and GFP:Mald1_{C-A}. **a** Immunoblot analysis of GFP:Mald1 and GFP:Mald1_{C-A} expressed in transgenic rice seeds and leaves. They were extracted with 1% SDS in 50 mM Tris-HCl (pH 7.5) and then separated by 8% SDS-PAGE. Seed extracts from transgenic rice harboring GluP::GFP:Mal (1) and transgenic rice harboring UbiP::GFP:Mal_{C-A}, leaf extracts from transgenic rice harboring UbiP::GFP:Mal (3) and transgenic rice harboring UbiP::GFP:Mal_{C-A} (4). Seed extracts of GluP::GFP:Mal (5) and UbiP::GFP:Mal (6) were treated with reducing agent before loading to SDS-PAGE. Aggregates were

detected by immunoblotting using anti-GFP antibody. **b** Separation of seed extracts of transgenic rice harboring GluP::GFP:Mal (1) and UbiP::GFP:Mal (2) by 6% SDS-PAGE. **c** Comparison of accumulation levels in transgenic rice seeds between GFP:Mald1 and GFP:Mald1_{C-A}. **d** Difference in extraction efficiency from mature seeds and leaves of transgenic rice harboring UbiP::GFP:Mal and UbiP::GFP:Mal_{C-A} that were treated with 0.2% SDS (1), 0.5% SDS (2) and 1% SDS (3) in 50 mM Tris-HCl (pH 7.5). EXT, extracted; PPT, not extracted arrow head represents the expressed GFP:Mald1 and GFP:Mald1_{C-A}

for solubilizing membrane proteins. When specifically expressed in the endosperm under the control of the glutelin promoter, 1% SDS was required for its extraction from mature seed. However, this extraction efficiency was much lower than that using the SDS-urea buffer. On the other hand, when expressed under the control of the ubiquitin promoter, it could be easily extracted from seed and leaf with all types of buffer used here. It is notable that extraction efficiency was almost same to the SDS-urea buffer. These results may be related to that the GFP:Mald1 directed by the maize ubiquitin promoter is mainly localized in aleurone layer of seed and is in ER lumen of leaf.

Next, to examine the cause of low extraction efficiency of GFP:Mald1 from transgenic rice seed that was directed by the endosperm-specific glutelin promoter, mature seed powder was treated with various extraction buffers containing reducing agent (10 mM DTT) or various detergents. As shown in Fig. 8b, extraction efficiency of the GFP:Mald1 was improved by the presence of reducing agent. On the other hand, there is little difference in extraction efficiency between GFP:Mald1 and GFP:Mald1(-KDEL) constructs, indicating that the presence of KDEL ER retention signal is not related to the extraction efficiency of GFP:Mald1. Next, when extraction efficiency from mature seed was compared

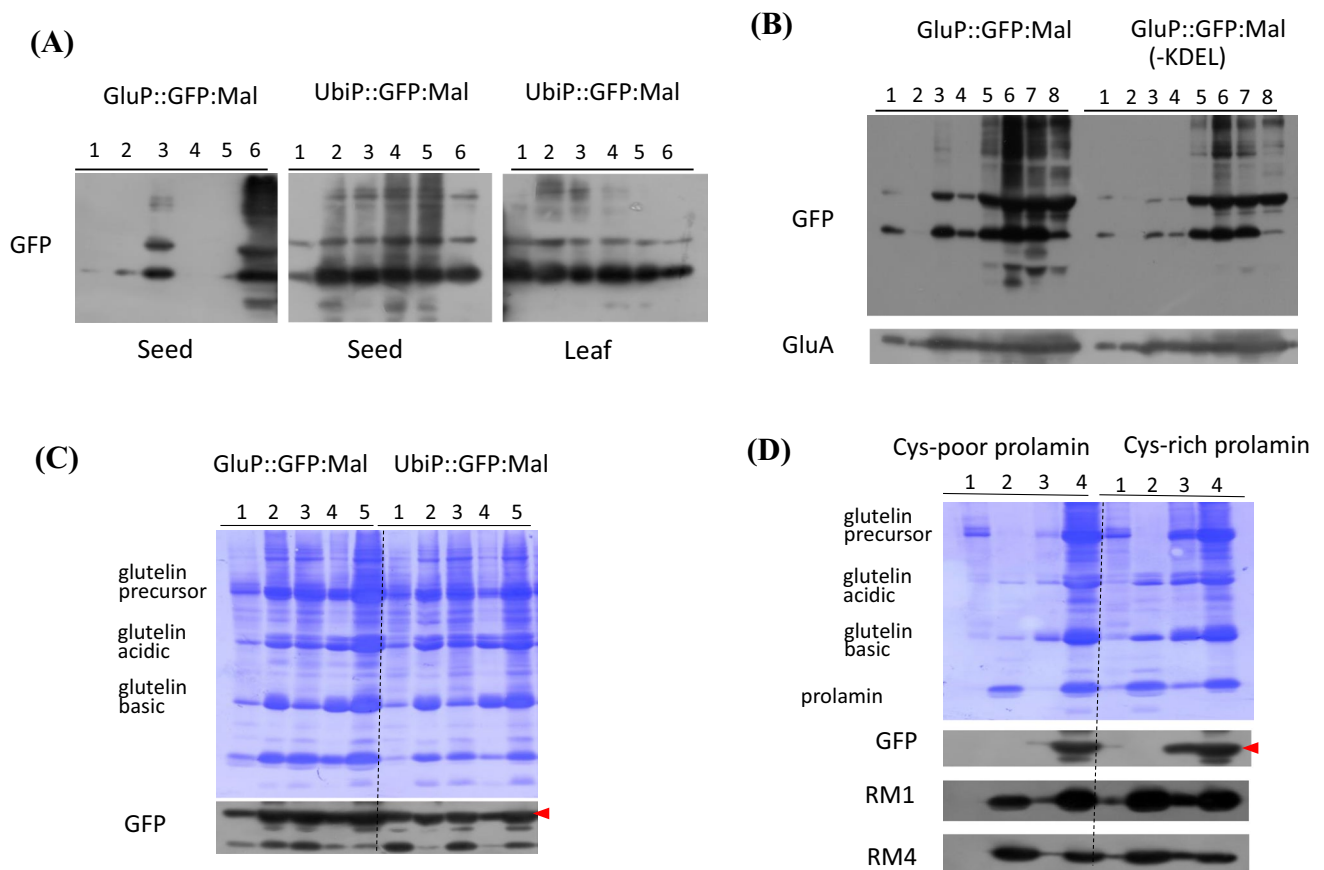


Fig. 8 Extraction of GFP:Mald1 from seed and leaves of transgenic rice. **a** Seed powder and frozen leaf powder of transgenic rice harboring GluP::GFP:Mal or UbiP::GFP:Mal construct were treated with saline buffer (0.5M NaCl in 50 mM Tris-HCl (pH 7.5)) (1) and saline buffer containing 0.2% SDS (2), 1% SDS (3), 1% TritonX-100 (4) and 1% Tween-20 (5) and urea-SDS buffer (6). **b** Seed powder of transgenic rice harboring GluP::GFP:Mal or GluP::GFP:Mal-KDEL was treated with saline buffer containing 1%CTAB (1), 10 mM DTT (2), 1% SDS (3), 0.2% SDS and 10 mM DTT (4), 0.5% SDS and 10 mM DTT (5), 1% SDS and 10 mM DTT (6), 6M urea and 10 mM DTT (7) and urea-SDS buffer (8). Extraction efficiency by these treatments

was examined by immunoblotting using anti-GFP and anti-GluA antibodies. **c** Seed powder of transgenic rice harboring GluP::GFP:Mal or UbiP::GFP:Mal was treated with saline buffer containing 1% SDS (1) or 1% SDS and 10 mM DTT (3) and urea-SDS buffer (5). Residues after treatment with 1% SDS or 1% SDS and 10 mM DTT were re-extracted with urea-SDS buffer (2 and 4). **d** Seed powder of transgenic rice harboring GluP::GFP:Mal was first treated with saline buffer (1), which was followed by extraction with 60% *n*-propanol or 60% *n*-propanol containing 10 mM DTT (2) and then with 1% lactic acid (3). Residues were treated with SDS-urea buffer (4). Arrow head represents the extracted GFP:Mald1

between GluP::GFP:Mald1 and UbiP::GFP:Mald1 constructs, GFP:Mald1 could not be extracted well from the former construct with 1% SDS (less than 20%), whereas almost half of the GFP:Mald1 could be extracted from the latter one (lane 1). Addition of 10 mM DTT to 1% SDS solution resulted in remarkable improvement of its extraction for the former construct (lane 3) (Fig. 8c). These results may be explained by the difference in expressed tissues and its intracellular deposition mode between both constructs. Furthermore, step-wise extraction of GFP:Mald1 from mature seed harboring GluP::GFP:Mald1 showed that it could be only extracted in the glutelin fraction after removal of Cys-rich prolamins with 60% *n*-propanol containing 10 mM DTT (lane 3), but not be extracted by removal of Cys-poor prolamins with 60% *n*-propanol (Fig. 8d). This result indicates

that GFP:Mald1 interacts with Cys-rich prolamins via disulfide bond when expressed in starchy endosperm cells.

