Plant Mol Biol (2009) 71:81–98
DOI 10.1007/s11103-009-9510-1

Characterization of the wheat endosperm transfer cell-specific protein TaPR60

Nataliya Kovalchuk · Jessica Smith · Margaret Pallotta · Rohan Singh · Ainur Ismagul · Serik Eliby · Natalia Bazanova · Andrew S. Milligan · Maria Hrmova · Peter Langridge · Sergiy Lopato

Received: 8 January 2009 / Accepted: 25 May 2009 / Published online: 10 June 2009

Abstract The *TaPR60* gene from bread wheat encodes a small cysteine-rich protein with a hydrophobic signal peptide, predicted to direct the TaPR60 protein to a secretory pathway. It was demonstrated by heterologous expression of recombinant TaPR60 protein that the signal peptide is recognized and cleaved in yeast cells. The full-length gene including promoter sequence of a *TaPR60* orthologue was cloned from a BAC library of *Triticum durum*. A transcriptional promoter-GUS fusion was stably transformed into wheat, barley and rice. The strongest GUS expression in wheat and barley was found in the endosperm transfer cells, while in rice the promoter was active inside the starchy endosperm during the early stages of grain filling. The *TaPR60* gene was also used as bait in a yeast two-hybrid screen. Five proteins were identified in the screen, and for some of these prey proteins, the interaction was confirmed by co-immunoprecipitation. The signal peptide binding proteins, TaUbiL1 and TaUbiL2, are homologues of animal proteins, which belong to proteolytic complexes, and therefore may be responsible for TaPR60 processing or degradation of the signal peptide. Other proteins that interact with TaPR60 may have a function in TaPR60 secretion or regulation of this process. Examination of a three dimensional model of TaPR60 suggested that this protein could be involved in binding of lipidic molecules.

Keywords Grain development · Endosperm · Transfer cells · LTP · Processing · Proteins–protein interaction

Abbreviations

ETC	Endosperm transfer cells
BETL	Basal endosperm transfer layers
BAP2	Basal layer type antifungal protein
DAP	Days after pollination
dbEST	GenBank expressed sequence tag database
GUS	β-Glucuronidase
BAC	Bacterial artificial chromosome
Y2H	Yeast two-hybrid
Y3H	yeast three-hybrid
BD	Binding domain
AD	Activation domain
UBL	Ubiquitin like domain
UBA	Ubiquitin associated domain
APP	Amyloid precursor protein
BCP	Blue copper protein

Introduction

After cellularization, the endosperm of cereals comprises four main tissues: endosperm transfer cells, embryo

The nucleotide sequences have been deposited in the GenBank database under GenBankAccession Numbers: FJ459807 (*TdPR60*), FJ459808 (*TaUbiL1*), FJ459809 (*TaUbiL2*), FJ459810 (*TaBCP*), FJ459811 (*TaPlp-P*), FJ459812 (*TaHVA22L*).

Electronic supplementary material The online version of this article (doi:10.1007/s11103-009-9510-1) contains supplementary material, which is available to authorized users.

N. Kovalchuk · J. Smith · M. Pallotta · A. Ismagul · S. Eliby · N. Bazanova · A. S. Milligan · M. Hrmova · P. Langridge · S. Lopato (✉)
Australian Centre for Plant Functional Genomics, PMB 1, Glen Osmond, SA 5064, Australia
e-mail: sergiy.lopato@acpfg.com.au

R. Singh
School of Agriculture, Food and Wine, The University of Adelaide, Waite Campus, Glen Osmond, SA 5064, Australia

surrounding region, starchy endosperm, and aleurone (Becraft 2001a; Olsen 2004a, b). Endosperm transfer cells (ETC) are characterized by the presence of cell wall ingrowths, which increase the surface of the cellular membrane up to 22-fold and make ETC very efficient in the uptake of nutrients from adjacent maternal vascular tissue to the endosperm (Wang et al. 1994; Becraft 2001b; Olsen 2004a).

Several groups of genes have been found to be expressed only or mainly in ETC (Thompson et al. 2001; Olsen 2004a), but functions have been assigned to just a few. Among such genes is *ZmTCRR-1* (Muniz et al. 2006). The product of this maize gene is a small type-A response regulator protein which belongs to the two-component signal transduction system. *ZmTCRR-1* mRNA transcripts were detected only in ETC, but the protein was detected in the conductive tissue deep inside the endosperm and is probably involved in intracellular signal transmission (Muniz et al. 2006).

Another ETC-specific gene with well characterized function is *ZmMRP-1*. It belongs to the subfamily of R1MYB transcription factors with a DNA binding domain containing a single MYB repeat. It can activate some, but not all, tested ETC-specific promoters. It has been shown to activate promoters of *BETL-1* and *BETL-2* (Gomez et al. 2002) as well as the promoters of *MEG1* (Gutierrez-Marcos et al. 2004) and *ZmTCRR-1* (Muniz et al. 2006). The specific *cis*-element, -TATCTCTATCTC-, recognized by this transcription factor, was recently identified in the promoter of *BETL-1* (Barrero et al. 2006).

Many ETC-specific genes encode low molecular weight cysteine-rich proteins with N-terminal hydrophobic signal peptides. However, there are little data in the literature about the functions of these proteins, the mechanism of proteolytic processing and the localization of the processed peptides. Four types of such proteins were found in maize basal endosperm transfer layers (BETL). BETL-1 and BETL-3 showed sequence homology to defensin-like proteins; BETL-2 had no homologous sequences; and BETL-4 had some homology to the Bowman-Birk family of α-amylase/trypsin inhibitors (Hueros et al. 1995, 1999). Members of these protein families have been shown to inhibit the growth of fungi and bacteria (Broekaert et al. 1997). Defensins can also alter the permeability of fungal plasma membranes and hence may act as regulators of transport through the plasma membrane (Thompson et al. 2001). Proteolytic processing and secretion into adjacent maternal tissue of BETL2 protein, which was later renamed basal layer type antifungal protein 2 (BAP2), was demonstrated by immunolocalization and immunodetection in different grain tissues (Serna et al. 2001). A gene with sequence similarity to *BETL3*, *OsPR9a*, was recently identified in rice. However, its expression in rice is not restricted to endosperm transfer cells, but was also detected in some flower tissues (Li et al. 2008). Expression of BETL proteins is greatly reduced in the maize *reduced grain filling1* (*rgf1*) mutant, which also showed decreased uptake of sugars in endosperm cells at 5–10 days after pollination (DAP) (Maitz et al. 2000).

Maternally expressed gene1 (*meg1*) encodes a further small protein which bears structural similarity to defensins. The *meg1* is expressed exclusively in the basal transfer region of maize endosperm and shows a parent-of-origin expression pattern during early stages and biallelic expression at later stages of endosperm development (Gutierrez-Marcos et al. 2004). The product of this gene is glycosylated and localized to the labyrinthine ingrowths of the walls of transfer cells.

Another class of ETC genes, encoding low molecular weight cysteine-rich proteins, was identified in the barley transfer cell domain of the endosperm coenocyte. The gene was designated *Endosperm 1* (*END1*) (Doan et al. 1996).The expression activity of the barley *END1* gene and its orthologue from wheat were studied using in situ hybridization (Doan et al. 1996; Drea et al. 2005). In both barley and wheat *END1* is expressed in the coenocyte above the nucellar projection during the free-nuclear division stage. After cellularization, *END1* transcripts accumulate mainly in the ventral endosperm over the nucellar projection, but from 8 DAP a low level of expression can also be detected in the modified aleurone and the neighboring starchy endosperm (Doan et al. 1996). The function of *END1* remains unknown.

Expression directed by the *OsPR602* promoter of the rice homologue of the barley *END1* gene, was studied in rice and barley. In rice the promoter of *OsPR602* was active in ETC and above the ETC in parts of the starchy endosperm. However, *GUS* expression was also detected in the maternal vascular tissue adjacent to ETC and vascular tissue of the lemma and palea (Li et al. 2008). Surprisingly, in barley the promoter of *OsPR602* was activated only in ETC and adjacent layers of starchy endosperm and the pattern of temporal and spatial *GUS* expression perfectly correlated with the expression of the *END1* gene from barley (Doan et al. 1996).

In this study, we cloned and characterized expression of an ETC-specific gene from a tetraploid wheat, *Triticum durum*. The gene, designated *TdPR60*, is an orthologue of *TaPR60* from hexaploid wheat (Li et al. 2008) and *END1* from barley (Doan et al. 1996). The spatial and temporal activities of the *TdPR60* promoter were analyzed by stable transformation of wheat, barley and rice with promoter-GUS reporter constructs. In an attempt to understand the function of the gene through the identification of proteins interacting with TaPR60, TaPR60 was used as bait in a yeast two-hybrid screen of a cDNA library prepared from developing wheat grain. Five proteins were found to

interact either with the hydrophobic signal peptide or the N-terminal part of the mature TaPR60 protein. The possible roles of the identified proteins in the proteolytic processing and secretion of TaPR60 are discussed.

Materials and methods

Gene cloning and plasmid construction

The full length cDNA sequence of *TaPR60* was used to probe a BAC library prepared from the genomic DNA of *Triticum durum* cv. Langdon (Cenci et al. 2003) using Southern blot hybridization as described elsewhere (Sambrook et al. 1989). Plasmid DNA from two BAC clones, which strongly hybridized with the probe, was isolated using a Large Construct Kit (QIAGEN). The *T. durum* homologue of *TaPR60* was amplified by PCR using BAC DNA as template and primers derived from the ends of the coding region of *TaPR60* cDNA. The *TdPR60* promoter sequence was first identified on the BAC clone by several consecutive sequencing reactions. As a result of 'walking' along the DNA, about 2,500 bp of sequence upstream from the *TdPR60* translation start codon was obtained. This sequence was subsequently used to design forward and reverse primers for the isolation of the promoter fragment. A 2,147 bp fragment of promoter with a full-length 5′-untranslated region of *TdPR60* was amplified by PCR using AccuPrime™ *Pfx* DNA Polymerase (Invitrogen) from DNA of BAC clone W60-1 (#227 N8) as a template. This was cloned into the pENTR/D-TOPO vector (Invitrogen), the cloned insert verified by sequencing and then subcloned into the pMDC164 plant transformation vector (Curtis and Grossniklaus 2003) using recombination cloning. The resulting construct was designated pTdPR60. Selectable marker genes conferred hygromycin resistance in plants and kanamycin resistance in bacteria. The resulting binary vector was introduced into *Agrobacterium tumefaciens* AGL1 strain by electroporation.

Plasmids for the mapping of protein binding sites were prepared in the following way: PCR amplicons of respective deletions of the coding region were generated using primers with introduced restriction sites and cDNA of TaPR60 as a template and subsequently cloned in the *Eco*R1 and *Bam*H1 restriction sites of the polylinker of the vector pGBKT7. Recombinant clones were confirmed by restriction analysis and verified by sequencing using the T7 primer.