Discussion

TPC7 and TPC 9 are derivatives of major birch pollen allergen Bet v 1, which have been developed through in vitro random recombination by means of DNA shuffling as versatile recombinant hypoallergenic allergen against multiple Fagales pollen allergens (Wallner et al. 2007). Since the TPC7 and TPC9 exhibit lower allergenicity and higher immunogenicity than the native Bet v 1, they are expected to act as the ideal tolerogens for allergen-specific immunotherapy against the major pollen allergens of trees belonging

to the Fagales order Bet v 1 family. Therefore, to create rice seed-based oral allergy vaccine against the pollen allergy of Fagales order, we have generated transgenic rice accumulating the Bet v 1 and its derivatives, TPC7 and TCP9. We previously reported that high amounts of TPC7 (540 µg/seed) and native Bet v 1 were deposited in spherical huge protein bodies (10–20 µm) in endosperm cells of transgenic rice seeds, when expressed as secretory protein under the control of the glutelin endosperm-specific promoter (Wang et al. 2013; Ogo et al. 2014).

It has been reported that more than 70% of birch pollen allergy patients develop an IgE-mediated hypersensitivity reaction termed as OAS after apple consumption (Geroldinger-Simic et al. 2011). This is mainly attributed to the immunological cross-reaction between the Bet v 1 specific IgE and apple major allergen Mal d 1, indicating the presence of common conformation IgE epitope(s) between both the allergens. Birch pollen allergen Bet v 1 is specifically expressed in pollen, while food allergen Mal d 1 is mainly expressed in vegetative tissues (fruits). Therefore, it is important to examine how these allergens are deposited in various tissues of transgenic rice, when these allergens are planned to be produced in transgenic rice and then to be used as rice-based allergy vaccines against these pollen and food allergy diseases.

As shown in Fig. 3, when the GFP:Betv1, GFP:TPC7 and GFP:Mald1 were expressed as secretory protein by ligating an N terminal signal peptide and a C terminal KDEL tag under the control of the maize ubiquitin promoter, accumulation levels of GFP:Betv1 and GFP:TPC7 were very low in various vegetative tissues such as leaf, root and embryo except for seed (endosperm). In contrast, the GFP:Mald1 highly accumulated in various tissues. When intracellular localization of the GFP:Mald1 was examined in maturing transgenic seeds, it was mainly deposited into novel huge PBs with a diameter of 5–10 µm in aleurone cells, whereas in starch endosperm cells and embryo it was sequestered into ER-derived PBs (PB-Is) with a diameter of 1 µm containing prolamin storage proteins and small ER-derived particles, respectively. This accumulation pattern was in remarkable contrast with the result that directed by rice endosperm-specific glutelin promoter, since the GFP:Mald1 is predominantly deposited in huge PBs with a size of 5–10 µm in the starchy endosperm cells like the Bet v 1 and TPC7 (Ogo et al. 2014; Wang et al. 2013). On the other hand, such huge PBs were not observed in vegetative tissues, although high amounts of GFP:Mald1 accumulate in vegetative tissues. The GFP:Mald1 is deposited into ER-derived small granules in vegetative tissues and embryo, which are very similar to those observed in production of rice prolamins (GFP:prolamin) expressed in vegetative tissues and calli of transgenic rice (Saito et al. 2009; Shigemitsu et al. 2013). Then, taking these results into consideration, we first

examined why the GFP:Mald1 is specifically deposited into huge PBs in aleurone cells, when directed by the maize ubiquitin promoter.

It has been suggested that huge PB formation is associated with the aggregate property of the recombinant proteins (Ogo et al. 2014). Furthermore, expression levels may also be important for the formation of huge PBs, since maize ubiquitin promoter directs higher expression in aleurone layer than in starchy endosperm of maturing seed, when investigated by the expression of GUS reporter gene (Takaiwa et al. 2007). Especially, requirement for high accumulation of recombinant proteins above a threshold level has been pointed out for the formation of PB in elastin-like polypeptide (ELPs) and hydrophobin fusion proteins (Gutierrez et al. 2013; Saberianfar et al. 2015).

PB formation is suggested to be efficiently initiated by ER retention via the C terminal KDEL signal of GFP:Mald1. Actually, localization of GFP:Mald1 in ER lumen was observed in various tissues including aleurone layer and seed embryo, since small granules with green fluorescence were observed along ER membrane network all over the whole cells, which were connected as ER network. That is, GFP:Mald1 is synthesized on rough ER, and subsequently transferred into the ER lumen. The size of PB gradually increases up to 10 µm through binding of small aggregates. As shown in Fig. 9, single or a few huge spherical PBs are finally formed by aggregating several small GFP:Mald1 particles. Aggregates are observed to be concentrated into huge PB through ER network connection. Finally, a single huge spherical PB is formed in many aleurone cells. Growth in PB size is accompanied by an increase in accumulation levels. Furthermore, PB formation is known to be mediated by aid of some chaperon proteins such as BiPs and PDILs that are implicated in folding and assembly. When subcellular localization of some BiPs and PDILs were examined by immuno-histochemical analysis, these chaperons are localized within huge PBs containing GFP:Mald1 as shown in Fig. S5, indicating that huge PBs in aleurone cells are formed in a similar manner to the huge PBs observed in the starchy endosperm (Ogo et al. 2014). Aleurone cells are filled with lipid bodies and protein storage vacuoles (PSVs) referred as aleurone particles. The 52 and 63 kD globulin and phytic acid are accumulated inside aleurone particles (Kurokawa et al. 2014). Antibodies against OsTiP3 and CNX as PSV and ER marker (Takaiwa et al. 2016; Wakasa et al. 2012), respectively, are not localized to huge PBs in the aleurone cells. Huge PBs are suggested to be independently formed by self-aggregation of Mal d 1. Generation mechanism of this huge PB is fundamentally same to that of Zera, ELPs and hydrophobin-I in plant cells (Conley et al. 2011; Saberianfar and Menassa 2017).