Plant transformation and analysis

The construct pTdPR60 was transformed into rice and barley using *Agrobacterium*-mediated transformation and the method developed by (Tingay et al. 1997) and modified by (Matthews et al. 2001). Rice (*Oryza sativa* L. ssp. Japonica cv. Nipponbare) and barley (*Hordeum vulgare* L. cv. Golden Promise) were used as donor plants. Wheat (*Triticum aestivum* L. cv. Bobwhite) was transformed using biolistic bombardment according to the following protocol. Immature seeds of wheat were surface-sterilized by immersing in 70% ethanol for 2 min, followed by incubation in 1% sodium hypochlorite solution with shaking at 125 rpm for 20 min and finally by three washes in sterile distilled water. Immature embryos (1.0–1.5 mm in length, semitransparent) were isolated aseptically and were placed, with the scutellum side up, on solid culture medium. Embryos developing compact nodular calli were selected using a stereomicroscope and used for bombardment 7–21 days after isolation. The cultures were kept in the dark at 25°C on solid MS (Duchefa, M0222) (Murashige and Skoog 1962) with 30 g/l sucrose and 2 mg/l 2,4-D (MS2). A 4,915 bp long fragment, containing the *TdPR60* promoter, *GUS* gene and *Nos* terminator, was excised from the pTdPR60 construct using *Pme*I and *Bsa*XI, gel purified and co-transformed together with the pUbi-hpt-35S ter cassette [3,676 bp *Pme*I/*Sma*I fragment of the plasmid pMDC32 (Curtis and Grossniklaus 2003)] into wheat using microprojectile bombardment. A DNA-gold coating was prepared according to the protocol of Sanford et al. (1993). Microprojectile bombardment was performed using the Biolistic PDS-1000/He Particle Delivery System (Bio-Rad). Before bombardment, immature embryos were pretreated for 4 h on MS2 medium supplemented with 100 g/l sucrose. Embryos (50/plate) were then placed in the centre of a plate to form a circle with a diameter of 10 mm. Bombardment conditions were 900 or 1,100 psi, with a 15 mm distance from the macrocarrier launch point to the stopping screen and a 60 mm distance from the stopping screen to target tissue. The distance between the rupture disk and the launch point of the macrocarrier was 12 mm. Sixteen hours after bombardment, the calli were transferred to MS2 medium and grown in the dark for 1 week. Two days after bombardment the treated calli were transferred to MS selection medium supplemented with 2 mg/l 2,4-D and 150 mg/l hygromycin B. After 3–6 selections (4–6 months) greening callus tissues were subcultured on MS regeneration medium supplemented with 1 mg/l kinetin and 5–10 mg/l zeatin. Regenerating plantlets were then transferred to jars with half-strength hormone-free MS medium supplemented with 50 mg/l hygromycin B. The fully developed plantlets were acclimated for 7–10 days at room temperature in a liquid medium containing fourfold diluted MS salts. Plants with strong roots were then transplanted into soil and grown under greenhouse conditions to maturity. Transgene integration was confirmed either by PCR using GUS specific primers or by Southern blot hybridization. In the last case genomic DNA from

transgenic lines was digested with *Xho*I and probed with the coding sequence of the hygromycin phosphotransferase selectable marker gene.

Histochemical and histological GUS assays

Histochemical and histological GUS assays were performed as described in (Li et al. 2008) using T_0, T_1 and T_2 (only for barley) transgenic plants and T_1, T_2 and T_3 seeds.

Northern blot hybridization and quantitative PCR analysis

Northern blot hybridization was performed as described elsewhere (Sambrook et al. 1989). The same filter membrane was hybridized consecutively with a full length cDNA of *TaPR60* and 3′ UTRs of each cDNA cloned in the yeast two-hybrid screen. Q-PCR analysis was performed using primers from coding and 3′ untranslated regions of *TaPR60*. cDNAs prepared from wheat (cv. Chinese Spring) grain at different days after pollination were used as a template. Grain of each stage was collected from the middle of several spikes from each of 3–4 plants and combined together; each data point is the mean of four replicate PCR reactions. The Q-PCR procedure was described in (Burton et al. 2004). *TaPR60* mRNA copy number was normalized against GAPDH mRNA copy number.

Protein expression and co-immuno-precipitation

The coding region of *TaPR60* with an N-terminal c-*Myc* epitope was subcloned into the pYES2.1 expressional vector using the pYES2.1 TOPO TA Expression Kit (Invitrogen). The recombinant TaPR60 protein that was expressed in the yeast (*Saccharomyces cerevisiae*) strain EGY contained an N-terminal c-*Myc* epitope as well as a C-terminal V5 epitope and 6His-tag. Yeast cells were collected at 0, 4, 8 and 24 h after induction of protein synthesis. Total yeast protein was extracted and analyzed by 15% PAGE–SDS-electrophoresis and detected on Western blots with antibodies (Clontech) against either c-*Myc* or 6His epitopes.

For co-immunoprecipitation experiments, the bait construct of TaPR60 and prey constructs of TaPR60 (negative control), TaUbiL1, TaBCP, TaClp-P, TaPI and TaHVA22L were linearized with either of *Bam*HI, *Sac*I or *Pst*I and used as templates for in vitro protein expression in the TNT Coupled Wheat Germ Extract System Kit (Promega). All prey proteins were co-expressed together with the bait TaPR60 and contained N-terminal HA epitopes, while the bait protein contained a c-*Myc* epitope. Co-immunoprecipitation was performed as described in the manual for the MATCHMAKER Co-IP Kit (Clontech), with the following modifications: 50 µl, rather than 10 µl, of in vitro translation mix was used for the immunoprecipitation; c-*Myc* antibodies directly coupled to agarose beads (Anti-c-Myc Agarose Conjugate, SIGMA) were used instead of soluble antibodies and protein A-agarose; and detection of in vitro translated proteins on the Western blot membranes was performed using immunodetection with respective antibodies instead of radioactive protein labeling and detection.

Molecular modeling of the TaPR60 protein

Three-dimensional (3D) molecular models of a putative wheat lipid binding protein (LBP) without signal peptide sequence were constructed using the Modeller 9v2 program (Sali and Blundell 1993). To identify the most suitable template for TaPR60, searches were performed through the Structure Prediction Meta-Server (Ginalski et al. 2003), MetaPP server (Rost et al. 2004), SeqAlert (Bioinformatics and Biological Computing, Weizmann Institute of Science, Israel), Protein Data Bank (PDB) (Berman et al. 2000) and 3D-PSSM Server (Kelley et al. 2000). The highest scoring templates suggested by these servers were lipid binding protein (LBP) from *Prunus persica* (PDP accession number 2alg, chain A; called hereafter 2alg:A) that also contained two lipid molecules, and a non-specific lipid transfer protein from wheat (PDP accession number 1n89). The 2alg:A containing 92 amino acid residues was retrieved from the PDB (Berman et al. 2000). The secondary structure predictions of the TaPR60 protein were performed with SAM T06 (Karplus et al. 1998) and with 3D-PSSM (Kelley et al. 2000). The positions of secondary structure elements and hydrophobic clusters were manually examined with hydrophobic cluster analysis (Callebaut et al. 1997). The malign3 module of Modeller 9v2 was used to align 2alg:A (91 residues) with TaPR60 (84 residues) and also to calculate the positional sequence identity between the two sequences. The sequence alignment generated by Malign3 was able to detect positions of eight paired cysteines that formed a basis for molecular modeling. The aligned 2alg:A and TaPR60 sequences were further used as input parameters to build 3D models on a Linux Red Hat workstation, running a Fedora Linux Core 4 operating system. The final 3D molecular model of TaPR60 was selected from 80 models that showed the lowest value of the Modeller objective function. The stereochemical quality and overall G-factors of the 2alg:A structure and TaPR60 model were calculated with PROCHECK (Ramachandran et al. 1963; Laskowski et al. 1993). *Z*-score values for combined energy profiles were evaluated by Prosa2003 (Sippl 1993); the plots were smoothed using a window size of 20 amino acid residues. Spatial superposition of the 2alg:A structure

and TaPR60 model was performed in the DeepView molecular browser (Guex and Peitsch 1997), using a fragment alternate fit routine. The distributions of hydrophobic and polar areas on the surface of TaPR60 were calculated with GRASP (Nicholls et al. 1991). The molecular graphics were generated with the PyMol (http://www.pymol.org/) and GRASP software packages. The molecular model of TaPR60 including its signal peptide was built with I-TASSER (Zhang 2008).

Yeast two-hybrid screening and yeast three-hybrid assay

The full-length coding region of the *TaPR60* cDNA was amplified using primers incorporating *Eco*RI and *Bam*HI restriction. The PCR product was cloned into the *Eco*RI-*Bam*HI sites of the pGBKT7 vector (Clontech). The resulting construct was used as a bait plasmid to screen a MATCHMAKER cDNA library prepared from the whole grain of *Triticum aestivum* L. cv. Chinese Spring at 0–6 DAP. Yeast transformation and screening were performed as described in the Yeast MATCHMAKER System 3 manual. Putative positive clones were verified by co-transformation with the bait construct and empty bait plasmid into yeast. The size of inserts was determined using PCR with primers derived from vector sequences flanking the insert. Clones containing the same size of inserts were analyzed by *Hae*III digestion, combined in groups, and inserts for several representatives from each group were sequenced. Full length cDNAs encoding TaUbiL1 and TaUbiL2 were isolated using the SMART™ RACE cDNA Amplification Kit (Clontech) and total RNA from wheat grain at 0–6 DAP.

The yeast three-hybrid system was used for the investigation of competition of two proteins for the same binding site on the third protein. The pBridge vector (Clontech) contains two distinct multiple cloning sites to allow constitutive expression of fusion BD-protein and a strong inducible expression of the second protein. The coding regions of *TaPR60* were subcloned into the *Eco*RI-*Bam*HI sites of pBridge as a fusion with the BD of the yeast GAL4; *TaHVA22* was cloned into *Not*I-*Bgl*II sites of the same vector under the *MET25* promoter. TaUbiL1, TaBCP and TaClp-P were expressed as activation domain fusions from the pGADT7 vector. The pGADT7 and pBridge vectors were used as negative controls. After yeast transformation with appropriate variants of constructs the yeast colonies were first grown on -Leu, -Trp media (selection for plasmids) and then transferred by replica plating to media lacking either -Leu, -Trp, -His (selection for BD-protein—AD-protein interactions) or -Leu, -Trp, -His, -Met (selection for protein–protein interactions plus induction of protein-competitor). The effect of the third protein competitor was indicated by expression of nutritional marker (absence of yeast growth on the -Leu, -Trp, -His, -Met media in the case were protein–protein interactions were disrupted).

Results

Cloning of the TdPR60 gene

The cDNA of *TaPR60* was isolated from a cDNA library prepared from the liquid part of the syncytial endosperm of *Triticum aestivum* at 3–6 DAP. A single cDNA of *TaPR60* was identified among approximately 200 cDNAs randomly selected for sequencing and grain-specific expression of the gene was demonstrated by quantitative (Q)-PCR (Li et al. 2008). The full length cDNA of *TaPR60* (Acc.No. Eu264062) was used as a probe to screen a bacterial artificial chromosome (BAC) library prepared from genomic DNA of *Triticum durum* cv. Langdon (Cenci et al. 2003). Two BAC clones were selected for further analysis on the basis of the strength of the hybridization signals. BAC DNA was isolated and used as a template for PCR with primers derived from the beginning and the end of the coding region of *TaPR60*. Only one from the two isolated BAC clones (#227 N8) gave the predicted PCR product. Sequencing of the PCR product revealed that the cloned insert is part of the genomic clone of the *TaPR60* orthologue from *T. durum.* The cloned gene was designated *TdPR60*. The coding region of *TdPR60* was interrupted by a single 172 bp long intron. Alignment of the protein sequences of TaPR60 and TdPR60, deduced from the genomic sequence, shows differences in four amino acid residues (Fig. 1). Analysis of the protein sequences of TaPR60 and TdPR60 with PSORT signal prediction computer software (http://psort.ims.u-tokyo.ac.jp/form.html)

Fig. 1 Alignment of protein sequences of TaPR60, TdPR60 and HvEND1. Identical amino acids are in *black boxes*, similar amino acids are in *grey boxes*