Bet v 1 and Mal d 1 belong to the same PR10 group and share about 55–68% amino acid sequence homology to each

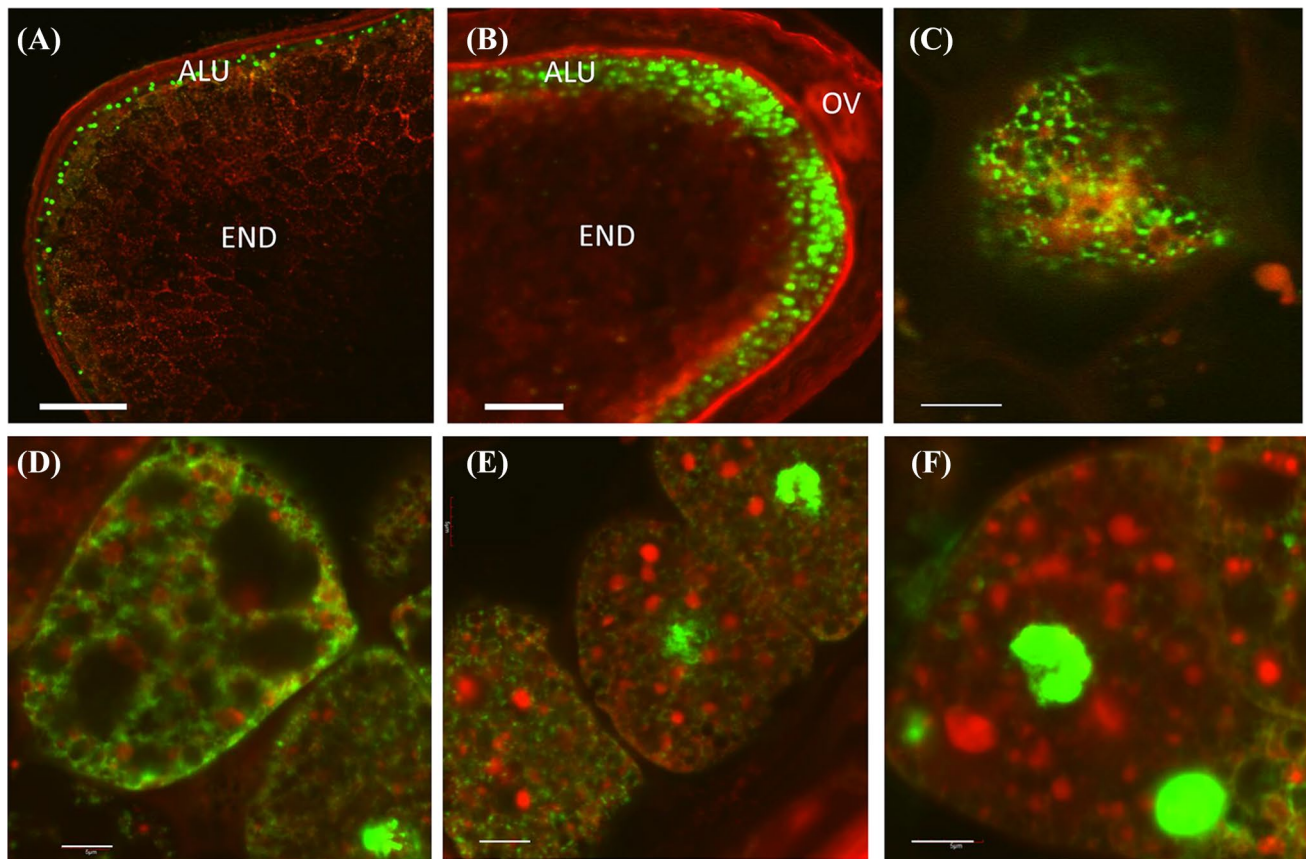


Fig. 9 Subcellular localization of GFP:Mald1 in aleurone cells. Green fluorescence by GFP represents the localization of GFP:Mald1. Red represents the rhodamine B stained particles. **a** Ventral site of sectioned seed. **b** Dorsal site of sectioned seed. **c–f** Huge PB for-

mation process of GFP:Mald1 in aleurone layer. *ALU* aleurone layer, *END* endosperm, *EMB* embryo, *OV* ovular vascular trace. Thick bars = 200 μ m, thin bars = 5 μ m

other. However, it is notable that their accumulation levels in various vegetative tissues are quite different, even although expressed under the control of the same constitutive ubiquitin promoter. The GFP:Mald1 highly accumulates in various tissues, whereas accumulation level of the GFP:Betv1 was reduced to low level in vegetative tissues except for seed (endosperm). When transcript levels in leaves and roots were examined between them by RT-PCR using GFP primers, they were almost identical. This result suggests that post-translational regulation is involved in such difference in accumulation levels. Non-correlation between transcript levels and accumulation levels of their translation products has been reported for several seed storage proteins in maize aleurone cells (Reyes et al. 2011). Furthermore, it has been demonstrated that foreign recombinant protein such as 7Crp peptide selectively accumulated in the endosperm rather than other tissues (embryo and vegetative tissues) even though expressed under the constitutive promoter (Takaiwa et al. 2007). In light of that accumulation levels of Bet v 1 and TCP7 are low in vegetative tissues except for endosperm, there is a possibility that they are more

sensitive to protease(s) than Mal d 1. This is also related to the evidences that proteolytic enzymes are low abundance in endosperm tissue compared to vegetative tissues (Benchabane et al. 2008). Difference in intracellular localization or post-translational modification such as N-glycosylation may be also implicated in such post-translational regulation.

In this study, we first demonstrated by deletion analysis of the Mal d 1 that the specific region between positions 41–90 is mainly responsible for high and stable accumulation in vegetative tissues. Participation of this region in high accumulation in vegetative tissues was further confirmed by domain swapping between the Mal d 1 and Bet v 1. As shown in Fig. 6, the exchange of this Mal d 1 region between positions 41–90 with the corresponding region of Bet v 1 (Bet:Mal:Bet) resulted in higher accumulation of the linked GFP in vegetative tissue than the native Bet v 1. This region includes the glycine rich P element motif highly conserved in PR10 proteins and N-glycosylation site in the Bet v 1 and TPC7. Furthermore, it is important to note that this region is localized on molecular surface of the tertiary structure, which may be sensitive to proteolysis.

It has been reported that the three dimensional structure (tertiary structure) of various PR10 proteins is very similar irrespective of divergence in their primary structures (40–68%). It is composed of a seven-stranded antiparallel β -sheet and three α -helical that form a large internal cavity. Such tertiary structure is similar on the outside but displays considerable differences inside due to variations in cavities arrangement. The region between positions 41–90 exhibits anti parallel β -2 (Gln41-Glu46), β -3 (Ile54-Thr57), β -4 (Thy67-Ile75) and β -5 (Ser81-Gly89) structure and constitutes protein surface of the Mal d 1 (Ahammer et al. 2017). Some amino acids in this region are suggested to be implicated in cross reactivity between the Bet v 1 and Mal d 1, since this region (positions 41–90) almost overlaps with the conformation IgE epitope of Bet v 1 located at the segment between positions 43 and 53 and Arg71, Asp73, His77, Ile87 and Lyn98, which was defined by the binding of monoclonal murine IgG1 BV16 to Bet v 1 (Holm et al. 2011; Mirza et al. 2000). Especially, mutation from Glu to Ser at position 46 in the Bet v 1 significantly decreased the IgE-binding capacity, indicating that Glu46 acts as crucial residue for IgE recognition. Actually, this mutation was suggested to cause the structural change of loop L4 containing the P-loop (Spangfort et al. 2003). On the other hand, five-point mutation of Mal d 1 at Thr11, Ile3, Thr57, Thr112 and Ile113 was reported to lead to marked reduction of binding capacity of Mal d 1 specific IgE (Ma et al. 2006; Holm et al. 2011). Furthermore, mutation of Ser112 resulted in reduction of Mal d 1 specific IgE binding activity. Many of these amino acids involved in Mal d 1 specific IgE binding are not identical with the BV16 epitope, suggesting that there may be difference in surface conformational structure responsible for IgE binding. Especially, it is notable that four amino acids of 16 amino acids in the BV16 B cell epitope were different and other residues except for the BV16 epitope within this region are also highly divergent. These results suggest that the region between positions 41–90 may be structurally and functionally different between the Mal d 1 and Bet v 1. Therefore, further study will be required to examine what structural difference is related to the variation in accumulation level between Mal d 1 and Bet v 1.

When the GFP:Mald1 was specifically expressed under the control of the endosperm-specific promoter, it was difficult to extract from mature seeds with various buffers without reducing agent. This may be explained by the findings that GFP:Mald1 is targeted into ER-derived huge PBs and into PB-Is through the interaction with Cys-rich prolamins via disulfide bond (Fig. 8) after transportation into ER lumen of the endosperm cells, since they are expressed in the same endosperm cells. Therefore, higher concentration of detergent (1–2% SDS) and reducing agent (DTT) are required for efficient extraction of GFP:Mald1 from transgenic rice seed. By contrast, when expressed under the control of

the constitutive promoter, the GFP:Mald1 could be easily extracted from various tissues with several detergents without reducing agent. Notably, it could be extracted from transgenic mature seeds even with saline solution. This may be explained by the observation that GFP:Mald1 is predominantly expressed in aleurone layer and embryo of maturing seed and are deposited into huge ER-derived PBs and small granules in ER, respectively, by self-aggregation without interaction with Cys-rich prolamins.