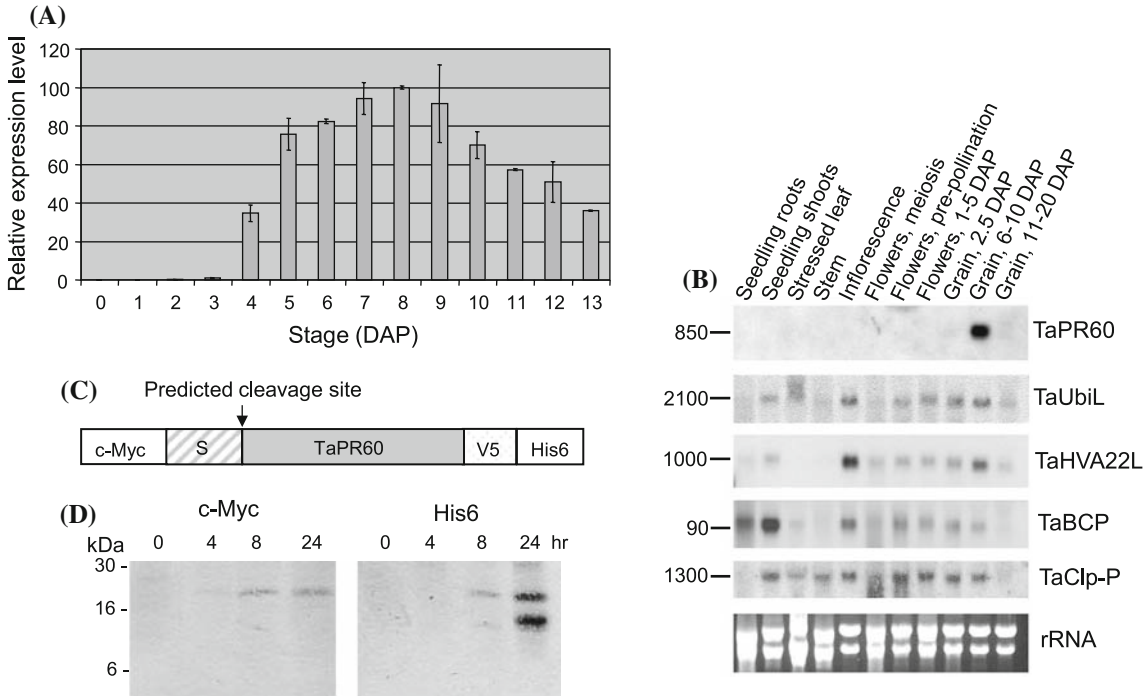


Fig. 2 Expression patterns of *TaPR60* and genes which encode interacting partners of *TaPR60*. **a** Relative expression of *TaPR60* in wheat grain at different stages of grain development as determined by quantitative real-time PCR. Expression is shown as a percentage of maximal expression seen at 8 days after pollination (DAP). **b** Expression of *TaPR60* and its interaction partners in different wheat tissues shown by northern blot hybridisation. Approximate sizes of the transcripts are given on the right. The ethidium bromide staining of the RNA (rRNA) is shown at the bottom. The tissues used are indicated at the top. **c** Schematic representation of recombinant TaPR60 protein expressed in yeast: c-*Myc* and V5—epitopes recognized by commercially available antibodies, His6—C-terminal His tag, S—signal peptide of TaPR60. (**d**) Detection of unprocessed and processed TaPR60 on Western blots using antibodies against c-*Myc* and His6 epitopes at 0, 4, 8, and 24 h after initiation of recombinant protein expression in yeast cells

indicated the presence of an N-terminal hydrophobic signal peptide of twenty amino acid residues in length, implying secretion of both TaPR60 and TdPR60 to either the vacuole or apoplast. The predicted cleavage site was situated between amino acid residues 20 and 21 (Fig. 2c), producing a mature protein of 87 amino acid residues with a predicted molecular mass of 9.5 kDa and a pI point of 8.24.

Spatial and temporal activity of the TdPR60 *promoter in wheat, barley and rice*

The promoter sequence of *TdPR60* was first identified by several consecutive sequencing reactions using DNA isolated from the BAC clone #227 N8 as template. As a result of 'walking' along the DNA, about 2.6 kbp of sequence upstream from the translation start codon were obtained. This sequence was subsequently used to design forward and reverse primers for the amplification from BAC DNA of a promoter fragment corresponding to 2,147 bp upstream of the translational start of *TdPR60*. The promoter fragment was cloned into the plant transformation vector pMDC164 (Curtis and Grossniklaus 2003), which harbors a hygromycin resistance marker gene for selection of transgenic plants, to provide the transcriptional GUS fusion promoter construct designated pTdPR60.

Computer analysis of the *TdPR60* promoter using PLACE software

(http://www.dna.affrc.go.jp/PLACE/signalscan.html) and a database of plant *cis*-acting regulatory DNA elements (Higo et al. 1999) revealed few *cis*-elements, which are known to be involved in endosperm-specific activation. Among them is the RY repeat or legumin box, CATGCAC, found in seed storage protein genes (Fujiwara and Beachy 1994). The RY repeats act together with the G box to provide seed-specific, ABA-dependent gene expression (Ezcurra et al. 1999). However, the G-box motif was not present in the *TdPR60* promoter. The promoter of *TdPR60* did contain a $(CA)_n$ element, CCAAACAC, which was originally identified in storage proteins in *Brassica napus*, where it is responsible for both embryo- and endosperm-specific transcription (Ellerstrom et al. 1996). Several core sites for binding of DOF proteins (AAAG) were found to be concentrated in the 250 bp long segment of the *TdPR60* promoter. They can potentially provide interaction with PBF-like DOFs, which

were demonstrated to bind to the prolamin box and specifically activate promoters in the endosperm (Diaz et al. 2005). However, a complete prolamin box was not identified. The (TATCTC) repeats, which specifically interact with *ZmMRP-1* transcription factor and are responsible for ETC-specific promoter activation in maize (Barrero et al. 2006), were not identified in the *TdPR60* promoter. Mapping of promoter segments responsible for the specific activation in ETC using biolistic bombardment was unsuccessful because of the extremely low frequency of successful transient transformation events observed in endosperm tissue. No activity of the promoter was observed after biolistic bombardment of isolated embryos at different stages of development, although bombardment with a positive control construct (pUbi-GUS) was successful (data not shown).

For stable transformation experiments, the pTdPR60 construct was transformed into the *Agrobacterium tumefaciens* strain AGL1 and the presence of plasmid in selected colonies was confirmed by PCR using specific primers. Transformed *Agrobacterium* was subsequently used to introduce constructs into rice and barley. A 4,915 bp long fragment, containing the *TdPR60* promoter, *GUS* gene and *Nos* terminator, was excised from the pTdPR60 construct and co-transformed together with a plant selectable marker cassette (Ubi-Hpt-Nos) into wheat using microprojectile bombardment. The integration of promoter-GUS fusions in transgenic plants was confirmed by PCR using primers derived from promoter and GUS sequences. Southern blot hybridization was also performed on some transgenic wheat and barley lines, providing data showing the number of inserted copies of the selectable marker *hpt* gene. The number of inserts in transgenic lines of barley varied from one (Line 2) to two (Line 3) and three (Lines 4–6). In wheat, the number of inserts varied from two to about 16 or more. The number of inserts in transgenic rice plants was not examined.

Six T_0 wheat lines were selected using the GUS staining assay, from which three were selected for further analysis. Ten T_1 progeny for each of the three lines were analyzed. Among the 30 progeny, there were two plants with very strong, two plants with relatively strong, and 13 plants with weak transgene expression; and 13 plants exhibited no GUS expression. All positive lines demonstrated the same pattern of GUS expression.

Eighteen T_0 barley lines were selected, from which two were used for the analysis of T_1 plants. From nine T_1 plants, one had strong, two plants had weak and six plants had no transgene expression. All T_0 and T_1 plants demonstrated the same pattern of GUS expression. The differences in expression levels between different transgenic lines showed no correlation with the number of inserts.

Twenty-four T_0 lines of transgenic rice were analyzed for GUS activity. Eight lines demonstrated strong promoter activity and the same pattern of gene expression. The T_1 generation was analyzed for two lines (six plants for each). All positive plants had the same patterns of transgene expression as T_0 plants. Wild type plants and/or plants transformed with a vector containing only the selectable marker cassette were used as negative controls. No differences were found between wild type plants and plants transformed with the control vector. The strength of the *TdPR60* promoter in barley and rice remained the same in the T_0 and T_1 generations. In contrast, promoter strength in wheat significantly decreased in the T_1 generation relative to the corresponding T_0 parents.

In wheat and barley the *TdPR60* promoter was active in endosperm transfer cells (Fig. 3a, b). The weak activity was also detected in adjacent layers of starchy endosperm cells. In wild type plants, the endogenous promoter was transcriptionally activated at about 3 DAP (Fig. 2a), however, the GUS activity in transgenic plants was detected only at 9 DAP (shown for barley at Fig. 3b). According to northern blot analysis and Q-PCR data, the endogenous promoter was active until approximately 15–20 DAP (Fig. 2a, b), but the GUS activity could be detected until at least 40 DAP. Neither northern blot analysis nor GUS staining identified the activity of the *TdPR60* promoter in any other plant tissues.

Surprisingly, in rice the *TdPR60* promoter was activated in a different way than in wheat and barley. The GUS activity was initially detected at 7–8 DAP mainly in the embryo surrounding region of endosperm. However, within 2–3 days strong GUS expression was found mainly inside the starchy endosperm and could be detected until at least 47–50 DAP (Fig. 4). The weak activity of the promoter in rice was found also in the layer of endosperm transfer cells (data not shown), however, the activity of the promoter inside the starchy endosperm was much stronger.

Proteolytic processing of TaPR60 in yeast

Attempts to raise specific antibodies against TaPR60 were unsuccessful, because recombinant TaPR60 could not be expressed in bacteria, and subsequent immunization attempts with three synthetic peptides, derived from the signal peptide and both ends of the mature protein, were also unsuccessful. However, TaPR60 was successfully expressed in yeast. The recombinant protein contained an N-terminal c-*Myc* epitope and C-terminal His-tag (Fig. 2c). The level of expression of the recombinant protein in yeast was low, providing insufficient protein for antibody preparation. However, while using this expression system, we observed that yeast is able to process by proteolytic cleavage at least part of the synthesized protein. Two protein products were detected with the antibody against the C-terminal His-tag, but only one product with larger

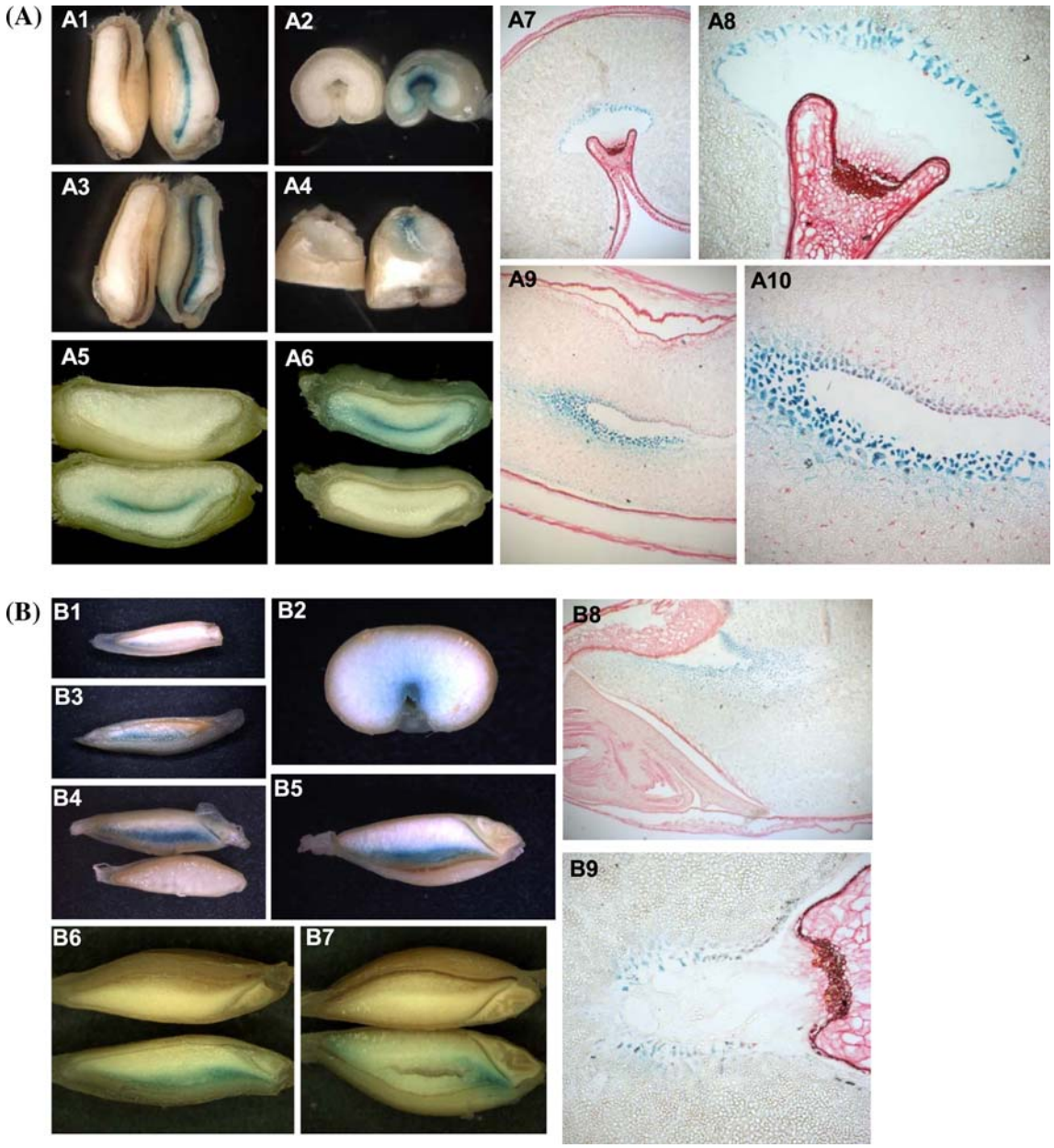


Fig. 3 Spatial and temporal GUS expression in wheat (**a**) and barley (**b**) directed by the TdPR60 promoter. GUS activity in wheat (**a**) and barley (**b**) detected in hand-cut longitudinal (A1, A3, A5, A6, B1, B3–B7) and transverse sections (A2, A4, and B2) of transgenic caryopsis, at 9 DAP (B1), 13 DAP (B3, B6), 15 DAP (A5, A6, B4), 17 DAP (B7), 24 DAP (A1–A4), and 34 DAP (B2, B5). No GUS activity is seen in wild type caryopses in the panels A1–A4, B4, B6 and B7. Grain was from T_0 *transgenic lines* (A1–A4, B1–B5) and T_1 progeny (A5, A6, B6, B7). Histochemical GUS assay counterstained with safranin in: 10 μm thick transverse sections of transgenic wheat (A7, A8) and barley (B9) caryopsis, and longitudinal sections of transgenic wheat (A9, A10) and barley caryopsis (B8) at 15 DAP (A9, A10), 16 DAP (B8), 31 DAP (A7, A8) and 34 DAP (B9)