Bet v 1 has been reported to generally form a mixture of monomers, dimers and higher order oligomers (Scholl et al. 2005). Dimerization was also reported for the Mal d 1 (Ma et al. 2006; Roulias et al. 2014; Rouvinen et al. 2010). When GFP:Mald1 was produced in transgenic rice, it displays multimeric formation in seed and leaf (Fig. 7). The Mal d 1 has a Cys residue like the Bet v 1. This Cys residue is suggested to be implicated in disulfide bond formation for self-aggregation (Cys-mediated dimerization) in various tissues or interaction with the Cys-rich prolamins in the endosperm of transgenic rice seeds. However, even though this Cys residue was substituted to Ala by mutagenesis, it was observed in transgenic rice that oligomeric (dimeric) formation was retained. Therefore, such oligomeric formation is suggested to be achieved by hydrogen bond and van der Waals mediated interaction. Notably, this mutation resulted in the alternation in physicochemical property of the accumulated Mal d 1, since it became difficult to extract the mutagenized Mal d 1 from vegetative tissues. That is, the mutated Mal d 1 lacking Cys residue was suggested to be falsely folded, leading to change in physicochemical property of Mal d 1. Therefore, Cys residue in Mal d 1 may be critical for the correct folding or oligomeric formation via disulfide bond.

It has been demonstrated in this study that localization and generation of such huge PBs in transgenic rice are dependent on the tissue specificity and strength of the promoter used for expression. When directed by the maize ubiquitin constitutive promoter, huge PBs were observed to be formed mainly in aleurone cells, whereas expression by the endosperm-specific glutelin promoter gave rise to their production only in starchy endosperm tissue. Higher expression above a threshold level is essentially required for huge PB formation. However, although various foreign recombinant proteins have been produced in our laboratory, such huge PBs could not be formed except for the PR10 proteins such as Bet v 1 and Mal d 1 even when the same strong endosperm-specific promoter has been utilized to drive expression of various foreign genes, suggesting that the specific physicochemical property such as aggregation property may be associated with the huge PB formation. Their tertiary structure leading to high aggregates may be responsible for the huge PB formation by helping to increase the accumulation levels. When the GFP was produced by

fusion with Mal d 1, it was highly and stably accumulated not only in seeds, but also in leaves and roots (Fig. S7). Notably, accumulation level of GFP has been significantly increased by fusion with Mal d 1. Furthermore, GFP:Mal d 1 fusion protein can be more easily extracted from various tissues of transgenic rice than that specifically expressed in transgenic rice seeds. Therefore, Mal d 1 may be utilized as fusion partner for the production of high-value recombinant proteins in plant like Zera and ELPs. Further works will be required whether production of foreign recombinant proteins can be enhanced by fusion with Mal d 1.

Acknowledgements We thank Ms. K. Miyashita, Y. Ikemoto and H. Yajima for technical assistance and Dr. Kenjiro Ozawa for encourage of this research.

Author contributions FT carried out the experimental design, data analysis and wrote the manuscript. YO and YW participated in experiments and contributed the data analysis.

References

- Agarwall P, Agarwal PK (2014) Pathogenesis related-10 proteins are small, structurally similar but with diverse role in stress signaling. *Mol Biol Rep* 41:599–611
- Ahammer L, Grutsch S, Kamenik AS, Liedl KR, Tollinger M (2017) Structure of the major apple allergen Mal d 1. *J Agric Food Chem* 65:1606–1612
- Benchabane M, Goulet C, Rivard D, Faye L, Gomord V, Michaud D (2008) Preventing unintended proteolysis in plant protein biofactories. *Plant Biotechnol J* 6:633–648
- Beuning L, Bowen J, Persson H, Barraclough D, Bulley S, Macrae E (2004) Characterisation of Mal d 1-related genes in *Malus*. *Plant Mol Biol* 55:369–388
- Conley AJ, Joensuu JJ, Richman A, Menassa R (2011) Protein body-inducing fusions for high-level production and purification of recombinant proteins in plants. *Plant Biotechnol J* 9:419–433
- Ebner C, Birkner T, Valenta R, Rumpold H, Breitenbach M, Scheiner O, Kraft D (1991) Common epitopes of birch pollen and apples—studies by western and northern blot. *J Allergy Clin Immunol* 88:588–594
- Egelkroun E, Rajan V, Howard JA (2012) Overproduction of recombinant proteins in plants. *Plant Sci* 184:83–101
- Fernandes H, Michalska K, Sikorski M, Jaskolski M (2013) Structural and functional aspects of PR-10 proteins. *FEBS J* 280:1169–1199
- Fritsch R, Bohle B, Vollmann U, Wiedermann U, Jahn-Schmid B, Krebitz M, Breiteneder H, Kraft D, Ebner C (1998) Bet v 1, the major birch pollen allergen, and Mal d 1, the major apple allergen, cross-react at the level of allergen-specific T helper cells. *J Allergy Clin Immunol* 102:679–686
- Geroldinger-Simic M, Zelniker T, Aberer W, Ebner C, Egger C, Greiderer A, Prem N, Lidholm J, Ballmer-Weber BK, Vieths S, Bohle B (2011) Birch pollen-related food allergy: clinical aspects and the role of allergen-specific IgE and IgG4 antibodies. *J Allergy Clin Immunol* 127:616–622
- Goto F, Yoshihara T, Sugimoto N, Toki S, Takaiwa F (1999) Iron fortification of rice seed by the soybean ferritin gene. *Nature Biotechnol* 17:282–286
- Gutiérrez SP, Saberianfar R, Kohalmi SE, Menassa R (2013) Protein body formation in stable transgenic tobacco expressing elastin-like polypeptide and hydrophobin fusion proteins. *BMC Biotechnol* 13:40
- Hofbauer A, Stoger E (2013) Subcellular accumulation and modification of pharmaceutical proteins in different plant tissues. *Curr Pharm Des* 19:5495–5502
- Holm J, Ferreras M, Ipsen H, Wurtzen PA, Gajhede M, Larsen JN, Lund K, Spangfort MD (2011) Epitope grafting, re-creating a conformational Bet v 1 antibody epitope on the surface of the homologous apple allergen Mal d 1. *J Biol Chem* 286:17569–17578
- Kurokawa S, Kuroda M, Mejima M, Nakamura R, Takahashi Y, Sagara H, Takeyama N, Satoh S, Kiyono H, Teshima R, Masumura T, Yuki Y (2014) RNAi-mediated suppression of endogenous storage proteins leads to a change in localization of overexpressed cholera toxin B-subunit and the allergen protein RAG2 in rice seeds. *Plant Cell Rep* 33:75–87
- Liu J, Ekramoddoullah AKM (2006) The family 10 of plant pathogenesis-related proteins: their structure, regulation, and function in response to biotic and abiotic stresses. *Physiol Mol Plant Pathol* 68:3–13
- Ma Y, Gadermaier G, Bohle B, Bolhaar S, Knulst A, Markovic-Housley Z, Breiteneder H, Briza P, Hoffmann-Sommergruber K, Ferreira F (2006) Mutational analysis of amino acid positions crucial for IgE-binding epitopes of the major apple (*Malus domestica*) allergen, Mal d 1. *Int Arch Allergy Immunol* 139:53–62
- Mari A, Ballmer-Weber BK, Vieths S (2005) The oral allergy syndrome: improved diagnostic and treatment methods. *Curr Opin Allergy Clin Immunol* 5:267–273
- Marzban G, Puehringer H, Dey R, Brynda S, Ma Y, Martinelli A, Zaccarini M, van der Weg E, Housley Z, Kolarich D, Altmann F, Laimer M (2005) Localisation and distribution of the major allergens in apple fruits. *Plant Sci* 169:387–394
- Matthes A, Schmitz-Eiberger M (2009) Apple (*Malus domestica* L. Borkh.) allergen Mal d 1: effect of cultivar, cultivation system, and storage conditions. *J Agric Food Chem* 57:10548–10553
- Mirza O, Henriksen A, Ipsen H, Larsen JN, Wissenbach M, Spangfort MD, Gajhede M (2000) Dominant epitopes and allergic cross-reactivity: complex formation between a Fab fragment of a monoclonal murine IgG antibody and the major allergen from birch pollen Bet v 1. *J Immunol* 165:331–338
- Ogo Y, Takahashi H, Wang S, Takaiwa F (2014) Generation mechanism of novel, huge protein bodies containing wild type or hypoallergenic derivatives of birch pollen allergen Bet v 1 in rice endosperm. *Plant Mol Biol* 86:111–123
- Oono Y, Wakasa Y, Hirose S, Yang L, Sakuta C, Takaiwa F (2010) Analysis of ER stress in developing rice endosperm accumulating β -amyloid peptide. *Plant Biotechnol J* 8:691–718
- Pagliarini G, Paris R, Arens P, Tartarini S, Ricci G, Smulders MM, van der Weg WE (2013) A qRT-PCR assay for the expression of all Mal d 1 isoallergen genes. *BMC Plant Biol* 13:51
- Puehringer H, Moll D, Hoffmann-Sommergruber K, Watillon B, Kattinger H, Machado MLD (2000) The promoter of an apple Ypr10 gene, encoding the major allergen Mal d 1, is stress- and pathogen-inducible. *Plant Sci* 152:35–50
- Qu LQ, 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
- Reyes FC, Chung T, Holding D, Jung R, Vierstra R, Otegui MS (2011) Delivery of prolamins to the protein storage vacuole in maize aleurone cells. *Plant Cell* 23:769–784
- Roulias A, Pichler U, Hauser M, Himly M, Hofer H, Lackner P, Ebner C, Briza P, Bohle B, Egger M, Wallner M, Ferreira F (2014) Differences in the intrinsic immunogenicity and allergenicity of Bet v 1 and related food allergens revealed by site-directed mutagenesis. *Allergy* 69:208–215

- Rouvinen J, Jänis J, Laukkanen ML, Jylhä S, Niemi M, Päivinen T, Mäkinen-Kiljunen S, Haahtela T, Söderlund H, Takkinen K (2010) Transient dimers of allergens. *PLoS ONE* 5:e9037
- Saberianfar R, Menassa R (2017) Protein bodies: how the ER deals with high accumulation of recombinant proteins. *Plant Biotechnol J* 15:671–673
- Saberianfar R, Joensuu JJ, Conley AJ, Menassa R (2015) Protein body formation in leaves of *Nicotiana benthamiana*: a concentration-dependent mechanism influenced by the presence of fusion tags. *Plant Biotechnol J* 13:927–937
- Saito Y, Kishida K, Takata K, Takahashi H, Shimada T, Tanaka K, Morita S, Satoh S, Masumura T (2009) A green fluorescent protein fused to rice prolamin forms protein body-like structures in transgenic rice. *J Exp Bot* 60:615–627
- Saraste M, Sibbald PR, Wittinghofer A (1990) The P-loop—a common motif in ATP- and GTP-binding proteins. *Trends Biochem Sci* 15:430–434
- Scholl I, Kalkura N, Shedziankova Y, Bergmann A, Verdino P, Knittelfelder R, Kopp T, Hantusch B, Betzel C, Dierks K, Scheiner O, Boltz-Nitulescu G, Keller W, Jensen-Jarolim E (2005) Dimerization of the major birch pollen allergen Bet v 1 is important for its in vivo IgE-cross-linking potential in mice. *J Immunol* 175:6645–6650
- Shigemitsu T, Masumura T, Morita S, Satoh S (2013) Accumulation of rice prolamin-GFP fusion proteins induces ER-derived protein bodies in transgenic rice calli. *Plant Cell Rep* 32:389–399
- Spangfort MD, Mirza O, Ipsen H, Van Neerven RJ, Gajhede M, Larsen JN (2003) Dominating IgE-binding epitope of Bet v 1, the major allergen of birch pollen, characterized by X-ray crystallography and site-directed mutagenesis. *J Immunol* 171:3084–3090
- Stoger E, Ma JK, Fischer R, Christou P (2005) Sowing the seeds of success: pharmaceutical proteins from plants. *Curr Opin Biotechnol* 16:167–173
- Tada Y, Utsumi S, Takaiwa F (2003) Foreign gene products can be enhanced by introduction into storage protein mutants. *Plant Biotechnol J* 1:411–422
- Takaiwa F (2013) Update on the use of transgenic rice seeds in oral immunotherapy. *Immunotherapy* 5:301–312
- Takaiwa F, Kikuchi S, 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, Wakasa Y (2007) Endosperm tissue is good production platform for artificial recombinant proteins in transgenic rice. *Plant Biotechnol J* 5:84–92
- Takaiwa F, Hirose S, Takagi H, Yang L, 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, Wakasa Y, Takagi H, Hiroi T (2015) Rice seed for delivery of vaccines to gut mucosal immune tissues. *Plant Biotechnol J* 13:1041–1055
- Takaiwa F, Yang L, Maruyama N, Wakasa Y, Ozawa K (2016) Deposition mode of transforming growth factor- β expressed in transgenic rice seed. *Plant Cell Rep* 35:2461–2473
- Takaiwa F, Wakasa Y, Hayashi S, Kawakatsu T (2017) An overview on the strategies to exploit rice endosperm as production platform for biopharmaceuticals. *Plant Sci* 263:201–209.
- Wakasa Y, Yasuda H, Takaiwa F (2006) High accumulation of bioactive peptide in transgenic rice seeds by expression of introduced multiple genes. *Plant Biotechnol J* 4:499–510
- Wakasa Y, Yasuda H, Oono Y, Kawakatsu T, Hirose S, Takahashi H, Hayashi S, Yang L, Takaiwa F (2011) Expression of ER quality control-related genes in response to changes in BiP1 levels in developing rice endosperm. *Plant J* 65:675–689
- Wakasa Y, Hayashi S, Takaiwa F (2012) Expression of OsBiP4 and OsBiP5 is highly correlated with the endoplasmic reticulum stress response in rice. *Planta* 236:1519–1527
- Wallner M, Stöcklinger A, Thalhamer T, Bohle B, Vogel L, Briza P, Breiteneder H, Vieths S, Hartl A, Mari A, Ebner C, Lackner P, Hammerl P, Thalhamer J, Ferreira F (2007) Allergy multivaccines created by DNA shuffling of tree pollen allergens. *J Allergy Clin Immunol* 120:374–380
- Wang S, Takahashi H, Kajiura H, Kawakatsu T, Fujiyama K, Takaiwa F (2013) Transgenic rice seeds accumulating recombinant hypoallergenic birch pollen allergen Bet v 1 generate giant protein bodies. *Plant Cell Physiol* 54:917–933
- Yasuda H, Hayashi Y, Jomori T, Takaiwa F (2006) The correlation between expression and localization of a foreign gene product in rice endosperm. *Plant Cell Physiol* 47:756–763
- Yasuda H, Hirose S, Kawakatsu T, Wakasa Y, Takiwa F (2009) Overexpression of BiP has inhibitory effects on the accumulation of seed storage proteins in endosperm cells of rice. *Plant Cell Physiol* 50:1532–1543