molecular weight was recognized by the antibody against the N-terminal c-*Myc* epitope (Fig. 2). The size of the detected proteins perfectly matched the predicted sizes of the full length recombinant protein (18.1 kDa) and the mature form of the protein (13.5 kDa). About 20–30% of synthesized protein molecules remained uncleaved 24 h after the induction of protein synthesis (Fig. 2d). The signal peptide was not detected. Successful expression of recombinant TaPR60 in yeast and the presence of a significant amount of uncleaved protein encouraged us to use TaPR60 as bait in a yeast two-hybrid screen.

Isolation of interaction partners of TaPR60

A yeast two-hybrid cDNA library prepared from immature wheat grain collected at 0–6 DAP was used for the screen.

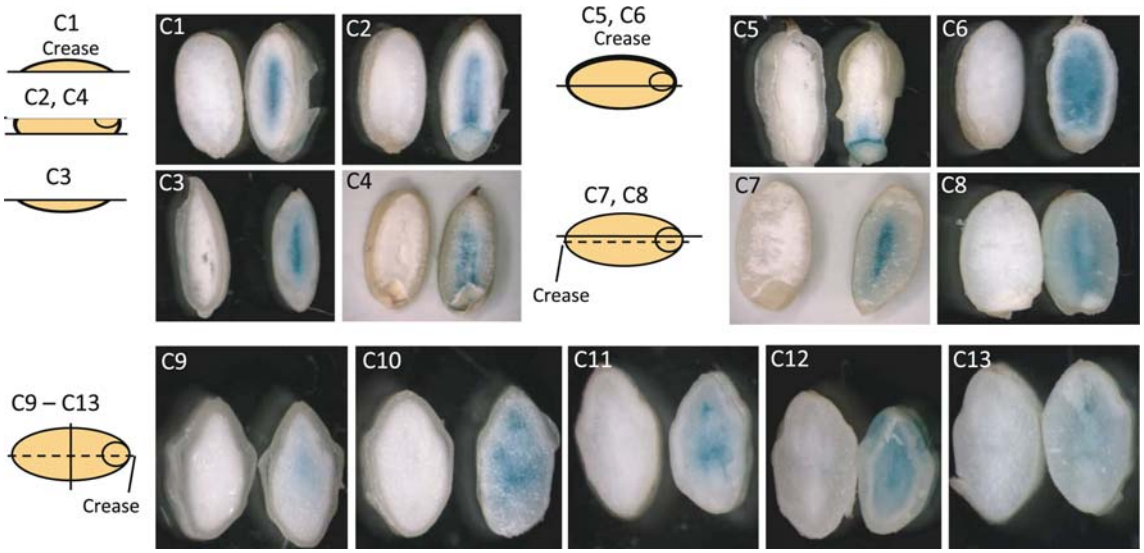


Fig. 4 Spatial and temporal GUS expression in rice (C) directed by the TdPR60 promoter. GUS activity in rice detected in hand-cut longitudinal (C1–C8) and transverse (C9–C13) sections of control (*left side* of each figure) and transgenic (*right side* of the figure) caryopses. GUS activity is detected in 8 DAP (C5), 10 DAP (C9), 13 DAP (C1, C2, C6, C10), 20 DAP (C3, C4, C11), 30 DAP (C7), 41 DAP (C12) and 47 DAP (C8, C13) old grain, but no GUS activity is seen in wild type caryopses (presented on the left side of same panels). A schematic representation of grain position during sectioning is provided

After screening several million colonies and a second round of selection for positive clones, four groups of clones were isolated. The clone inserts represented five different cDNAs. The first group, comprising approximately 75% of all clones, contained partial cDNA sequences of two close homologues, possibly homoeologues, encoding ubiquitin-like proteins, designated TaUbiL1 and TaUbiL2. The cloned fragments of both cDNAs lacked approximately 200 bp at the beginning of the coding region. The absent 5′ ends of cDNAs of *TaUbiL1* and *TaUbiL2* were cloned using 5′ RACE-PCR.

Three independent clones contained the same insert encoding the full length coding region of the catalytic subunit of a Clp-P protease, designated *TaClp-P*. Inserts from the third group of clones, comprising five independent clones, encoded the full length CDS of a blue copper protein, designated *TaBCP*. The fourth group of two clones contained sequence for a HVA22 protein family member, which was designated *TaHVA22L*. The domain structures and the lengths of partial sequences, as well as the network of protein–protein interactions identified, either by the yeast two-hybrid screen, the yeast two-hybrid assay and/or co-immunoprecipitation, are shown in Fig. 5.

The fact that five different proteins interacted with the relatively short TaPR60 protein suggested the possibility that TaPR60 binds all of them with the same ‘sticky’ segment. However, mapping of the binding sites of TaPR60 revealed the binding sites for the different proteins were located to different parts of the TaPR60 protein (Fig. 6a, b). Both TaUbiL1 and TaUbiL2 proteins interact strongly with a ten amino acid-long segment of the signal peptide. TaBCP also binds the signal peptide, but the interaction is dependent on the presence of an intact cleavage site. The N-terminal half of the mature protein is sufficient for the strong interaction with TaClp-P. In contrast, binding of TaHVA22L to the same region seems to be much stronger in the non-processed than in the processed protein. TaClp-P interacts with the segment of TaPR60, which is approximately twice the length of the segment that is needed for the strong binding of TaHVA22L. The yeast two-hybrid screen with ‘processed’ TaPR60 as bait, in which bait sequence starts from amino acid residue number 21, was repeated twice but no binding partners were detected. A subsequent screen using the full length TaBCP as bait identified several clones encoding TaUbiL1 and TaUbiL2, but TaPR60 was not isolated (data not shown). Most of the clones isolated with TaHVA22L as bait contained cDNA of TaClp-P (data not shown).

We used the yeast three-hybrid system based on the pBridge vector (Clontech) to identify if the TaHVA22L protein can compete with TaClp-P for the interaction with TaPR60. Two proteins were expressed from the pBridge vector: TaPR60 constitutively as a bait protein and TaHVA22L at several-fold excess as an inducible competitor. TaClp-P, TaUbiL1, TaBCP or empty vector were used as prey. It was demonstrated that induction of TaHVA22L does not affect interaction of TaPR60 with TaUbiL1 and TaBCP. However, TaHVA22L did disrupt the interaction with TaClp-P.

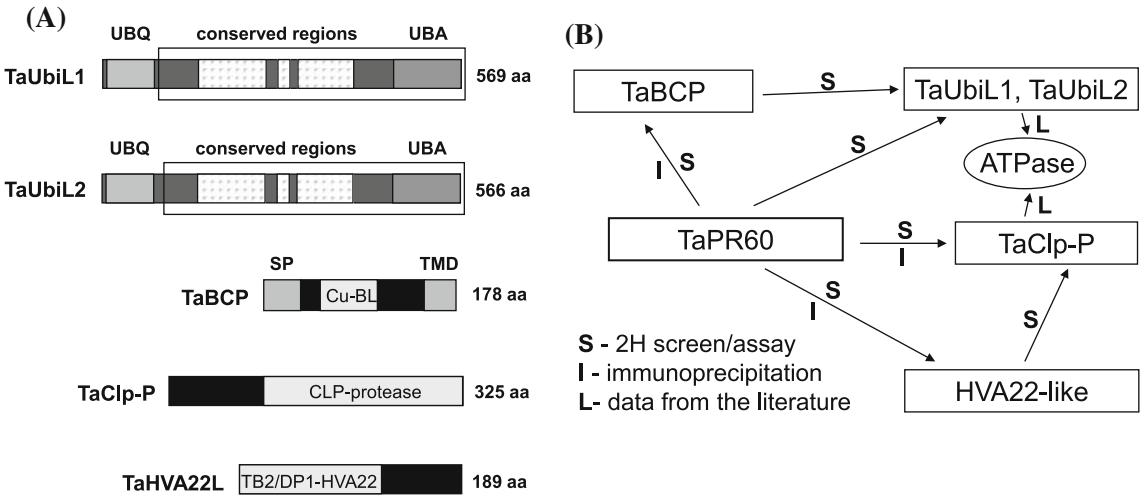


Fig. 5 Domain structure and interactions of proteins identified in the Y2H screen. **a** Domain structure of identified proteins: *UBQ* Ubiquitin like domain, *UBA* Ubiquitin activation domain, *SP* signal peptide, *Cu-BL* copper binding domain, *TMD* trans-membrane domain, *CLP-protease* proteolytic domain, *TB2/DP1-HVA22* characteristic fold for HVA22 family proteins. **b** Protein–protein interactions demonstrated in the Y2H screen/assay, co-immunoprecipitation and/or described in the literature

Validation of the interactions identified in the yeast two-hybrid screen

The expression at the mRNA level for each of the identified genes was studied in various wheat tissues and at different stages of flower and grain development, using the northern blot hybridization technique. It was demonstrated that mRNAs of all the proteins that were identified in the yeast two-hybrid screen are expressed in grain at the same stage of development as *TaPR60*, with overlapping expression profiles. At the time of induction of *TaPR60* gene expression, expression of *TaUbiL1* and *TaHVA22* was at their maximum levels. There was also correlation in the expression levels of *TaUbiL*, *TaBCP* and *TaClp-P* in late flowers and early grain.

We have confirmed by co-immunoprecipitation the interactions between TaPR60 and some of the proteins identified in the yeast two-hybrid screen. The proteins were expressed from the linearized bait (TaPR60) and prey vectors (TaPR60, TaClp-P, TaBCP, and TaHVA22L) in cell-free extracts prepared from wheat germ. The recombinant TaPR60 contained the c-*Myc* epitope, whereas other proteins contained the HA epitope. TaPR60 was co-expressed in pairs with each of the interacting partners and with itself, but with a different epitope tag. Successful expression was confirmed by immunodetection on Western blots using c-*Myc* and HA antibodies (Fig. 6c). The c-*Myc* antibody coupled to agarose was used for co-immunoprecipitation. Co-precipitated proteins were identified by Western blot hybridization using the anti-HA antibody (Fig. 6d). TaBCP and TaClp-P were found to co-precipitate with TaPR60. A very weak band for TaHVA22L was also detected on the wet nitrocellulose membrane after staining, but this could not be seen on the digitized scan of the membrane. In Fig. 6d, the position of this band is shown with the lowest arrow. TaPR60 could not be detected in the precipitate. TaPR60 was used as a negative control, because no homodimerization of this protein has been demonstrated in the Y2H assay (data not shown). Interaction of TaPR60 with TaUbiL1 and TaUbiL2 could not be confirmed by co-immunoprecipitation because all attempts to express ubiquitin-like proteins either in cell free extracts or in bacteria were unsuccessful.