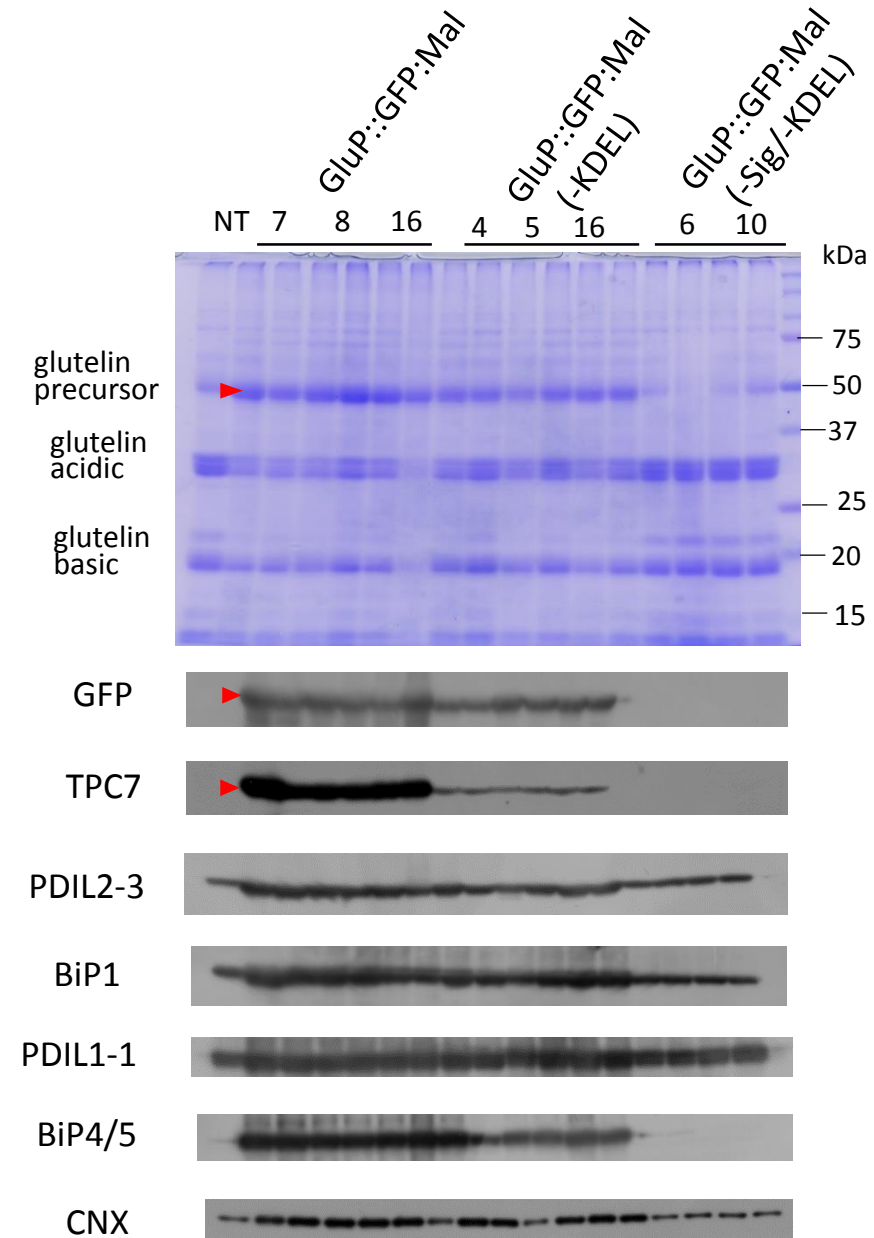


Fig. S1

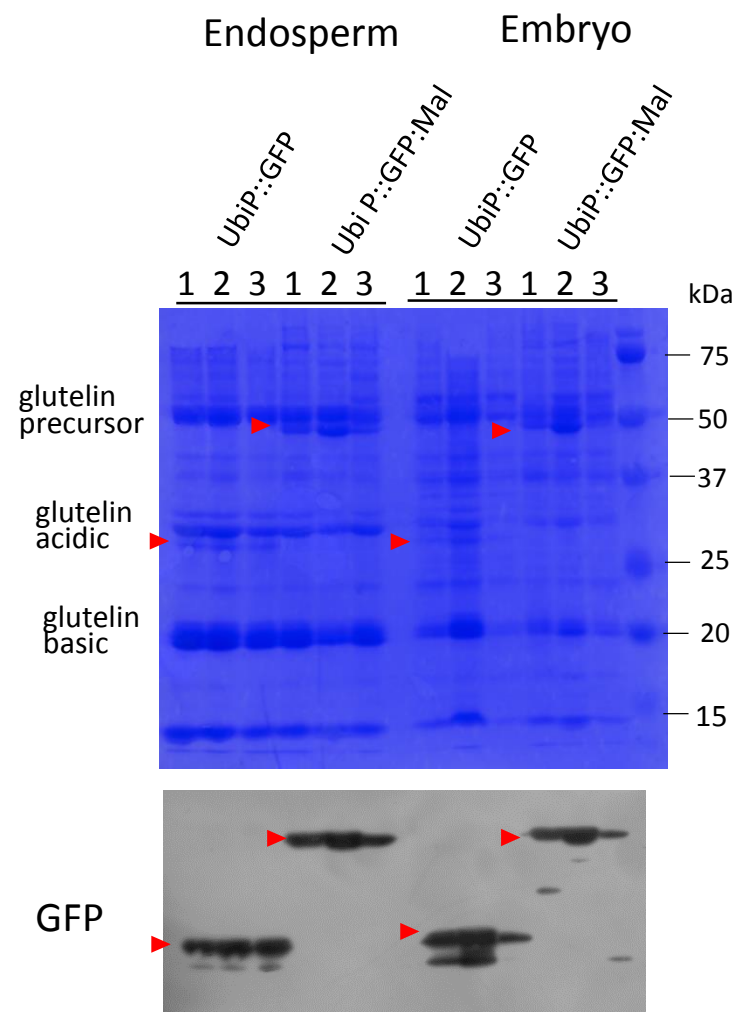


Fig S2

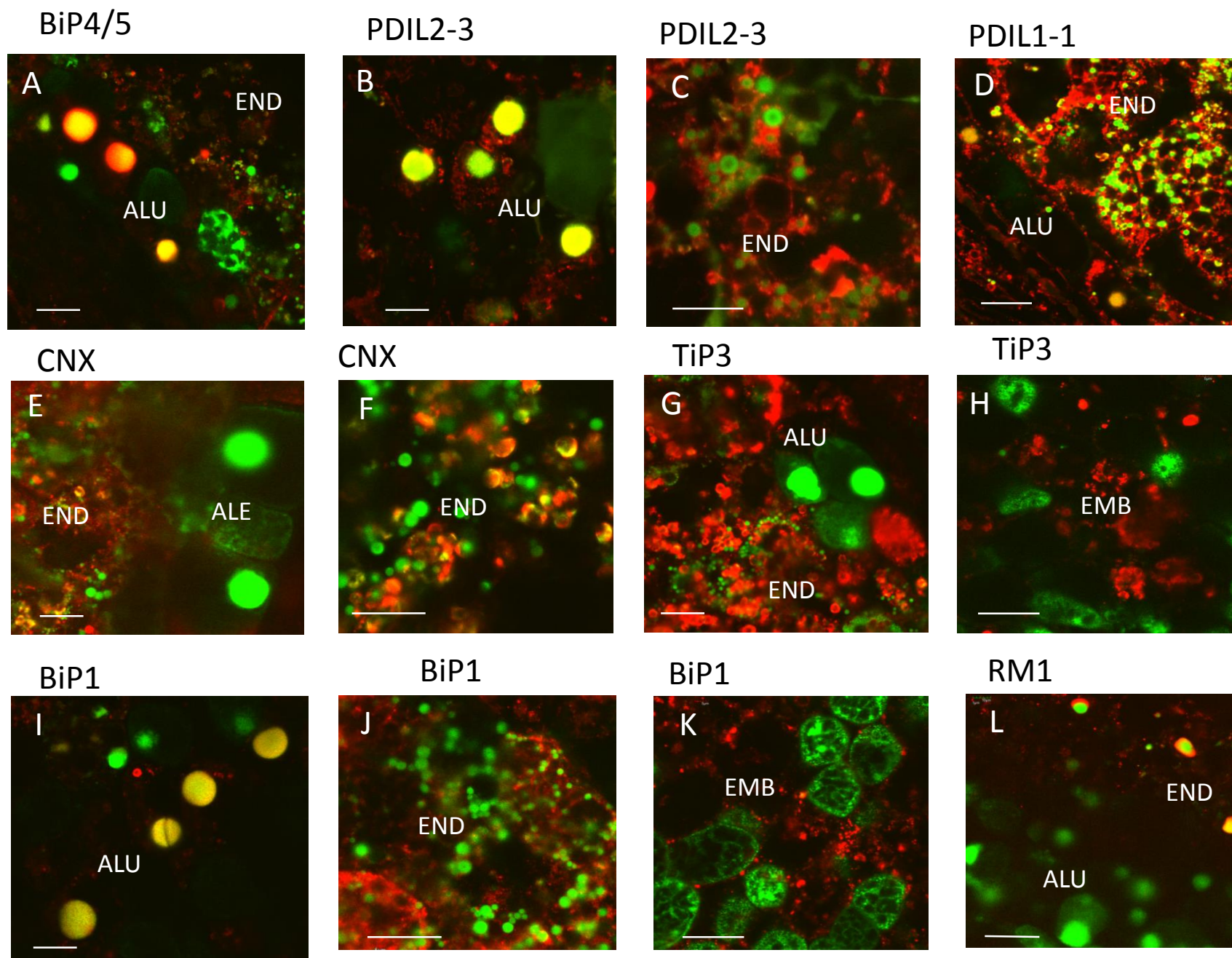


Fig. S3

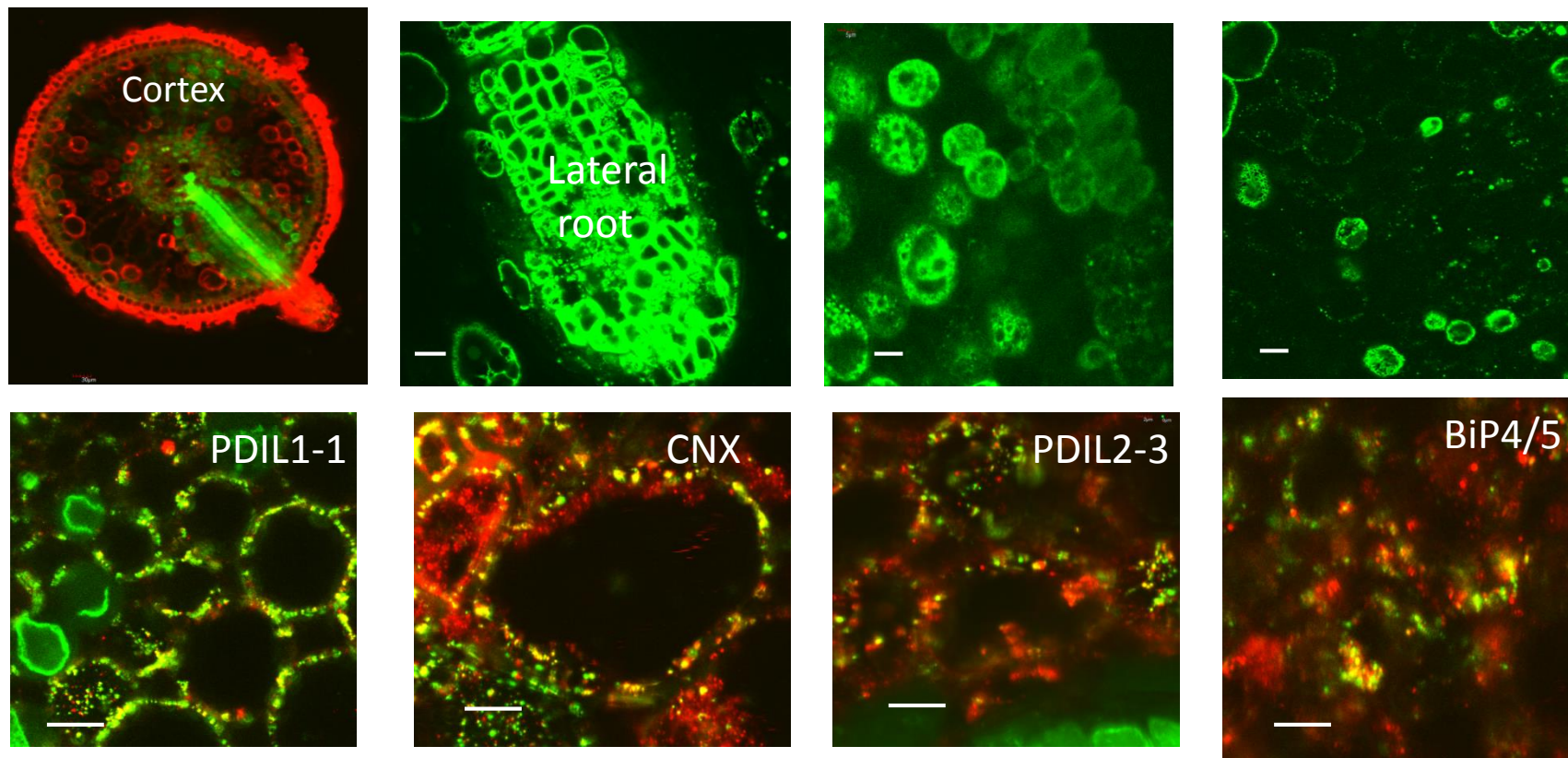


Fig. S4

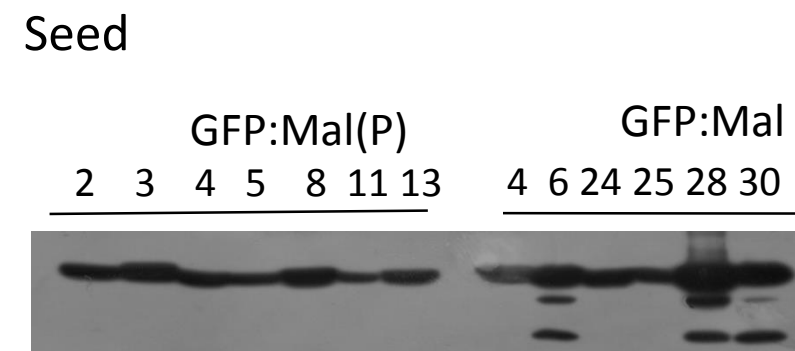
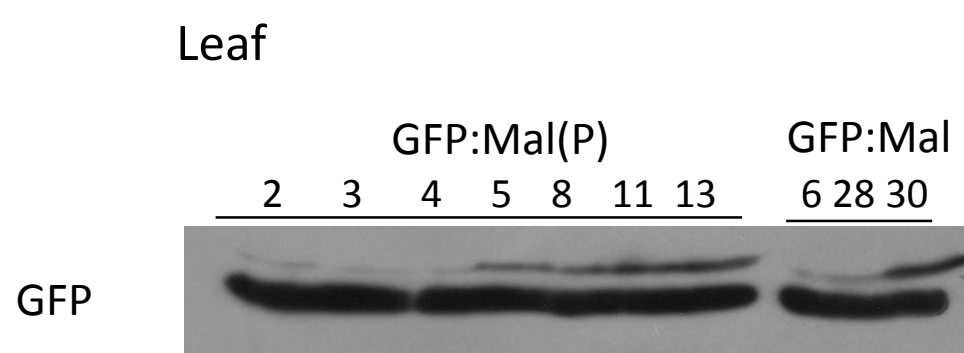