Discussion

Several families of ETC-specific genes encoding different small cysteine-rich proteins with N-terminal hydrophobic signal peptides have been identified in developing grain (Hueros et al. 1995; Doan et al. 1996; Serna et al. 2001; Gutierrez-Marcos et al. 2004). Although data obtained for mRNA expression, and protein localization and structure allow speculation about the possible role of some of these proteins in grain development, their exact function remains unclear. In wheat, reverse genetic approaches are still too complex, expensive and unpredictable, mainly because of the large and complex genome, the paucity of genomic sequence, and the expenses and resources required for transformation and maintenance of sufficient numbers of transgenic plants. Forward genetics in wheat is also problematic because of the redundancy in the hexaploid genome, and, in the case of small cysteine-rich proteins, redundancy in function among this family of proteins.

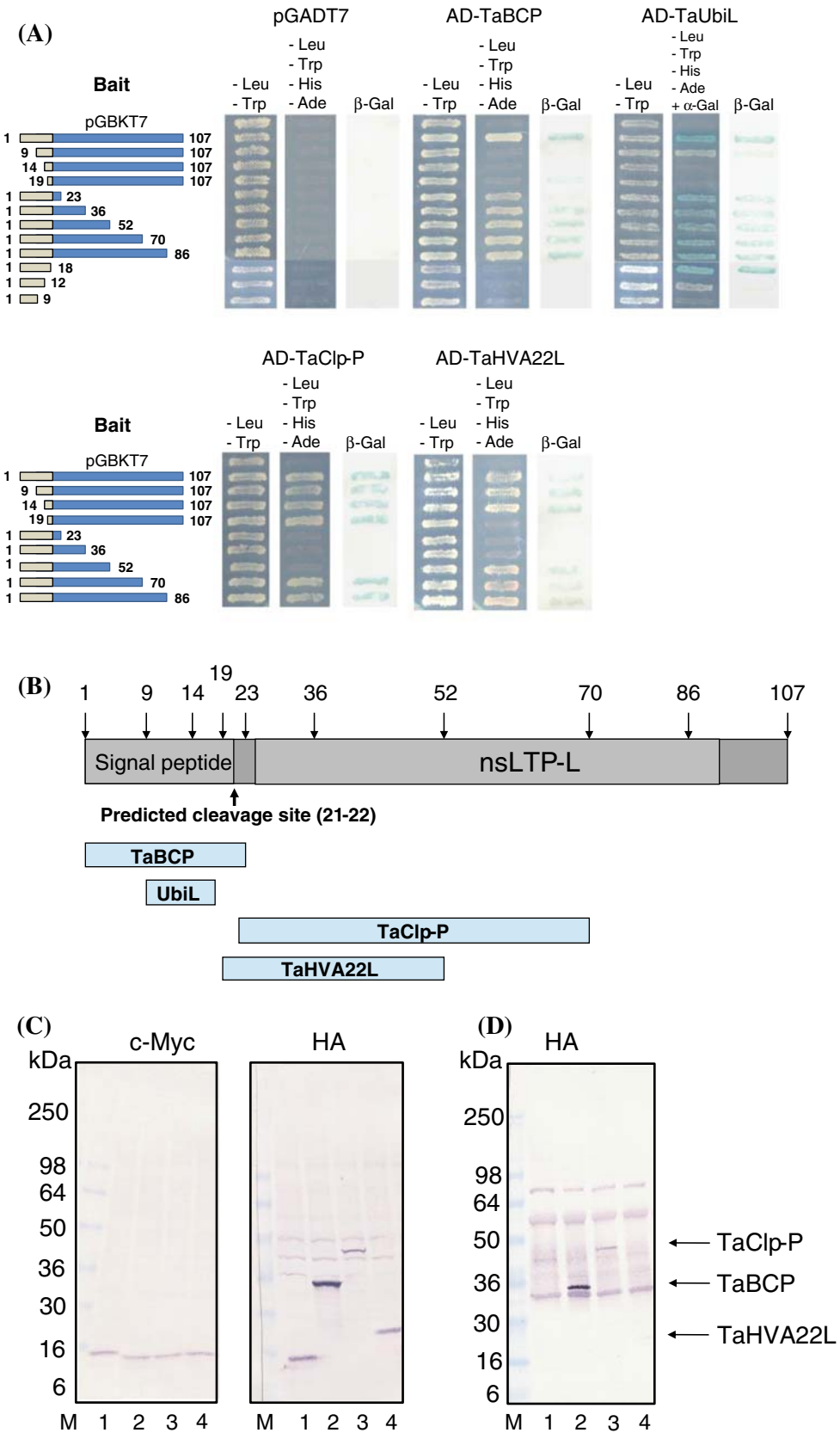


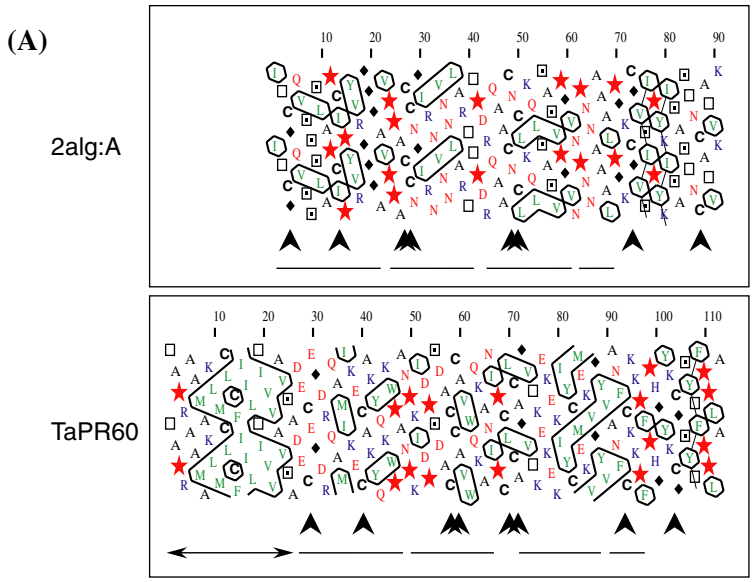
Fig. 6 Interaction partners of TaPR60 identified in the Y2H screen. **a** Interaction between full length and different truncated versions of identified proteins demonstrated by the Y2H assay. Truncated forms of TaPR60 are shown on the *left side*; first and last amino acid residues are indicated. **b** Interactions with particular segments of the TaPR60 protein, as mapped in the yeast two hybrid assays, are shown as *blue boxes* under a schematic picture of TaPR60. **c** Successful expression in cell free extracts of TaPR60 and its interacting partners, as identified in the Y2H screen, demonstrated by Western blot detection. **d** Results of co-immunoprecipitation of TaPR60 with differently tagged TaPR60 (negative control) and identified proteins demonstrated by Western blot. Proteins were expressed in pairs: *lane 1* TaPR60 with c-*Myc* and HA tags, *lane 2* TaPR60 (c-*Myc*) and TaBCP (HA), *lane 3* TaPR60 (c-*Myc*) and TaClp-P (HA), *lane 4* TaPR60 (c-*Myc*) and TaHVA22L (HA), *lane M* molecular weight markers. Co-precipitated proteins are indicated with arrows

Besides these genetic approaches, one of the first steps to elucidate the function of genes is to study the temporal and spatial patterns of gene expression. It was demonstrated earlier by quantitative RT-PCR that *TaPR60* is an endosperm-specific gene (Li et al. 2008). Analysis of the promoter of the *T. durum* orthologue of *TaPR60* provided additional and more precise data on expression of the *END1*-like genes. The anticipated utility of the promoter in biotechnological applications prompted analysis of the promoter in three agriculturally important plants: wheat, barley and rice. The spatial and temporal activities of the promoter in wheat and barley were nearly identical. In wheat and barley the promoter was active in endosperm transfer cells and adjacent layers of starchy endosperm. In barley, the spatial and temporal activity of the *TdPR60* promoter was practically indistinguishable from the activity of the promoter of the rice gene *OsPR602* (Li et al. 2008). Surprisingly, in rice the *TdPR60* promoter was activated 1–2 days earlier and was detected mainly inside of the starchy endosperm (Fig. 4), in contrast to its activity in wheat and barley. Unlike the *OsPR602* promoter in rice (Li et al. 2008), no activity of the wheat promoter was identified in rice flowers or any other tissues. This suggests at least partial incompatibility of transcription factors and *cis*-elements responsible for ETC-specific activation of wheat and rice promoters. Use of strongly diverged or different *cis*-elements in rice and wheat is supported by the observation that identical or near-identical *cis*-elements in the *TaPR60* and *OsPR602* promoters could not be identified. The expression pattern of *TdPR60* suggests involvement of the gene product either in signal transduction or nutrient transfer into the endosperm. Although transcripts of *TaPR60* were detected in endosperm as early as 3 DAP (Fig. 2a), at the beginning of coenocyte cellularization, GUS activity was not observed before 9 DAP, when cellularization of the endosperm was near completion. This result implies that the gene product is not involved in the endosperm cellularization process. The difference between timing of mRNA and protein expression can be explained by possible translational regulation provided by the 5′UTR sequence, which was included in the promoter construct. However, we have no experimental evidence for a role of the 5′UTR.

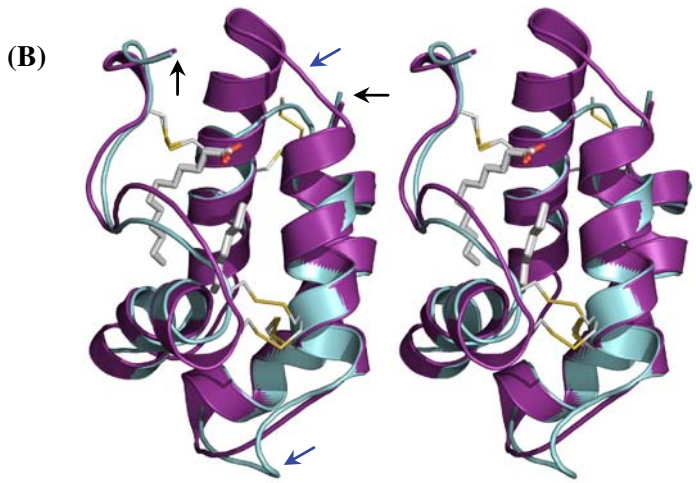
Molecular modeling is another complementary approach, which can assist in predicting and better understanding the function of proteins. This has been demonstrated in our current structural proteomics era, where 3D structures or molecular models are becoming integral components of multidisciplinary biological research. An example of usefulness of molecular modeling has recently been demonstrated (Rodrigues et al. 2008). A molecular model of the wheat TaPR60 protein was constructed using as a template a putative *Prunus persica* LBP (Protein DataBank accession number 2alg:A; Pasquato et al. 2006) that was identified by several prediction servers listed in the Materials and Methods section. The LBP protein represents a canonical all alpha protein categorized in the SCOP protein classification system. More precisely, these structures belong to the family of 'Plant Lipid Transfer and Hydrophobic Proteins' (Andreeva et al. 2008). Other hydrophobic proteins are also classified in this protein group, such as a soybean hydrophobic protein (Baud et al. 1993). The positional sequence identity and similarity between the 2alg:A and TaPR60 sequences were 17 and 31%, respectively. This rather low sequence identity but significant similarity emphasized complexity of modeling (Sanchez and Sali 1998; John and Sali 2003; Sippl and Wiederstein 2008). However, the analysis of sequences *via* hydrophobic cluster analysis (HCA) (Callebaut et al. 1997) indicated that the positions of eight cysteine residues were highly conserved in both sequences (Fig. 7a; marked by arrowheads). This observation suggested that the cysteine residues in TaPR60 could also be paired, and the associated parts of the sequences could form α-helical secondary structural elements (Fig. 7a; marked by lines).

The 3D models of TaPR60 with and without signal peptide were generated (Fig. 7) and their stereochemical parameters (Ramachandran et al. 1963; Laskowski et al. 1993) and combined energy profiles (Sippl 1993) were found to be acceptable (data not shown). Both proteins (Fig. 7) contained a series of highly conserved α-helices folded in all alpha protein class according to SCOP protein classification (Andreeva et al. 2008). The major structural differences between the TaPR60 model and its template stemmed from one-eight-residue deletion and one-four-residue insertion in TaPR60 (Fig. 7b, marked by blue arrows). Two lipoid molecules (lauric acid and heptane) were also modeled in the central hydrophobic cavity of TaPR60 (contained in 2alg:A) that was formed in the central region of the helical bundle (Fig. 7b); the volume of this cavity was approximately 1,240 $Å^3$, compared to the volume of the cavity of 2alg:A, which equaled 1,205 $Å^3$. The side chains of predominantly hydrophobic and aromatic residues (Val, Ala, Phe, Trp) lining up the cavity were rotated out to the sides of the protein fold, so a rather large hydrophobic cavity could be formed that could internalize possibly several, and at least two lipid or lipid-like molecules. Three amino acid residues (Trp36, Cys45, Lys52) were making contacts of less than 3 Å with the two lipid molecules through hydrophobic interactions.