Fig. S5

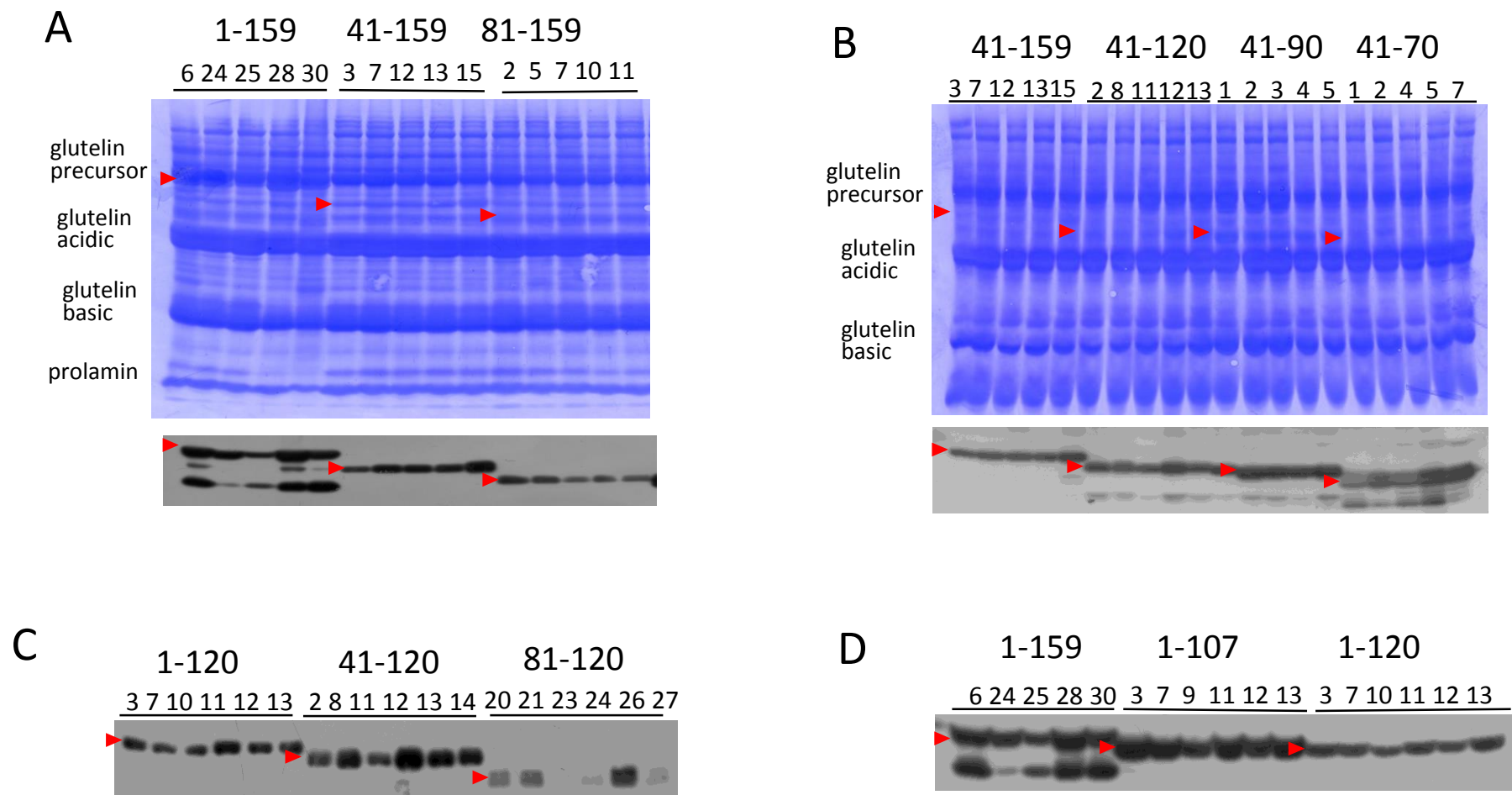
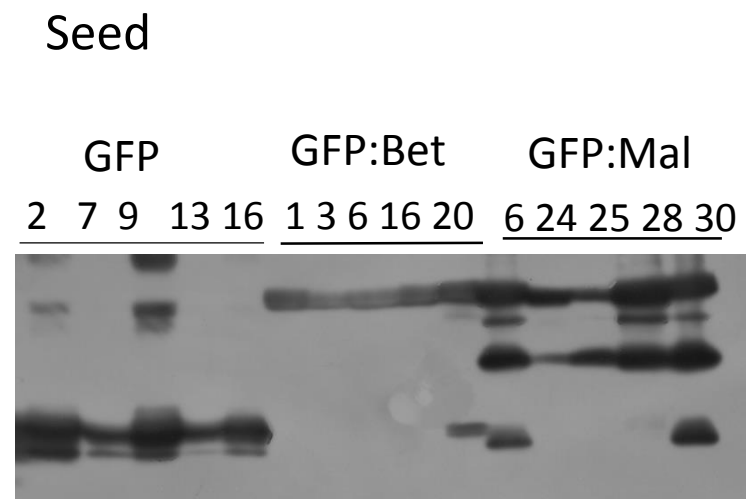
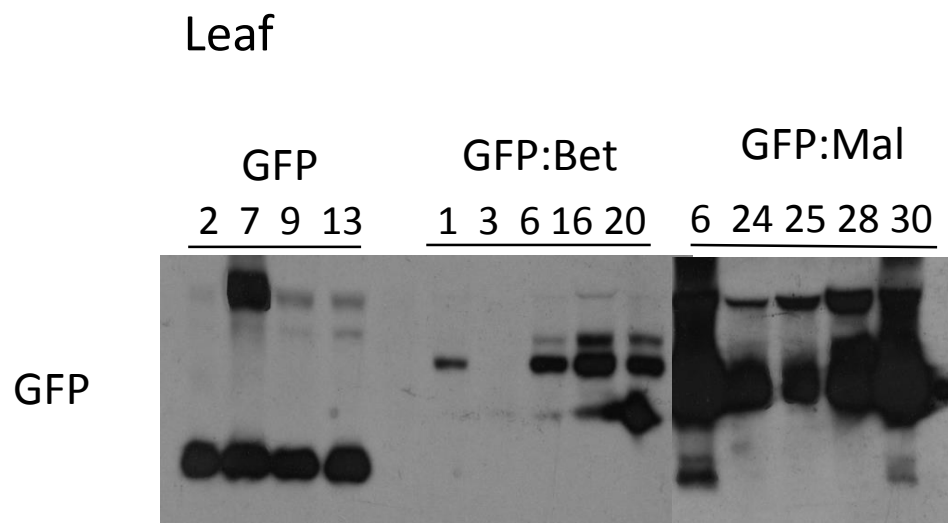


Fig. S6



Supplementary Figures

Figure S1. Immunoblot analysis of GFP:Mald1 fusion proteins expressed in transgenic rice seeds under the control of the endosperm-specific glutelin 2.3 kb *GluB-1* promoter. Total proteins were extracted from mature seeds with SDS-urea buffer, and then separated by 13% SDS-PAGE. Proteins were visualized with CBB staining and accumulation level of GFP:Mald1 was analyzed by immunoblotting using anti-GFP or anti-TPC7 antibody. Effect on expression of chaperon proteins by accumulation of GFP:Mald1 were examined by immunoblotting using several anti-chaperone proteins.

Figure S2. Immunoblot analysis and SDS-PAGE of GFP containing KDEL tag expressed as secretory protein in transgenic rice seed under the control of the ubiquitin promoter. GFP accumulates in both embryo and endosperm of mature seed.

Figure S3. Subcellular localization of GFP:Mald1 in aleurone and starchy endosperm cells in maturing seeds at 15-18 DAFs. Green signal indicates GFP fluorescence or immunofluorescence of GFP immune-stained with anti-GFP antibody. Red signal indicates immunofluorescences of OSBiP-1, OSBiP4/5, OSPDIL1-1, OSPDIL2-3, CNX, OsTiP3 and RM1 (Cys-rich 14kDa prolamin) immunostained with their specific anti-rabbit antibodies. Signals were detected using second antibody with the Alexa-488-conjugated goat anti-mouse IgG (green) and Alexa-564-conjugated goat anti-rabbit IgG (red). The images correspond to the merge channels resulting from the combination of anti-mouse IgG recognizing GFP (green) and anti-rabbit IgG recognizing chaperons (red). Co-localization is represented by yellow or orange color. Scale bars are 5 μ m. ALU, alurone layer; END, endosperm; EMB, embryo

Figure S4. Intracellular localization of GFP/Mald1 and various chaperone proteins in root tissue. Transverse section of transgenic rice root. The images correspond to the merged channel resulting from the combination of GFP of GFP:Mald1 (shown in green) and chaperon proteins (shown in red). Colocalization is shown in orange. Bars 5 μ m

Figure S5. Effect of mutation of Mal d 1 P-loop motif on accumulation level in leaf and seed. P-loop motif of Mal d 1 was mutagenized and this modified Mal d 1 was expressed as fusion protein with GFP in leaves and seeds of several independent transgenic rice lines (GFP:Mal(P)). Accumulation levels in these leaves and mature seeds were compared with those of GFP:Mald1 as control.

Figure S6. Effect of deleted Mal d 1 on accumulation levels of linked GFP in seeds of transgenic rice. Total proteins extracted from mature seeds of 4-5 representative independent transgenic lines for each expression construct were subjected to immunoblot analysis using anti-GFP antibody

Figure S7. Comparison of accumulation levels of GFP, GFP:Betv1 and GFP:Mald1 in leaf and seed. GFP, GFP:Betv1 and GFP:Mald1 were expressed under the control of the maize ubiquitin promoter.