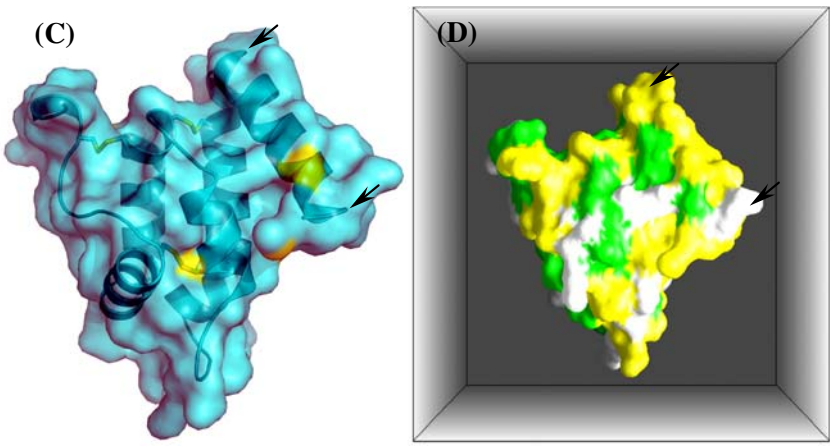
Molecular modeling of a full length sequence of TaPR60, including its hydrophobic signal peptide, was performed with the I-TASSER (Zhang 2008). The model of TaPR60 indicated that the signal peptide comprising 20 amino acid residues (Fig. 7a) formed an additional α-helix that had a significant hydrophobic surface (Fig. 7c, d; the



Hydrophobic cluster analysis of 2alg:A and TaPR60 (with signal peptide)



Molecular model of TaPR60 and 3D structure of 2alg:A (less signal peptide) with bound lipoid molecules



Molecular model of TaPR60 (with signal peptide)

Fig. 7 Molecular model of the TaPR60 protein. **a** Hydrophobic cluster analysis of 2alg:A and TaPR60. Positions of N-terminal hydrophobic signal peptide (*large arrow*), four paired conserved cysteines (*arrowheads*) and α-helices (*lines*) are marked. **b** Superposition of the TaPR60 model (*cyan*) on the template crystal structure of 2alg:A (*magenta*) showing secondary structural elements. A root-mean-square-deviation value for 69 structurally equivalent residues is 1.5 Å over the Cα backbone positions. The dispositions of bound lauric acid (*left*) and heptane (*right*), in *cpk* colours, internalised in protein cavities, and the positions of four invariant disulfide bridges (*yellow*) are also shown. The 8-residue deletion and 4-residue insertion in TaPR60 are marked by *blue arrows*. The *right-hand-side* and *left-hand-side* arrows indicate N- and C-terminal parts of the proteins, respectively. **c** Molecular fold and surface of TaPR60 with a hydrophobic signal peptide. The molecular surface with cysteines (*yellow*) is projected over the surface of TaPR60. **d** Molecular surface of TaPR60 with a hydrophobic signal peptide. *Yellow* and *green* indicate hydrophobic (Ala, Val, Leu, Ile, Met, Phe, Pro) and polar (Ser, Thr, Tyr, His, Cys, Asn, Gln, Try) areas, respectively, while *white* areas denote surface-exposed charged residues (Lys, Arg, Asp, Glu) and Gly, calculated by GRASP. In panels (**c**) and (**d**) the *lower* and *upper* arrows indicate the positions of N- and C-terminal parts of the hydrophobic signal peptide, respectively

signal peptide is marked by black arrows). It could be concluded that this additional structural element could quite dramatically change the overall hydrophobic character of the TaPR60 protein.

In summary, molecular modeling suggested that TaPR60 could be involved in binding of lipid or lipid-like molecules. As the cavity shape remains practically the same in the processed and unprocessed TaPR60, the protein can potentially keep the lipid and lipid-like molecule(s) in the cavity during processing and secretion. Thus one of the possible functions of this protein can be transport of lipids to or through membranes (e.g. co-secretion outside the cell with molecules of lipid-like origin).

An additional approach to reveal the function of a gene is identification of interaction partners of the gene product and determination of function through the assignment to particular protein complex(es) with already known function(s). For example, identification of several well known splicing factors among interaction partners of TaRSZ38 assigned this factor to the spliceosome and suggested its role as a splicing factor (Lopato et al. 2006). Before applying the yeast two-hybrid system to heterologous proteins, which potentially can be processed by proteolytic cleavage, one should first determine if this protein will be processed in the same or a similar way in yeast. It was shown that yeast can indeed proteolytically process TaPR60. However, this processing is not efficient and a substantial portion of the synthesized protein remained unprocessed even 24 h after the induction of protein expression (Fig. 2d). Possible reasons for incomplete processing include a high level of expression (saturation of the capacity of the processing machine); inefficient processing in heterologous systems (lack of compatibility with the processing machine); and/or inefficient processing due to conformational changes caused by the additional peptide sequences from the c-*Myc*, V5 epitope, and His-tag. The sequences of the epitopes, which were added to both ends of recombinant TaPR60, could potentially interfere with the correct folding of the protein. Regardless of the reason, the incomplete processing permitted use of the Y2H system to identify proteins, which can interact with TaPR60. It should be noted, however, that during the yeast two-hybrid screen partial cleavage of TaPR60 most probably leads to the presence of baits of two types: the first one comprising the yeast GAL4 binding domain (BD) fused to full length TaPR60, and the second (the result of signal peptide cleavage) comprising the BD coupled to the signal peptide only. This theoretically could lead to the enrichment of clones for preys which interact with the signal peptide only.

A relatively small protein, TaPR60 interacted with five proteins, as summarized in Fig. 5. Only two of the five proteins belong to the same family: TaUbiL1 and TaUbiL2 (TaUbiLs) showed 96.7% amino acid sequence identity and can be considered as either close homologues or homoeologues (the same gene from different genomes). They both contain an N-terminal ubiquitin-like region (UBL), placing them within a large family of ubiquitin-like proteins. In addition the C-terminal part of TaUbiLs contains an ubiquitin-associated (UBA) domain. Both domains have been implicated in targeting and degradation of proteins by the ubiquitin-proteosome pathway. The long conserved region is situated between the UBL and UBA domains. The protein is enriched for asparagine-proline repeats, NPXΨ, where X is any amino acid and Ψ is a hydrophobic amino acid. Among these repeats is one NPF motif, which is the core-binding element to the protein domain known as the Eps15 homology (EH) domain. This domain was found in a variety of proteins involved in endocytosis, vesicle sorting and cytoskeletal architecture (Mayer 1999). There are no data on the function of TaUbiL1 and TaUbiL2 or similar proteins in plants. However, these proteins have considerable structural and sequence similarity to several relatively well studied proteins from mammals and yeast, such as PLIC-1 and PLIC-2 (Wu et al. 1999), Ubiquilin1 (Mah et al. 2000; Wu et al. 2002), XDRP1 (Funakoshi et al. 1999b), and Dsk2 (Biggins et al. 1996; Funakoshi et al. 2002). The protein sequence alignment of TaUbiL1 and TaUbiL2 to animal and yeast proteins is shown in Supplementary Fig. 1. It has been demonstrated that human Ubiquilin 1 (UBQLN1) plays a central role in both positive and negative regulation of proteosomal degradation of various proteins including cyclin A, γ-aminobutyric acid receptor, and Hepatitis C virus RNA-dependent RNA polymerase (Funakoshi et al. 1999a; Bedford et al. 2001; Gao et al. 2003). Mapping of segments of TaPR60 for involvement in the interaction with TaUbiL1 and TaUbiL2 revealed that both proteins specifically interact with a 10 amino acid-long segment of the hydrophobic signal peptide and the remainder of the TaPR60 protein is not necessary for this interaction (Fig. 6a, b). The molecular model of the unprocessed TaPR60 implies that the hydrophobic signal peptide is exposed on the surface of the protein and easily available for the interaction with TaUbiLs (Fig. 3c). These findings suggest that TaUbiL1 and TaUbiL2 are involved in the regulation of the signal peptide processing and/or utilization of the cleaved signal peptide. This speculation is supported by the data obtained using down-regulation of human UBQLN1. Down-regulation of this protein dramatically increased the rate of amyloid precursor protein (APP) maturation and trafficking through the secretory pathway, leading to increased secretion of APP (Hiltunen et al. 2006). Therefore, UBQLN1 works as a modulator of trafficking from the trans-Golgi network to the cell surface in the secretory pathway. UBQLNs were shown to be localized to the ER and vesicular structures of several types

of cells (Mah et al. 2000). It was also demonstrated that UBQLNs can regulate endoproteolysis (Massey et al. 2005). It was suggested that the mechanism of this regulation is by binding in close vicinity to the cleavage site and physically protecting it from proteolytic cleavage. It was also demonstrated that the ER-bound proteosome is involved in endoproteolysis (Massey et al. 2005). At the same time UBQLN-like proteins were shown to be a part of the proteosome as they interact with components of the 26S proteosome through the UBL domain (Kleijnen et al. 2000). The ability of yeast cells to process recombinant TaPR60 can be explained by the presence in yeast of machinery for signal protein processing, similar to other eukaryotes, including plants. Yeast contain at least two genes encoding TaUbiL-like proteins, *Dsk2* (Funakoshi et al. 2002) and *Rad23* (Chen and Madura 2002), which possibly function as TaUbiLs during TaPR60 processing in yeast cells.

The high frequency of independent clones for TaUbiL1 and TaUbiL2 isolated in the Y2H screen may be explained by the ability of these proteins to interact equally well with both types of baits (processed and un-processed). The cloned sequences were slightly different in length but none contained a UBL domain. That indicates that the UBL domain is not important for protein–protein interactions with TaPR60. Apparently there was strong selection in the screen for clones which contain TaUbiLs with a full length conserved region, because shorter clones, with the UBA domain only, were not found. This suggests that TaUbiL1 and TaUbiL2 interact with TaPR60 through the conserved domain rather than through the C-terminal UBA domain. However, experimental verification of this hypothesis by mapping of protein interacting segments of TaUbiLs is required.

The third protein identified in the Y2H screen, TaBCP, also interacted with the signal peptide (Fig. 5a, b). It belongs to the large family of phytocyanins, small blue copper proteins from non-photosynthetic parts of the plant (Ryden and Hunt 1993). TaBCP has an N-terminal hydrophobic signal peptide (the first 23 amino acid residues) and a second hydrophobic sequence at the C-terminal end. PSORT software predicts plasma membrane localization of TaBCP with a certainty of 0.919. TaBCP contains a single copper binding site, which is formed by two histidines, one cysteine, and one methionine or a glutamine. In some phytocyanins, amino acid residues important for copper binding were lost during evolution and proteins lost their ability to bind copper (Hunt et al. 1985). The role of most phytocyanins is still unclear. Some were demonstrated to be involved in pollen tube guidance regulation (Kim et al. 2003; Dong et al. 2005), induced by some abiotic stresses, or be involved in plant defense responses (Kreps et al. 2002; Provart et al. 2003). Other data also suggest that phytocyanins function as signaling molecules because they were found to be nodule-specific in legumes (Fedorova et al. 2002) and to bind to tobacco S-RNase, which is involved in self-pollen rejection (Cruz-Garcia et al. 2005). TaBCP co-immunoprecipitated with TaPR60 and interacts with TaUbiL1 and TaUbiL2 suggesting that these four proteins are in the same complex. Alternatively, it could be aberrant performance of the Y2H system, whereby the yeast homologues of TaUbiL build a bridge between the two wheat proteins in the Y2H assay, for example if Dsk2 interacts with the signal peptides of both TaPR60 and TaBCP. However if this were the case, TaPR60 also should produce "homodimers" through Dsk2 in the yeast-two hybrid assay, which did not happen.

One of the proteins identified in the screen with TaPR60 is a Clp protease. TaClp-P binds to the N-terminal part of the mature TaPR60 and does not require the TaPR60 signal peptide for the interaction. TaClp-P is a proteolytic subunit of Clp proteases, which have been found in both prokaryotes and eukaryotes. In plants, *Clp* genes are constitutively expressed and their products were found in plastids, mitochondria and in the cytoplasm (Peltier et al. 2004). Clp proteins are directly involved in the proteolysis of misfolded, unfolded and damaged proteins and are required for cell division, development of sporulation, genetic competence and quality control of protein translation [reviewed in (Sokolenko et al. 2002)]. It is tempting to speculate that TaClp-P is involved in proteolytic degradation of the signal peptide. Several families of peptidases are involved in signal peptide processing and degradation [listed in (Sokolenko et al. 2002)]. However, we could not find any data on the involvement of plant Clp proteases in processing of N-terminal signal peptides. However, in the bacteria *Bacillus subtilis* inactivation of ClpP improves the processing of the precursor of α-amylase exposed on the outer surface of the cytoplasmic membrane (Pummi et al. 2002). These findings extend the regulatory role of Clp protease to protein secretion most likely through the regulation of the turnover of signal peptidases. A mutation in the gene for another ClpP-like protease from *Bacillus subtilis*, TepA, leads to a reduced rate of the precursor translocation across the membrane because of the accumulation of signal peptides in the membrane or the translocation channel. It is likely that this Clp-like protease is required for the degradation of proteins or signal peptides that are inhibitory to protein translocation (Bolhuis et al. 1999). These data suggest that TaClp-P is involved either in degradation of the signal peptide of TaPR60 after proteolytic cleavage of the peptide by a signal peptidase, or in the regulation of the level of signal peptidase(s). The regulatory subunit of Clp protease has ATPase activity (Peltier et al. 2004), and it is interesting to note that human Ubiquilins have been shown to interact with ATPases (Wada et al. 2006). Therefore, the

regulatory ATPase subunit of protease could potentially connect ClpP with TaUbiL1 and TaUbiL2.

The TaHVA22L protein is a homologue of HVA22 from barley (Shen et al. 2001) and a possible orthologue of Arabidopsis AtHVA22a (Chen and Madura 2002). The level of barley HVA22 transcript increases during grain development, as well as in response to elevated concentrations of ABA and some environmental stresses. Phylogenetic analyses show that TaHVA22L is in fact closer to AtHVA22a than to barley HVA22. The Arabidopsis gene shows very little response to ABA and drought, suggesting that it, and, by association, *TaHVA22L*, are housekeeping genes. The role of their gene products in plants, however, remains unclear. The *HVA22* homologue from yeast (*Saccharomyces cerevisiae*), *Yop1p*, is involved in vesicular traffic (Brands and Ho 2002). The yeast double mutant of *yop1* and *sey1*, the latter being a homologue of the Arabidopsis *RHD3* gene (Bajaj and Mohanty 2005), is defective in vesicular traffic as evidenced by the accumulation of transport vesicles and a decrease in invertase secretion (Brands and Ho 2002). Hence it is possible that TaHVA22L acts as a positive regulator of secretion. TaHVA22L interaction with TaClp-P may be stronger than with TaPR60, as several independent clones of TaClp-P but no clones of TaPR60 were found in the Y2H screen when full length TaHVA22L was used as bait (data not shown). According to results of the Y3H assay, TaHVA22L can disrupt interaction between TaPR60 and TaClp-P. In a similar experiment it was demonstrated that induction of over-expression of TaHVA22L does not influence interaction between TaPR60 and either TaUbiL1 or TaBCP. These results correlate with the results of the mapping of protein–protein interactions and demonstrate that TaHVA22L and TaClp-P compete for the same protein segment of TaPR60, but may interact with this protein simultaneously with TaUbiL1 and/or TaBCP.

Taken together the data demonstrate that TaPR60 interacts with several proteins, which are involved in the regulation of secretion and degradation of signal peptides in other organisms. According to the molecular modeling predictions, TaPR60 could belong to the family of lipid-binding proteins and may be involved in lipid transfer. The identified interacting proteins could help TaPR60 to fulfill its main function, which could be the transfer of lipid or lipid-like molecules through the cell membrane to the apoplast. Whether TaPR60 can interact with other proteins after processing and secretion remains to be investigated.

Acknowledgments We thank Professor U. Grossniklaus for providing us with the collection of pMDC vectors, and Dr. U. Langridge and R. Hosking for assistance with growing plants. This work was supported by the Australian Research Council, the Grains Research and Development Corporation and the Government of South Australia.

References

Andreeva A, Howorth D, Chandonia JM, Brenner SE, Hubbard TJP, Chothia C, Murzin AG (2008) Data growth and its impact on the SCOP database: new developments. Nucleic Acids Res 36:D419–D425. doi:10.1093/nar/gkm993

Bajaj S, Mohanty A (2005) Recent advances in rice biotechnology-towards genetically superior transgenic rice. Plant Biotechnol J 3:275–307. doi:10.1111/j.1467-7652.2005.00130.x

Barrero C, Muniz LM, Gomez E, Hueros G, Royo J (2006) Molecular dissection of the interaction between the transcriptional activator ZmMRP-1 and the promoter of BETL-1. Plant Mol Biol 62:655–668. doi:10.1007/s11103-006-9047-5

Baud F, Pebaypeyroula E, Cohenaddad C, Odani S, Lehmann MS (1993) Cristal-structure of hydrophobic protein from soybean—a member of a new cysteine-rich family. J Mol Biol 231:877–887. doi:10.1006/jmbi.1993.1334

Becraft PW (2001a) Cell fate specification in the cereal endosperm. Semin Cell Dev Biol 12:387–394. doi:10.1006/scdb.2001.0268

Becraft PW (2001b) Plant steroids recognized at the cell surface. Trends Genet 17:60–62. doi:10.1016/S0168-9525(00)02165-X

Bedford FK, Kittler JT, Muller E, Thomas P, Uren JM, Merlo D, Wisden W, Triller A, Smart TG, Moss SJ (2001) GABA(A) receptor cell surface number and subunit stability are regulated by the ubiquitin-like protein Plic-1. Nat Neurosci 4:908–916. doi:10.1038/nn0901-908

Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) The protein data bank. Nucleic Acids Res 28:235–242. doi:10.1093/nar/28.1.235

Biggins S, Ivanovska I, Rose MD (1996) Yeast ubiquitin-like genes are involved in duplication of the microtubule organizing center. J Cell Biol 133:1331–1346. doi:10.1083/jcb.133.6.1331

Bolhuis A, Matzen A, Hyyrylainen HL, Kontinen VP, Meima R, Chapuis J, Venema G, Bron S, Freudl R, van Dijl JM (1999) Signal peptide peptidase- and ClpP-like proteins of *Bacillus subtilis* required for efficient translocation and processing of secretory proteins. J Biol Chem 274:24585–24592. doi:10.1074/jbc.274.35.24585

Brands A, Ho THD (2002) Function of a plant stress-induced gene, *HVA22*. Synthetic enhancement screen with its yeast homolog reveals its role in vesicular traffic. Plant Physiol 130:1121–1131. doi:10.1104/pp.007716

Broekaert WF, Cammue BPA, DeBolle MFC, Thevissen K, De-Samblanx GW, Osborn RW (1997) Antimicrobial peptides from plants. Crit Rev Plant Sci 16:297–323. doi:10.1080/713608148

Burton RA, Shirley NJ, King BJ, Harvey AJ, Fincher GB (2004) The CesA gene family of barley. Quantitative analysis of transcripts reveals two groups of co-expressed genes. Plant Physiol 134:224–236. doi:10.1104/pp.103.032904

Callebaut I, Labesse G, Durand P, Poupon A, Canard L, Chomilier J, Henrissat B, Mornon JP (1997) Deciphering protein sequence information through hydrophobic cluster analysis (HCA): current status and perspectives. Cell Mol Life Sci 53:621–645. doi:10.1007/s000180050082

Cenci A, Chantret N, Kong X, Gu Y, Anderson OD, Fahima T, Distelfeld A, Dubcovsky J (2003) Construction and characterization of a half million clone BAC library of durum wheat (*Triticum turgidum* ssp. durum). Theor Appl Genet 107:931–939. doi:10.1007/s00122-003-1331-z

Chen L, Madura K (2002) *Rad23* promotes the targeting of proteolytic substrates to the proteasome. Mol Cell Biol 22:4902–4913. doi:10.1128/MCB.22.13.4902-4913.2002

Cruz-Garcia F, Hancock CN, Kim D, McClure B (2005) Stylar glycoproteins bind to S-RNase in vitro. Plant J 42:295–304. doi:10.1111/j.1365-313X.2005.02375.x

Curtis MD, Grossniklaus U (2003) A gateway cloning vector set for high-throughput functional analysis of genes in planta. Plant Physiol 133:462–469. doi:10.1104/pp.103.027979

Diaz I, Martinez M, Isabel-LaMoneda I, Rubio-Somoza I, Carbonero P (2005) The DOF protein, SAD, interacts with GAMYB in plant nuclei and activates transcription of endosperm-specific genes during barley seed development. Plant J 42:652–662. doi:10.1111/j.1365-313X.2005.02402.x

Doan DNP, Linnestad C, Olsen OA (1996) Isolation of molecular markers from the barley endosperm coenocyte and the surrounding nucellus cell layers. Plant Mol Biol 31:877–886. doi:10.1007/BF00019474

Dong J, Kim ST, Lord EM (2005) Plantacyanin plays a role in reproduction in Arabidopsis. Plant Physiol 138:778–789. doi:10.1104/pp.105.063388

Drea S, Leader DJ, Arnold BC, Shaw P, Dolan L, Doonan JH (2005) Systematic spatial analysis of gene expression during wheat caryopsis development. Plant Cell 17:2172–2185. doi:10.1105/tpc.105.034058

Ellerstrom M, Stalberg K, Ezcurra I, Rask L (1996) Functional dissection of a *napin* gene promoter: identification of promoter elements required for embryo and endosperm-specific transcription. Plant Mol Biol 32:1019–1027. doi:10.1007/BF00041385

Ezcurra I, Ellerstrom M, Wycliffe P, Stalberg K, Rask L (1999) Interaction between composite elements in the *napA* promoter: both the B-box ABA-responsive complex and the RY/G complex are necessary for seed-specific expression. Plant Mol Biol 40:699–709. doi:10.1023/A:1006206124512

Fedorova M, van de Mortel J, Matsumoto PA, Cho J, Town CD, VandenBosch KA, Gantt JS, Vance CP (2002) Genome-wide identification of nodule-specific transcripts in the model legume *Medicago truncatula*. Plant Physiol 130:519–537. doi:10.1104/pp.006833

Fujiwara T, Beachy RN (1994) Tissue specific and temporal regulation of beta-conglycinin gene—roles of the RY repeat and other *cis*-acting elements. Plant Mol Biol 24:261–272. doi:10.1007/BF00020166

Funakoshi A, Fukamizu Y, Miyasaka K (1999a) Mechanism of cholecystokinin-A-receptor antagonist on human pancreatic exocrine secretion—localization of CCK-A receptor in the human duodenum. Digestion 60:75–80. doi:10.1159/000051459

Funakoshi M, Geley S, Hunt T, Nishimoto T, Kobayashi H (1999b) Identification of XDRP1; a *Xenopus* protein related to yeast Dsk2p binds to the N-terminus of cyclin A and inhibits its degradation. EMBO J 18:5009–5018. doi:10.1093/emboj/18.18.5009

Funakoshi M, Sasaki T, Nishimoto T, Kobayashi H (2002) Budding yeast Dsk2p is a polyubiquitin-binding protein that can interact with the proteasome. Proc Natl Acad Sci USA 99:745–750. doi:10.1073/pnas.012585199

Gao L, Tu H, Shi ST, Lee KJ, Asanaka M, Hwang SB, Lai MMC (2003) Interaction with a ubiquitin-like protein enhances the ubiquitination and degradation of hepatitis C virus RNA-dependent RNA polymerase. J Virol 77:4149–4159. doi:10.1128/JVI.77.7.4149-4159.2003

Ginalski K, Elofsson A, Fischer D, Rychlewski L (2003) 3D-Jury: a simple approach to improve protein structure predictions. Bioinformatics 19:1015–1018. doi:10.1093/bioinformatics/btg124

Gomez E, Royo J, Guo Y, Thompson R, Hueros G (2002) Establishment of cereal endosperm expression domains: identification and properties of a maize transfer cell-specific transcription factor, ZmMRP-1. Plant Cell 14:599–610. doi:10.1105/tpc.010365

Guex N, Peitsch MC (1997) Swiss-model and the Swiss-PdbViewer: an environment for comparative protein modeling. Electrophoresis 18:2714–2723. doi:10.1002/elps.1150181505

Gutierrez-Marcos JF, Costa LM, Biderre-Petit C, Khbaya B, O'Sullivan DM, Wormald M, Perez P, Dickinson HG (2004) *Maternally expressed gene1* is a novel maize endosperm transfer cell-specific gene with a maternal parent-of-origin pattern of expression. Plant Cell 16:1288–1301. doi:10.1105/tpc.019778

Higo K, Ugawa Y, Iwamoto M, Korenaga T (1999) Plant *cis*-acting regulatory DNA elements (PLACE) database: 1999. Nucleic Acids Res 27:297–300. doi:10.1093/nar/27.1.297

Hiltunen M, Lu A, Thomas AV, Romano DM, Kim M, Jones PB, Xie Z, Kounnas MZ, Wagner SL, Berezovska O, Hyman BT, Tesco G, Bertram L, Tanzi RE (2006) Ubiquilin 1 modulates amyloid precursor protein trafficking and a beta secretion. J Biol Chem 281:32240–32253. doi:10.1074/jbc.M603106200

Hueros G, Varotto S, Salamini F, Thompson RD (1995) Molecular characterization of *Bet1*, a gene expressed in the endosperm transfer cells of maize. Plant Cell 7:747–757

Hueros C, Gomez E, Cheikh N, Edwards J, Weldon M, Salamini F, Thompson RD (1999) Identification of a promoter sequence from the *BETL1* gene cluster able to confer transfer-cell-specific expression in transgenic maize. Plant Physiol 121:1143–1152. doi:10.1104/pp.121.4.1143

Hunt LT, George DG, Yeh LSL (1985) Ragweed allergen RA-3—relationship to some type-1 copper-binding proteins. J Mol Evol 21:126–132. doi:10.1007/BF02100086

John B, Sali A (2003) Comparative protein structure modeling by iterative alignment, model building and model assessment. Nucleic Acids Res 31:3982–3992. doi:10.1093/nar/gkg460

Karplus K, Barrett C, Hughey R (1998) Hidden Markov models for detecting remote protein homologies. Bioinformatics 14:846–856. doi:10.1093/bioinformatics/14.10.846

Kelley LA, MacCallum RM, Sternberg MJE (2000) Enhanced genome annotation using structural profiles in the program 3D-PSSM. J Mol Biol 299:499–520. doi:10.1006/jmbi.2000.3741

Kim S, Mollet JC, Dong J, Zhang KL, Park SY, Lord EM (2003) Chemocyanin, a small basic protein from the lily stigma, induces pollen tube chemotropism. Proc Natl Acad Sci USA 100:16125–16130. doi:10.1073/pnas.2533800100

Kleijnen MF, Shih AH, Zhou PB, Kumar S, Soccio RE, Kedersha NL, Gill G, Howley PM (2000) The hPLIC proteins may provide a link between the ubiquitination machinery and the proteasome. Mol Cell 6:409–419. doi:10.1016/S1097-2765(00)00040-X

Kreps JA, Wu YJ, Chang HS, Zhu T, Wang X, Harper JF (2002) Transcriptome changes for *Arabidopsis* in response to salt, osmotic, and cold stress. Plant Physiol 130:2129–2141. doi:10.1104/pp.008532

Laskowski RA, Macarthur MW, Moss DS, Thornton JM (1993) Procheck—a program to check the stereochemical quality of protein structures. J Appl Cryst 26:283–291. doi:10.1107/S0021889892009944

Li M, Singh R, Bazanova N, Milligan AS, Shirley N, Langridge P, Lopato S (2008) Spatial and temporal expression of endosperm transfer cell-specific promoters in transgenic rice and barley. Plant Biotechnol J 6:465–476. doi:10.1111/j.1467-7652.2008.00333.x

Lopato S, Borisjuk L, Milligan AS, Shirley N, Bazanova N, Parsley K, Langridge P (2006) Systematic identification of factors involved in post-transcriptional processes in wheat grain. Plant Mol Biol 62:637–653. doi:10.1007/s11103-006-9046-6

Mah AL, Perry G, Smith MA, Monteiro MJ (2000) Identification of ubiquilin, a novel presenilin interactor that increases presenilin protein accumulation. J Cell Biol 151:847–862. doi:10.1083/jcb.151.4.847

Maitz M, Santandrea G, Zhang ZY, Lal S, Hannah LC, Salamini F, Thompson RD (2000) *rgf1*, a mutation reducing grain filling in maize through effects on basal endosperm and pedicel development. Plant J 23:29–42. doi:10.1046/j.1365-313x.2000.00747.x

Massey LK, Mah AL, Monteiro MJ (2005) Ubiquilin regulates presenilin endoproteolysis and modulates gamma-secretase components, Pen-2 and nicastrin. Biochem J 391:513–525. doi:10.1042/BJ20050491

Matthews PR, Wang MB, Waterhouse PM, Thornton S, Fieg SJ, Gubler F, Jacobsen JV (2001) Marker gene elimination from transgenic barley, using co-transformation with adjacent 'twin T-DNAs' on a standard *Agrobacterium* transformation vector. Mol Breed 7:195–202. doi:10.1023/A:1011333321893

Mayer BJ (1999) Endocytosis: EH domains lend a hand. Curr Biol 9:R70–R73. doi:10.1016/S0960-9822(99)80014-1

Muniz LM, Royo J, Gomez E, Barrero C, Bergareche D, Hueros G (2006) The maize transfer cell-specific type-A response regulator ZmTCRR-1 appears to be involved in intercellular signalling. Plant J 48:17–27. doi:10.1111/j.1365-313X.2006.02848.x

Murashige T, Skoog F (1962) A revised medium for rapid growth and bio assays with tobacco tissue cultures. Physiol Plant 15:473–481. doi:10.1111/j.1399-3054.1962.tb08052.x

Nicholls A, Sharp KA, Honig B (1991) Protein folding and association—insides from the interfacial and thermodynamic properties of hydrocarbons. Proteins 11:281–296. doi:10.1002/prot.340110407

Olsen OA (2004a) Dynamics of maize aleurone cell formation: the "surface-"rule. Maydica 49:37–40

Olsen OA (2004b) Nuclear endosperm development in cereals and *Arabidopsis thaliana*. Plant Cell 16:S214–S227. doi:10.1105/tpc.017111

Pasquato N, Berni R, Folli C, Folloni S, Cianci M, Pantano S, Helliwell JR, Zanotti G (2006) Crystal structure of peach Pru p 3, the prototypic member of the family of plant non-specific lipid transfer protein pan-allergens. J Mol Biol 356:684–694. doi:10.1016/j.jmb.2005.11.063

Peltier JB, Ripoll DR, Friso G, Rudella A, Cai Y, Ytterberg J, Giacomelli L, Pillardy J, van Wijk KJ (2004) Clp protease complexes from photosynthetic and non-photosynthetic plastids and mitochondria of plants, their predicted three-dimensional structures, and functional implications. J Biol Chem 279:4768–4781. doi:10.1074/jbc.M309212200

Provart NJ, Gil P, Chen WQ, Han B, Chang HS, Wang X, Zhu T (2003) Gene expression phenotypes of *Arabidopsis* associated with sensitivity to low temperatures. Plant Physiol 132:893–906. doi:10.1104/pp.103.021261

Pummi T, Leskela S, Wahlstrom E, Gerth U, Tjalsma H, Hecker M, Sarvas M, Kontinen VP (2002) ClpXP protease regulates the signal peptide cleavage of secretory preproteins in *Bacillus subtilis* with a mechanism distinct from that of the Ecs ABC transporter. J Bacteriol 184:1010–1018. doi:10.1128/jb.184.4.1010-1018.2002

Ramachandran GN, Ramakrishnan C, Sasisekharan V (1963) Stereochemistry of polypeptide chain configurations. J Mol Biol 7:95–99. doi:10.1016/S0022-2836(63)80023-6

Rodrigues JCM, Tucker MR, Johnson SD, Hrmova M, Koltunow AGM (2008) Sexual and apomictic (asexual) seed formation in *Hieracium* requires the plant Polycomb-group gene fertilization independent endosperm (FIE). Plant Cell 20:2372–2386. doi:10.1105/tpc.108.059287

Rost B, Yachdav G, Liu JF (2004) The predictprotein server. Nucleic Acids Res 32:W321–W326. doi:10.1093/nar/gkh377

Ryden LG, Hunt LT (1993) Evolution of protein complexity—the blue copper-containing oxidases and related proteins. J Mol Evol 36:41–66. doi:10.1007/BF02407305

Sali A, Blundell TL (1993) Comperative protein modeling by satisfaction of spatial restrains. J Mol Biol 234:779–815. doi:10.1006/jmbi.1993.1626

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning. A laboratory manual, vol 1, 2 and 3, 2nd edn. Cold Spring Harbor Laboratory Press, New York

Sanchez R, Sali A (1998) Large-scale protein structure modeling of the *Saccharomyces cerevisiae* genome. Proc Natl Acad Sci USA 95:13597–13602. doi:10.1073/pnas.95.23.13597

Sanford JC, Smith FD, Russell JA (1993) Optimizing the biolistic process for different biological applications. Methods Enzymol 217:483–509. doi:10.1016/0076-6879(93)17086-K

Serna A, Maitz M, O'Connell T, Santandrea G, Thevissen K, Tienens K, Hueros G, Faleri C, Cai G, Lottspeich F, Thompson RD (2001) Maize endosperm secretes a novel antifungal protein into adjacent maternal tissue. Plant J 25:687–698. doi:10.1046/j.1365-313x.2001.01004.x

Shen QX, Chen CN, Brands A, Pan SM, Ho THD (2001) The stress- and abscisic acid-induced barley gene *HVA22*: developmental regulation and homologues in diverse organisms. Plant Mol Biol 45:327–340. doi:10.1023/A:1006460231978

Sippl MJ (1993) Recognition of errors in 3-dimentional structures of proteins. Proteins 17:355–362. doi:10.1002/prot.340170404

Sippl MJ, Wiederstein M (2008) A note on difficult structure alignment problems. Bioinformatics 24:426–427. doi:10.1093/bioinformatics/btm622

Sokolenko A, Pojidaeva E, Zinchenko V, Panichkin V, Glaser VM, Herrmann RG, Shestakov SV (2002) The gene complement for proteolysis in the *Cyanobacterium synechocystis* sp. PCC 6803 and *Arabidopsis thaliana* chloroplasts. Curr Genet 41:291–310. doi:10.1007/s00294-002-0309-8

Thompson RD, Hueros G, Becker HA, Maitz M (2001) Development and functions of seed transfer cells. Plant Sci 160:775–783. doi:10.1016/S0168-9452(01)00345-4

Tingay S, McElroy D, Kalla R, Fieg S, Wang MB, Thornton S, Brettell R (1997) *Agrobacterium tumefaciens*-mediated barley transformation. Plant J 11:1369–1376. doi:10.1046/j.1365-313X.1997.11061369.x

Wada S, Hamada M, Satoh N (2006) A genomewide analysis of genes for the heat shock protein 70 chaperone system in the ascidian Ciona intestinalis. Cell Stress Chaperones 11:23–33. doi:10.1379/CSC-137R.1

Wang HL, Offler CE, Patrick JW, Ugalde TD (1994) The cellular pathway of photosynthate transfer in the developing wheat-grain. 1. Delineation of a potential transfer pathway using fluorescent dyes. Plant Cell Environ 17:257–266. doi:10.1111/j.1365-3040.1994.tb00291.x

Wu AL, Wang J, Zheleznyak A, Brown EJ (1999) Ubiquitin-related proteins regulate interaction of vimentin intermediate filaments with the plasma membrane. Mol Cell 4:619–625. doi:10.1016/S1097-2765(00)80212-9

Wu SL, Mikhailov A, Kallo-Hosein H, Hara K, Yonezawa K, Avruch J (2002) Characterization of ubiquilin 1, an mTOR-interacting protein. Biochim Biophys Acta-Mol Cell Res 1542:41–56

Zhang Y (2008) I-TASSER server for protein 3D structure prediction. BMC Bioinform 9:40–47. doi:10.1186/1471-2105-9-40